Molecular mechanisms in the first step of ABA mediated response in Coffea ssp
Michelle Guitton Cotta

To cite this version:

HAL Id: tel-02008137
https://tel.archives-ouvertes.fr/tel-02008137
Submitted on 5 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
MOLECULAR MECHANISMS IN THE FIRST STEP OF ABA-MEDIATED RESPONSE IN Coffea ssp.
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, holidays 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
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’Éthiopie, 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 arômes et saveurs avec des teneurs limitées en caféine. *C. arabica* est une plante allotéraploïde (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. À 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 caficole emploie plusieurs millions de personnes, ce qui souligne son importance économique et sociale dans les pays de la zone intertropicale ou les cafériers 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 60kg par an).

Comme de nombreuses grandes productions végétales, le caférier 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 cafériers 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 *Cc:DREB1D*) 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 \textit{PYR/PYL/RCAR-SnRK2-PP2C} chez \textit{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:
  o les différents tissus et organes de \textit{C. canephora} et de \textit{C. arabica}?
  o les racines et les feuilles des clones tolérants et sensibles à la sécheresse de \textit{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 \textit{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 \textit{PYR/PYL/RCAR-SnRK2-PP2C} de \textit{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 \textit{C. canephora} cultivés en serre en condition de stress hydrique,
5. de comparer les profils d’expression obtenus \textit{in vivo} à ceux obtenus \textit{in silico} pour cette même espèce,

**Principaux résultats**

**Identification des gènes \textit{PYR/PYL/RCAR-SnRK2-PP2C} du système tripartite chez \textit{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 \textit{C. canephora} codant pour les protéines du système tripartite \textit{PYR/PYL/RCAR-SnRK2-PP2C}. Ainsi, les séquences de ces protéines préalablement identifiées chez \textit{Arabidopsis thaliana}, mais également chez l’oranger (\textit{Citrus sinensis}), le riz asiatique (\textit{Oryza sativa}), la vigne (\textit{Vitis vinifera}), la tomate (\textit{Solanum lycopersicum}) et la pomme de terre (\textit{Solanum tuberosum}) ont servi de séquences de référence pour rechercher les séquences orthologues de \textit{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 \textit{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 (DT : 14, 73 et 120) et sensible (DS : 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 peu 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 cafiers 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 DT 14 et 73 de *C. canephora*. A l’inverse, les
conditions de sécheresse répriment fortement l’expression du gène \textit{CcPYL4} dans ces deux organes et quelques soient les clones.

- **Famille des gènes \textit{PP2C}**

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

  Nos résultats montrent très clairement une augmentation de l’expression des gènes \textit{CcABI2}, \textit{CcAHG3} et \textit{CcHAI} en condition de sécheresse dans les feuilles et les racines des clones D\textsuperscript{T} et D\textsuperscript{S} clones de \textit{C. canephora}. La sécheresse induit également l’expression des gènes \textit{CcAHG2} et \textit{CcHAB} mais seulement dans les feuilles des clones tolérants D\textsuperscript{T} 14, 73 et 120. Concernant le gène \textit{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 \textit{CcABI1}, \textit{CcABI2} et \textit{CcAHG3} augmente spécifiquement en condition de sécheresse. Enfin, d’un point de vue quantitatif, \textit{CcHAI} est le gène le plus fortement surexprimé en condition de sécheresse dans les feuilles et dans les racines.

- **Famille des gènes \textit{SnRK2}**

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

  Dans les feuilles, le gène \textit{CcSnRK2.2} est le seul qui présente une augmentation significative de son expression en condition de sécheresse pour les clones tolérants (D\textsuperscript{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\textsuperscript{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 \textit{CcSnRK2.2}, \textit{CcSnRK2.6} et \textit{CcSnRK2.7}. A l’inverse, la sécheresse diminue l’expression du gène \textit{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 \textit{CcSnRK2.11} s’exprime dans les feuilles mais pas dans les racines. Enfin, aucune expression des gènes \textit{CcSnRK2.1}, \textit{CcSnRK2.12} et \textit{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 cafétier, 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. À 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 CO₂ (ETR/A : electron transport rate/CO₂ assimilation rate) en condition de sécheresse ratio que les autres clones tolérants (D₁4 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 $\Psi_{pd}$ 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 pourraient être utilisées dans les programmes de sélection des cafésiers pour la création de nouvelles variétés plus tolérantes à la sécheresse.
Figures List

Review of related literature

Figure 1 Coffee trade statistic for the last four years of crop production .......................................................... 28
Figure 2 Global coffee consumption. During last years the demand increase in many countries including traditional markets, exporting counties and emerging markets .......................................................... 30
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 .......................................................... 32
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) .......................................................... 33
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 .......................................................... 35
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 .......................................................... 36
Figure 7 Conceptual diagram of the stress tolerance/stress avoidance model of low-$\Psi_w$ responses. .......................................................... 38
Figure 8 C. canephora clones (A: Drought tolerant, DT; B: Drought susceptible, DS) grown in greenhouse and submitted to drought conditions .......................................................... 39
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) .......................................................... 40
Figure 10 Predawn leaf water potential ($\Psi_{pd}$) 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. $\Psi_{pd}$ values (expressed in mega-Pascal, MPa) were measured once a week during the 2009 dry season (23-month-old plants) .......................................................... 41
Figure 11 Typical root systems of four clones of Robusta coffee grown under full irrigation .......................................................... 41
Figure 12 The evolution of predawn leaf water potential ($\Psi_{pd}$) 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, $\Psi_{pd}$ evolutions are presented .......................................................... 42
Figure 13 The productivity (measured in liters of cherries per plant) and $\Psi_{pd}$ 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)

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.

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.

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).

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.

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.

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.

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.

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. ................................................................. 55

Figure 22 A schematic representation of the group A PP2C, AtABI1 and the SnRK2, AtOST1. AtABI1 consists of a PP2C (catalytic) domain (brown) in addition to the 11 motifs (green) (Bork et al., 2006) at its C-terminal. AtOST1 consists 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. ................................................................. 55

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. ................................................................. 57

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.................. 58

Chapter 1

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 ........................................... 98

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-APP2C (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 ≤ 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.

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. ........................................ ................................................... .................................................. 106

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. 107

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. ........................................ ................................................... .................................................. 108

Chapter 2

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 .............................................. .................. 120

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) ................................................... ................................... 122

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) ........................................................................ 123
Figure 4 Expression profiles of SnRK2 genes in leaves of *C. canephora DT* (clone 14) and DS (clone 22) and *C. arabica DT* (I59) and DS (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) ................... 125

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 DT* 14 (circles) and DS 22 clones (triangles) and *C. arabica DT* I59 (squares) and DS 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 .......................................................... 127
Tables List

Review of related literature

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

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

Chapter 1

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.

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 and Gene IDs in Phytozome 10.3.

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 and Gene IDs in Phytozome 10.3.

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 and Gene IDs in Phytozome 10.3.
Summary

Preface .................................................................................................................................................. 23
Reference ............................................................................................................................................ 25
1 Review of Related Literature ........................................................................................................... 28
  1.1 International market ..................................................................................................................... 28
  1.2 Global Climate Change: impacts in coffee production ................................................................. 31
    1.2.1 Impacts in coffee areas .......................................................................................................... 31
    1.2.2 Impact in term of abiotic stress .............................................................................................. 33
    1.2.3 Impact on coffee plants .......................................................................................................... 34
  1.3 Coffea genus .................................................................................................................................. 34
    1.3.1 Coffea arabica ......................................................................................................................... 35
    1.3.2 Coffea canephora ..................................................................................................................... 35
    1.3.3 Other Coffea species ............................................................................................................... 36
  1.4 Drought responses in plants ........................................................................................................... 37
    1.4.1 Coffee genetic diversity and drought ....................................................................................... 39
    1.4.2 Physiological responses .......................................................................................................... 41
    1.4.3 Biochemical responses .......................................................................................................... 42
    1.4.4 Molecular responses ............................................................................................................. 43
  1.5 ABA structure and biological roles ............................................................................................... 44
    1.5.1 ABA biosynthesis, catabolism, conjugation and transport .................................................... 45
  1.6 The PYL/PP2C/SnRK2: the first steps of ABA sensing and signaling ............................................. 48
    1.7 Evolution of ABA sensing and signaling ....................................................................................... 49
      1.7.1 The tripartite system: PYL-PP2C-SnRK complex .................................................................. 51
      1.7.2 PYR-PYL/RCARs: ABA receptors: ....................................................................................... 52
      1.7.3 PP2Cs phosphatases ............................................................................................................. 54
  1.8 SnRK2 kinases ............................................................................................................................... 56
REFERENCES .......................................................................................................................................... 59
PRESENTATION OF THE PHD PROJECT ............................................................................................ 71
CHAPTER 1 ............................................................................................................................................ 72
Molecular mechanisms of ABA-mediated response to drought in leaves and roots of Coffea canephora ................................................................................................................................. 74
Abstract ................................................................................................................................................ 75
INTRODUCTION ..................................................................................................................................... 76
MATERIAL AND METHODS .................................................................................................................. 77
RESULTS ................................................................................................................................................. 80
Gene expression profiles in *Coffea arabica* and *Coffea canephora* leaves revealed transcriptional regulations of key genes involved in ABA signaling.
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). Variations 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 identify 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 intracellular 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 Snf1-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.


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


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 a 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.](image)

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.

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

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 year. 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).
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.

<table>
<thead>
<tr>
<th>Calendar year</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exporting countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>20.178</td>
<td>20.146</td>
<td>20.271</td>
<td>20.458</td>
</tr>
<tr>
<td>Indonesia</td>
<td>3.842</td>
<td>4.100</td>
<td>4.292</td>
<td>4.458</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>3.387</td>
<td>3.463</td>
<td>3.656</td>
<td>3.681</td>
</tr>
<tr>
<td>Importing countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>European Union</td>
<td>41.018</td>
<td>41.875</td>
<td>42.215</td>
<td>41.638</td>
</tr>
<tr>
<td>USA</td>
<td>22.232</td>
<td>23.417</td>
<td>23.767</td>
<td>24.441</td>
</tr>
<tr>
<td>Japan</td>
<td>7.131</td>
<td>7.435</td>
<td>7.494</td>
<td>7.695</td>
</tr>
<tr>
<td>TOTAL</td>
<td>143.430</td>
<td>147.811</td>
<td>150.389</td>
<td>152.204</td>
</tr>
</tbody>
</table>

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.
5(9, 2) 5$/7'(17$785(32

SV D FRQVHTXHQFH RI JOREDO ZDUPLQJ, FRIIHH-JURZLQJ JHRJUDSKLFDO UHILRQV FRXOG DOVR VXIIIHU LPSRUWDQW JHRJUDSKLFDO GHOFRDOL]DWLRQ (SVVDG $QVDHG $QVDHG, 2004). Q PDULQDO UHJLRQV ZLWKRFXS RI VXULIQI GU/ VHDVRQW, WKLV FRXOG OHG LQ GHUHVHFLQJ FRIIHH ILOHGV DV PXFK DV 80% (2OQDDD $5DOKR, 2006). SQDOJLQI WUHFRPHW RI UHFRQW FOLFDWLRQ FKDQJHV E[WUDSRODZDUPLQJ WUHFRQW WKH KLWWRULFDO WHQGHQFLQLQJ WKH WPSHUDWXUH DQG SUHFLSLWDWLRQ WR 2010 LQ FRIIHH SURGXFLQJ DUHDV LQ 9HUDFUX]. QHLFR, WKH DQDO\]LQJ EHJDVLQJ SUHGLFW WKDW FRIIHH SURGXFLQJ DUHDV LQ 9HUDFUX]. QHLFR, WKH DQDO\]LQJ EHJDVLQJ SUHGLFW WKDW FRIIHH SURGXFLQJ DUHDV LQ 9HUDFUX]. QHLFR, WKH DQDO\]LQJ EHJDVLQJ SUHGLFW WKDW FRIIHH SURGXFLQJ DUHDV LQ 9HUDFUX]. QHLFR, WKH DQDO\]LQJ EHJDVLQJ SUHGLFW WKDW FRIIHH SURGXFLQJ DUHDV LQ 9HUDFUX]. QHLFR, WKH DQDO\]LQJ EHJDVLQJ SUHGLFW WKDW FRIIHH SURGXFLQJ DUHDV LQ 9HUDFUX]
6mR 3DXOR. WKH SRWHQWLDO DUHD IRU SURGXFWRQ ZRXOG GHFOLQH IURP 70-75% RI WKH VVDWHV WR 20-25%. ZKLOH FRIIHH DUHD ZRXOG EH UHGXFHG \& 10% LQ 3DUDQG DQG SURGXFWRQ ZRXOG EH HOLPLQDWHQ LQ \*RLDV VVDWH (LJXUH 4). 7KH QHZ DUHDV VXLWDEOH IRU FRIIHH SURGXFWRQ WKDW FRXOG HPHUJH LQ 6DQWD &DWDULQD DQG SLR *UDQGH GR 6XO ZLQO RQQO SDUWLQDOI FRPSHQVYVWKH ORV RI DUHD LQ RWWKHU VVDWHV (3LQWR & SVVDG, 2008).

7KHVH IRUWKFRPLQJ VFHQDULRV UHTXLUH QHZ DSSURDFKHV WKDW GHYHORS LQQRWDWLYH WVUDWHLHV WR PDQJH WKH FURS SURGXFWRQ VVWHP DQG UHGXFH WKH LPSDFW RI FOLPDWFK FDQJH. 6WUDWHLHV VXFK DV WKH GHYHORS LQJ DIURUWHV\& SURGXFWRQ VVWHP, LQFUHSDV LULJDLWRQ, DQG JUHHQ, DULFXOWXUDO SUDFWLFHV PDQJHDLQLQJ FKDV RI UHFRPH WR ERUHU KDV EHHQ GHPRQV WUDWHG (-DUDPLOOR, 2009).

1.2.2. PSDFW LQ WHUP RI DELRWLF VWUHV

8QGHU GURXIKW DQG KLJK WHPSHUDWUXVHV, VRPH FRIIHH SHVWV DQG GLVHDVHV VKRXOG DOVR EHRFRPH PRUH VHYHUH. 7KH RFFXUHQFH RI OHDI PLQHUV (AUXFRWSWH UHFRQH) GLVHDVHV KDQ EHQQH LQFUHDLQ LQ RYHU UHHQW \*UDUV LQ FRIIHH \& UHOGV DV D FRQVXHG FRQGLWRQV (SVLVL HV DW DO., 2012). 2Q WKH RWKHU KDQG, OHDI UXWV GLVHDVHV LV ULYQJ ZLWK ZDUPHU WHPSHUWUXVHV. /S)DQWHU, WKH QSPEHU RI FIJOH OLIXV HQQHULQJ RQV RI +SKRWHPXV KDPSLV KDQ EHHQ LQFUHDLQ LQ XQGHU WKH VDPH FOLPDWLF FRQGLWRQV, DV D UHVVXOW, D WKHUPDO WROHIDQFH RI WKH FRIIHH EHUU ERUHU KDQ EHHQ GHPRQWUDWHG (-DUDPLOOR HV DW DO., 2009). Q WKH FDVH RI
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 $2.62 \times 10^3$ 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.](image)

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). \textit{C. racemosa} had the longest root system in comparative analyses with other coffee species (\textit{C. canephora}, \textit{C. arabica}, \textit{C. liberica} and \textit{C. congensis}) and the root system is mainly contrasting with \textit{C. congensis} root system which survived in a natural environment completely different of \textit{C. racemosa} (Dublin, 1968).

Medina Filho \textit{et al.} (1977b) had evaluated the genetic material of \textit{C. racemosa} from Campinas (Brazil), and they verify that triploids (\textit{C. arabica} x \textit{C. racemosa}) as well as individuals belonging to the second generation backcrosses to \textit{C. arabica} were highly resistant to drought, while Catuai and Acaia cultivars of \textit{C. arabica} (positive controls of the experiment), were highly sensitive. While these cultivars lose a lot of leaves the plants which derivate of \textit{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 \textit{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 \textit{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 \textit{et al.}, 2006) (Figure 7).
In most cases, the plant first response is avoid low $\Psi_w$. Tissue $\Psi_w$ 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 $\Psi_w$ 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 $\Psi_w$ of the plant will equilibrate with that the water source (most of cases $\Psi_w$ of the soil). When soil water content and $\Psi_w$ are low, the $\Psi_w$ of the plant tissue must also decrease, either through water loss or by adjustment made by the plant (dehydration avoidance) to achieve a low $\Psi_w$ while avoiding a water loss. The main mechanism of dehydration avoidance are accumulation of solutes and cell wall hardening (Verslues et al., 2006).

The $\Psi_w$ of a walled cell, such as a plant cell, is governed by the equation: $\Psi_w = \Psi_s + \Psi_p$, where $\Psi_s$ is the osmotic potential and $\Psi_p$ is the pressure potential (turgor pressure). At a given $\Psi_w$, a higher $\Psi_p$
can be achieved by accumulating solutes inside the cell, thus lowering $\Psi_s$. The accumulation of additional solutes in response to low $\Psi_w$ is termed osmotic adjustment (Zhang et al., 1999). Osmotic adjustment refers to the active accumulation of additional solutes in response to low $\Psi_w$ (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 $\Psi_w$ (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 $\Psi_w$) is the phytohormone abscisic acid (ABA). It accumulates in response to abiotic stress and regulates the processes involved several the aspects of the low-$\Psi_w$ 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 $\Psi_w$. 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-$\Psi_w$ 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, DT; B: Drought susceptible, DS) grown in greenhouse and submitted to drought conditions.

5?DUGLQJ &RIIHD DUDELFD, WKH VWXG\ RI SRSXODWLRQV IURP (WKLRSLD JURZLQI XQGHU FRQWUDVWVLQJ FOLPDWLF FRQGLWLRQV QRW DOVR UHYHDOHG WKDW WKLV VSDFH RI JURXS RI &. FDQHSKUUDQW YDULQG VRLO PRLVWXUH FRQGLWLRQV 2007), LV ZKLFK WKH UHVSRQHV WKDW WKHV H[QHDFWH UHYHDOHG WKDW WKLV SHULRG RI DURXQG 200 GD\V ZLWKRXW UDLQIDOOV (PEUDSD &HUUDGRV).


LQWHUPEHUV RWKHU WKH SUURQHFWLQJ YDOXH FDUHG HDUWLDOO\ WR WKH 6*1 JURXS RI &. FDQHSKUUDQW YDULQG &RQLODQW KDYH EHHQ FKDUDFWHUL]HG DV YLJRURXV SODQWY ZLWK KLJK SURGXFWLYLW\ WKURXJKRXW UDUV XQGHU GURXJKW WUHVHV (HUUmR, 2000; RQVHFD, 2004). LQIHUSULQW DQDO\HV DOVR UHYHDOHG WKDW WKHVH &RQLODQW FORQHV EHRQJ WR WKH 6*1 JURXS RI &. FDQHSKUUDQW (DPERW HW DO., 2008; ORQW DJQHWL]HG HW DO., 2012).

5?DUGLQJ &RIIHD DUDELFD, WKH VWXG\ RI SRSXODWLRQV IURP (WKLRSLD JURZLQI XQGHU FRQWUDVWVLQJ FOLPDWLF FRQGLWLRQV QRW DOVR UHYHDOHG WKDW WKLV VSDFH RI JURXS RI &. FDQHSKUUDQW YDULQG VRLO PRLVWXUH FRQGLWLRQV 2007), LV ZKLFK WKH UHVSRQHV WKDW WKHV H[QHDFWH UHYHDOHG WKDW WKLV SHULRG RI DURXQG 200 GD\V ZLWKRXW UDLQIDOOV (PEUDSD &HUUDGRV).

0DMRU GLIIHUKHFWH EHWWHU WKHVH WZLQJ SXEL, DQG WKH LQGLFDWLRQ EHHQ WKH GHU VHUW RQH 2008 DQG 2009. LV ZKLFK 10, RU ZLWKRXW 1 (1) LQGLFDWLRQ DQG WKH EHHQ WKH EHHQ YDOXH LWV XQGHU FRQWUDVWLFK &RQLODQW DOVR UHYHDOHG WKDW WKHVH WROHUDQW FORQHV EHRQJ WR WKH 6*1 JURXS RI &. FDQHSKUUDQW (DPERW HW DO., 2008; ORQW DJQHWL]HG HW DO., 2012.)


0DMRU GLIIHUKHFWH EHWWHU WKHVH WZLQJ SXEL, DQG WKH LQGLFDWLRQ EHHQ WKH GHU VHUW RQH 2008 DQG 2009. LV ZKLFK 10, RU ZLWKRXW 1 (1) LQGLFDWLRQ DQG WKH EHHQ WKH EHHQ YDOXH LWV XQGHU FRQWUDVWLFK &RQLODQW DOVR UHYHDOHG WKDW WKHVH WROHUDQW FORQHV EHRQJ WR WKH 6*1 JURXS RI &. FDQHSKUUDQW (DPERW HW DO., 2008; ORQW DJQHWL]HG HW DO., 2012.)
1.4.2 3KVLORJLFDO UHVSRQVHV

7KH RXLOORX (6*1) JURXS RI & FDQHSKRUD DSSHDUV WR EH PRUH WROHUDQW WR ZDWHU GHI'LFLW WKDQ 5REXVWD (6*2) (ORQWDJIRQ & /HUR\, 1993). 6*1 JURXS PDLQWDLQ VWRPDWDO RSHQLQJ DOQG FRQVHTXHQWO\ DFWLHY SKRWRVQWKLHVLV, ZKLOH VWRPDWD RI 6*2 SODQWV ZHUB FRPSOHWHO\ FORVHG XQGHU GURXJKW FRQGLWLRQV.

%HVLGHV WKDW, PRUH HIILFLHQW URRW ZDWHU DEVRUSWLQRIU WRK HKH 6*1 SODQWV FRXOG H[SODLQ LWV GURXJKW WROHUDQH DOEHLW LWV PDLQWHDQFH RI VWRPDWDO RSHQLQJ (%R\HU, 1969). 3KVLORJLFDO DQDO\VHV DOVR VXJH \XGHLW \KDOQG \EU FRXOG EH D GLUHFW FRQVHTXHQFH RI EHWJXU URRW GHYHORSPHQW (3LQKHLUR HW \DO\, 2005) (LJXUH 11).

2QH RI WKH SKVLORJLFDO SDUDPHWHUV WKDW GLVWLQJLQVK WKH GURXJKW-VXVFHSWLEOH (*) FORQH 22 RI & FDQHSKRUD YDU. &RQLORQ IURP WKH 7 FORQHV 14, 73 DOQG 120 LV WKH UDHV RI GHIUHVH LQ WKH SUHGDZQ OHDI

Drought-tolerant clones
Cloned 14
Cloned 120

Drought-sensitive clones
Cloned 46
Cloned 109A

LJXUH 10 3UHGDLQ OHDID ZDWHU SROLQWLDO (Ω(N) PHDVXUHG LQ SODQWV RI & DUDELF, SXL (58%, WULDQIOH) DOQG S3S5 (59, VTXDUH) FOWLYDUV ZHUB JURXQ XQGHU FRQWURO (&, RSHQ VPEROV) DOQG GURXJKW (, EODFN VPEROV) FRQGLWLRQV. (Ω(N) YDOXHV (H[SUHVVHG LQ PHID-3VDFO, 003) ZHUB PHDVXUHG RFQH D ZHHN GUXULQJ WKH 2009 GUL VHDFV (23-PWQ-WK-ROP SODQWV).

6RXUFH: O\R\DWR HWD\DO\, 2016.
water potential ($\Psi_{pd}$) (RDPW) (Pinheiro et al., 2004). To reach the imposed $\Psi_{pd}$ 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 $\Psi_{pd}$ 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 ($\Psi_{pd}$) 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, $\Psi_{pd}$ evolutions are presented. Source: Marraccini et al., 2011.](image)

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$_2$O$_2$ 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 ($\Psi_{pd}$) 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 DT I59 compared with the DS 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 $\Psi_{pd}$ 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).
1.5 $5%$ VWUXFWXUH DQG ELROJLFDQ UROHV

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 compartmentation 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 compartment 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 cassette (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).

![ABA metabolic pathways](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).
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).

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 ubiquitous in plants and it is also produced by some phytopathogenic fungi, bacteria and metazoans ranging from sea sponges 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).

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 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.
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 ARABIDOPSIS THALIANA 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 (Mosqueta et al., 2011) suggesting that a single receptor modified is sufficient to activate this
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 \textit{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).

\subsection*{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 \textit{A. thaliana} and rice contained for example 80 and 78 PP2C genes, respectively (Xue et al., 2008). Phylogenetic analyses from \textit{Arabidopsis} and soybean were supported by gene structure and protein motifs and led to subdivide the PP2C genes (Figure 21).
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 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.

---

**Figure 22** A schematic representation of the group A PP2C, AtABI1 and the SnRK2, AtOST1. AtABI1 consists of a PP2C (catalytic) domain (brown) in addition to the 11 motifs (green) (Bork et al., 2006) at its C-terminal. AtOST1 consists 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).
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.](image)

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


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


REVIEW OF RELATED LITERATURE


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 Cotta¹,², Érica Cristina da Silva Rêgo³, Stéphanie Sidibe-Bocs⁴, Tatiana Santos Costa³, Fernanda de Araújo Carneiro¹, Jean-François Dufayard⁴, Dominique This², Pierre Marraccini³,⁴, Alan Carvalho Andrade³,*#

¹Departamento de Química, Laboratório Central de Biologia Molecular (LCBM), UFLA, Lavras, Minas Gerais, Brazil
²Montpellier SupAgro, UMR AGAP, F-34398 Montpellier, France.
³Embrapa Recursos Genéticos e Biotecnologia (LGM), Parque EB, CP 02372, 70770-917 Brasília, DF, Brazil.
⁴CIRAD, 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/](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 (DT14, 73 and 120) and -susceptible (D522) *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 DT clones, suggesting that they may constitute key-genes for drought tolerance in these clones. The DT 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 (DuMattta 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 ancestors (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; DaMattta 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\textsuperscript{T} and D\textsuperscript{S} clones of \textit{C. canephora} (Marraccini \textit{et al.}, 2011, 2012; Vieira \textit{et al.}, 2013), some of them (e.g. \textit{RD29} and \textit{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 \textit{et al.}, 2014). Under stress conditions, this phytohormone, synthesized in roots and leaves during periods of water depletion (Thompson \textit{et al.}, 2007), is perceived by PYR/PYL/RCAR receptors that are the first component of the ABA tripartite systems (Klingler \textit{et al.}, 2010). Once formed, the PYL-ABA complex bind to the clade A phosphatase type 2C (PP2C) inactivating them (Hao \textit{et al.}, 2011; Ma \textit{et al.}, 2009; Park \textit{et al.}, 2009). Then, the subclass III SNF1-related kinase (SnRK2) proteins are activated by dephosphorylation allowing expression of downstream stress responsive genes (Cutler \textit{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 \textit{C. canephora} (Denoeud \textit{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 \textit{C. canephora}, (ii) to characterize these orthologs according to gene structure, protein functional domains, phylogeny, syntenic and (iii) to evaluate the expression profile of those genes in leaves and roots of contrasting (D\textsuperscript{T} and D\textsuperscript{S}) clones \textit{C. canephora} subjected or not to drought conditions.

**MATERIAL AND METHODS**

\textit{Plant material}

Drought-tolerant (D\textsuperscript{T}: 14, 73 and 120) and -susceptible (D\textsuperscript{S}: 22) clones of \textit{C. canephora} Conilon were grown in greenhouse conditions (under controlled temperature 25\textdegree C, relative humidity of 70\% and photosynthetic flux PPF 900 \(\mu\text{mol} \text{s}^{-1}\)) 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 (\(\Psi_{pd}\)) of around -3.0 MPa. From each clone, biological triplicate samples...
CHAPTER 1: THE PYR/PP2C/SnRK2 TRIPARTITE SYSTEM IN C. CANEPHORA

(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/](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
CHAPTER 1: THE PYL/PP2C/SnRK2 TRIPARTITE SYSTEM IN *C. canephora*

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](http://www.ncbi.nlm.nih.gov/sites/entrez?db=taxonomy)). The HMM (Hidden 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/](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/](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 2.4 µg 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 1 µl of diluted single-stranded cDNA and 0.2 µM (final concentration) of each primer in a final volume of 10 µl 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 \((1+E)^{\Delta\Delta C_t}\), where \(\Delta C_t = C_{\text{target gene}} - C_{\text{reference gene}}\) and \(\Delta\Delta C_t = \Delta C_t - \Delta C_{\text{internal calibrator}}\).

**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 *CcABII* 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 \textit{CcPYL7a}, \textit{CcPYL7b}, \textit{CcSnRK2.1}, \textit{CcSnRK2.12-13} coding-genes located on chr0 were not showed on the resumèd phylo-analyses, however, they are represented in the complete ones (Fig. S2 and S4).

\textit{ABA (PYR/PYL/RCAR) receptors}

Among the nine PYR/PYL/RCAR proteins, \textit{CcPYR1}, \textit{CcPYL8a} and \textit{CcPYL9} showed high sequence identity (72\%, 83\% and 84\%) with tomato sequences while \textit{CcPYL2} and \textit{CcPYL4} shared 84\% and 74\% of identity with the potato proteins, and \textit{CcPYL7a} and \textit{CcPYL7b} had 54\% and 53\% of identity with the same grape locus while the \textit{CcPYL8a} and \textit{CcPYL8b} proteins shared 82\% of identity. Finally, the \textit{CcPYL13} showed 62\% of sequence identity with grape GSVIVG01013161001 protein. BLASTP results showed that \textit{CcPYL4}, \textit{CcPYL7a} and \textit{CcPYL7b} proteins were highly homologous to \textit{AtPYL6} from \textit{A. thaliana}, \textit{CsPYL5} from \textit{C. sinensis} and \textit{VvRCAR6} from \textit{V. vinifera}, respectively (Table S2).

The \textit{CcPYL4}, \textit{CcPYL7a} and \textit{CcPYL7b} proteins also shared high identity respectively with the Solyc10g076410, Solyc10g085310 and Solyc03g095780 proteins of \textit{S. lycopersicum}. All these coffee PYR/PYL/RCAR proteins (including in \textit{CcPYL7a} and \textit{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 \textit{CcPYL} genes were identified on different ancestral blocks of the seven eudicot chromosomes such as the G2 (\textit{CcPYR1}, \textit{CcPYL8a}, \textit{CcPYL8b} and \textit{CcPYL9}), G4 (\textit{CcPYL2}), G6 (\textit{CcPYL4}) or G7 (\textit{CcPYL13}) groups. All \textit{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 \textit{CcPYR1} and \textit{CcPYL8b} genes also belonged to the same paralogous region (Fig. 2a). A different situation was observed for the \textit{CcPYL8a} and \textit{CcPYL9} genes that derived from different paralogous regions.

Manual curation of \textit{CcPYL} genes revealed that \textit{CcPYR1}, \textit{CcPYL2}, \textit{CcPYL7a}, \textit{CcPYL7b} and \textit{CcPYL13} did not contain introns, while one intron was found in \textit{CcPYL4} and two in \textit{CcPYL8a}, \textit{CcPYL8b} and \textit{CcPYL9} genes (Fig. 2b). No evidence of 5’ UTRs regions was found for \textit{CcPYL2}, \textit{CcPYL7a} and \textit{CcPYL7b} genes. \textit{CcPYL8a} was the only gene presenting an intron (of 316 bp length) on the 5’ UTR
region. For \textit{CcPYL8b}, a 3' UTR extension was based on sequence alignments with a corresponding EST of \textit{C. arabica} (GR997267) expressed in leaf, fruit, flower, root and calli tissues. The \textit{CcPYL8b} gene was also extended in its 5' UTR using the similarities found in \textit{PYL9} genes of tomato (LOC101258886) and potato (LOC102591194) (Table S2). In the same way, an extension was also found in \textit{CcPYL13} 5' UTR region based on the GT013431 EST sequence of \textit{C. arabica} expressed in fruits. Phylogenetic analyses revealed that the ABA receptors \textit{CcPYL8a}, \textit{CcPYL8b} and \textit{CcPYL9} belong to the subfamily I together with \textit{AtPYL7-10} from \textit{A. thaliana}, \textit{CsPYL8-9} from \textit{C. sinensis} and \textit{VvPYL8-9} from \textit{V. vinefera} (Fig. 3a, Fig. S2). The \textit{CcPYL4} protein was located in the subfamily II together with \textit{AtPYL4-6} and \textit{CsPYL4-5} while \textit{CcPYL13} was closely related to \textit{AtPYL11-13}. Finally, the subfamily III contained the \textit{CcPYR1} and \textit{CcPYL2} proteins, the first being related to the \textit{AtPYR1}, \textit{AtPYL1} and \textit{CsPYR1} proteins, and the second to \textit{AtPYL2-3} and \textit{CsPYL2} proteins.

\textbf{Phosphatase type 2C (PP2C) proteins}

The majority of coffee PP2Cs were identical to phosphatases proteins from \textit{Solanaceae} (Table S3). Among them, the \textit{CcABI1}, \textit{CcABI2}, \textit{CcHAB} and \textit{CcHAI} were highly similar to potato proteins while \textit{CcABI2}, \textit{CcAHG3} and \textit{CcHAB} were related to tomato sequences. On the other hand, the \textit{CcAHG2} protein presented 57% of identity with a grape sequence. The catalytic domain of PP2Cs composed of 11 conserved motifs with \textit{Mg^{2+}/Mn^{2+}} [\textit{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 (\textit{CcAHG2}), G4 (\textit{CcHAB}, \textit{CcABI1}, \textit{CcABI2}) and G6 (\textit{CcHAI} and \textit{CcAHG3}) ancestral blocks (Fig. 2a). Even thought \textit{CcHAB}, \textit{CcABI1} and \textit{CcABI2} in one hand, and \textit{CcHAI} and \textit{CcAHG3} in another, evolved from the same ancestral block genes, all these genes belonged to a different paralogous region. \textit{CcAHG2} and \textit{CcAHG3} contained three introns while four were observed in \textit{CcABI1}, \textit{CcABI2}, \textit{CcHAI} and \textit{CcHAB} genes (Fig. 2b). Only \textit{CcABI1} and \textit{CcABI2} genes contained introns in their corresponding 5' UTR regions. The phylogenetic analyses revealed that \textit{CcABI1-2} and \textit{CcHAB} proteins evolved together with \textit{AtABI1-2}, \textit{AtHAB1-2}, \textit{CsABI1} and \textit{CsHAB1} while \textit{CcAHG2}, \textit{CcAHG3} and \textit{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 DT and DS 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 DT clones 14 and 120. Expression levels of CcPYL8a gene did not changed significantly from I to NI conditions in leaves of DT clones 14, 73 and 120 but decreased significantly under drought in DS 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 DT 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 DT clone 73 and DS clone 22, while the increase was much more reduced in DT 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 DT 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 DT 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 \( \text{CcAHG2} \) was expressed in drought-stressed leaves of \( \text{D}^\text{T} \) clones but not in those of \( \text{D}^\text{S} \) clone 22, highly suggests a key function of this gene in leaves of \( \text{C. canephora} \) \( \text{D}^\text{T} \) clones submitted drought.

**SnRK2 gene family in \( \text{C. canephora} \)**

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

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

In contrast to subgroup III, the main targets of subgroup II \( \text{SnRK2s} \) are stress-responsive genes coding transcription factors (Kulik et al., 2011). For example, Zhang et al. (2010) showed that over-expression of wheat \( \text{SnRK2.8} \) 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, \( \text{A. thaliana} \) over-expressing \( \text{SnRK2.7} \) from wheat showed enhanced photosystem II activity and root growth (Zhang et al., 2011). Even though \( \text{SnRK2.7} \) and \( \text{SnRK2.8} \) might be functionally redundant, \( \text{SnRK2.7} \) was shown to be expressed in roots, leaves and flowers of Arabidopsis while \( \text{SnRK2.8} \) was mainly expressed in roots, indicating different tissue specificities of these two kinases (Mizoguchi et al.,...
2010). The up-regulated expression of \textit{CcSnRK2.7} and \textit{CcSnRK2.8} in roots of drought-stressed \textit{C. canephora}, led us to propose key functions of both kinases in coffee roots.

\textit{How the tripartite system PYL-PP2C-SNRK2 of ABA perception could explain DT and DS phenotypes of \textit{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 DT clone 120 was the only one presenting significant increase of leaf ABA content under drought. Whatever the \textit{C. canephora} clone, no significant differences of ABA contents were observed in roots, therefore indicating that DT and DS phenotypes were probably due to altered ABA signalling pathway rather than deficiencies of ABA synthesis.

Previous studies revealed that transport rate/\text{CO}_2 assimilation (ETR/A) ratio was significantly higher under drought in DT clone 73 compared to DT 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 (\textit{CcAPX1}), a prephenate-dehydrogenase like protein (\textit{CcPDH1}) and a non-symbiotic haemoglobin (\textit{CcNSH1}) was already reported in this clone, suggesting its protection involved strong induction of antioxidant and osmoprotection systems (Vieira \textit{et al.}, 2013). The up-regulated expression of \textit{SnRK2.2}, \textit{SnRK2.7} and \textit{SnRK2.8} upon drought in its leaves could participate in activating such pathways.

Another interesting result concerned \textit{CcAHG2} that was expressed only in leaves of all DT clones but not in those of DS clone 22. Because \textit{CcAHG2} lacks ortholog in \textit{A. thaliana}, its biochemical function is unknown. Despite this, there are two \textit{CcAHG2} orthologous in Solanum species (Solyc08g082260 in \textit{S. lycopersicum} and PGSC0003DM3400076209 in \textit{S. tuberosum}). Further research is therefore needed to know if \textit{CcAHG2} could be used as a molecular marker of drought tolerance in coffee.

Compared to DS clones of \textit{C. canephora}, it was already reported that DT clone 120 had a deeper root system that should allow greater access to soil water (Pinheiro \textit{et al.}, 2005). Interestingly, \textit{CcPYL8b},
but also of \textit{SnRK2} (\textit{CcSnRK2.2}, \textit{CcSnRK2.6} and \textit{CcSnRK2.7}) and \textit{PP2C} (\textit{CcABI1}, \textit{CcABI2} and \textit{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\textsuperscript{T} clone 14 also had a root depth similar to clone 120 (Pinheiro \textit{et al.}, 2005), it did not showed up-regulate expression of \textit{PYL}, \textit{SnRK2} and \textit{PP2C} genes in roots, indicating the existence of different mechanisms amongst the D\textsuperscript{T} coffee clones regarding water deficit (Vieira \textit{et al.}, 2012). Whatever it is, the differences observed for \textit{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\textsuperscript{T} clones. Of course, this does not preclude the involvement of other hormonal regulatory pathways in the establishment of drought tolerance phenotypes in coffee. For example, up-regulated expression of subclass II \textit{SnRK2} genes by salicylic acid, ethylene, and jasmonates, has already been reported (Kulik \textit{et al.}, 2011). The occurrence of such entangled crosstalks between biotic and abiotic pathways might exist in coffee, as suggested by the fact that the D\textsuperscript{T} clone 14 was also recently identified as resistant to multiple races of root-knot nematodes \textit{Meloidogyne} (Lima \textit{et al.}, 2015).

\textbf{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 Científico e Tecnológico/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


CHAPTER 1: THE PYL/PP2C/SnRK2 TRIPARTITE SYSTEM IN C. CANEPHORA


Chapter 1: The PYL/PP2C/SnRK2 Tripartite System in C. Canephora


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\(^7\) (14, 73 and 120) and D\(^5\) (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.
6XSSRUWLQJ ,QIRUPDWLRQ
SGGLWLRQDO 6XSSRUWLQJ ,QIRUPDWLRQ PDI EH IRXQG RQOLQH LQ WKH 6XSSRUWLQJ ,QIRUPDWLRQ WDE IRU WKLV DUWLFOH:

(a) \textbf{CcPYLs}

<table>
<thead>
<tr>
<th>CcPYL1 Ce08_g02750</th>
<th>1</th>
<th>2</th>
<th>12</th>
<th>9</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>9</th>
<th>10</th>
<th>82</th>
</tr>
</thead>
<tbody>
<tr>
<td>CePYL2 Ce08_g10450</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>CePYL4 Ce02_g05990</td>
<td>36</td>
<td>103</td>
<td>61</td>
<td>4</td>
<td>128</td>
<td>171</td>
<td>168</td>
<td>21</td>
<td>16</td>
<td>223</td>
</tr>
<tr>
<td>CePYL7a Ce00_g17440</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CePYL7b Ce00_g23730</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CePYL8a Ce02_g01800</td>
<td>63</td>
<td>13</td>
<td>56</td>
<td>31</td>
<td>28</td>
<td>30</td>
<td>81</td>
<td>71</td>
<td>54</td>
<td>169</td>
</tr>
<tr>
<td>CePYL8b Ce08_g15960</td>
<td>123</td>
<td>21</td>
<td>71</td>
<td>103</td>
<td>152</td>
<td>93</td>
<td>122</td>
<td>40</td>
<td>61</td>
<td>116</td>
</tr>
<tr>
<td>CePYL9 Ce02_g39180</td>
<td>243</td>
<td>28</td>
<td>84</td>
<td>142</td>
<td>169</td>
<td>89</td>
<td>140</td>
<td>41</td>
<td>42</td>
<td>65</td>
</tr>
<tr>
<td>CePYL13 Ce02_g15060</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>18</td>
<td>36</td>
<td>67</td>
<td>129</td>
</tr>
</tbody>
</table>

(b) \textbf{CcPP2Cs}

<table>
<thead>
<tr>
<th>CcABI1 Ce08_g11010</th>
<th>195</th>
<th>19</th>
<th>46</th>
<th>214</th>
<th>36</th>
<th>37</th>
<th>14</th>
<th>14</th>
<th>24</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>CcABI2 Ce06_g17740</td>
<td>112</td>
<td>39</td>
<td>48</td>
<td>100</td>
<td>28</td>
<td>35</td>
<td>22</td>
<td>17</td>
<td>45</td>
<td>62</td>
</tr>
<tr>
<td>CcAHG2 Ce08_g16010</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>61</td>
<td>82</td>
<td>164</td>
</tr>
<tr>
<td>CcAHG3 Ce02_g07430</td>
<td>261</td>
<td>74</td>
<td>214</td>
<td>227</td>
<td>17</td>
<td>9</td>
<td>6</td>
<td>82</td>
<td>124</td>
<td>208</td>
</tr>
<tr>
<td>CcHAB Ce04_g01620</td>
<td>82</td>
<td>12</td>
<td>63</td>
<td>73</td>
<td>23</td>
<td>23</td>
<td>17</td>
<td>11</td>
<td>32</td>
<td>47</td>
</tr>
<tr>
<td>CcHAI Ce01_g13400</td>
<td>103</td>
<td>16</td>
<td>151</td>
<td>156</td>
<td>39</td>
<td>58</td>
<td>40</td>
<td>108</td>
<td>150</td>
<td>135</td>
</tr>
</tbody>
</table>

(c) \textbf{CcSnRK2s}

<table>
<thead>
<tr>
<th>CcSnRK2.1 Ce00_g19520</th>
<th>0</th>
<th>139</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>CcSnRK2.2 Ce07_g05710</td>
<td>219</td>
<td>28</td>
<td>16</td>
<td>18</td>
<td>58</td>
<td>44</td>
<td>47</td>
<td>27</td>
<td>25</td>
<td>136</td>
</tr>
<tr>
<td>CcSnRK2.6 Ce02_g18420</td>
<td>86</td>
<td>16</td>
<td>65</td>
<td>52</td>
<td>33</td>
<td>24</td>
<td>15</td>
<td>12</td>
<td>51</td>
<td>33</td>
</tr>
<tr>
<td>CcSnRK2.7 Ce07_g14700</td>
<td>85</td>
<td>97</td>
<td>10</td>
<td>83</td>
<td>32</td>
<td>23</td>
<td>18</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>CcSnRK2.8 Ce10_g06790</td>
<td>38</td>
<td>13</td>
<td>9</td>
<td>34</td>
<td>28</td>
<td>15</td>
<td>9</td>
<td>22</td>
<td>37</td>
<td>36</td>
</tr>
<tr>
<td>CcSnRK2.10 Ce02_g22790</td>
<td>93</td>
<td>38</td>
<td>82</td>
<td>74</td>
<td>80</td>
<td>89</td>
<td>45</td>
<td>70</td>
<td>63</td>
<td>46</td>
</tr>
<tr>
<td>CcSnRK2.11 Ce08_g11200</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>22</td>
<td>33</td>
<td>41</td>
</tr>
<tr>
<td>CcSnRK2.12 Ce00_g35430</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>CcSnRK2.13 Ce00_g07830</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>11</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

JLIXUH 61+HDW PDS YLXDO\LJ\DLRQ RI WKH \&F3<5/3\(</D), \&F332\&(E) DQG \&F6Q5.2\&(F) JIQHQ IDPLOLHV. )URP OHIW WR ULJKW, WKH OLEUDULHV FRUUHVSQG WR URWW, VVDPHQ, SLVWLO, OHDI, SHULVSHUP (120, 150 DQG 180 GDW DUWHI SROOLQD\DLRQ-\"S\"3\) DQG HQGRVSHUP (180, 260 DQG 320 \"S\"3) IURP \& FDQHISKRUD 515-6HT GDWD. 7UDQVFULSW DEQGDQFQH ZDV QRUPDOL\HG ZLWK 51.0 DQG WKH OHIW RI JIQHQ H\$UHV\$LQ VR LQGLFDW\HG ZLWK D FRORXU VFDOH. IURP ZKLWH (ZHDO) H\$UHV\$HG) WR UHG (VVURQIO\HG) (DGDSWHG IURP. KWWZ://ZZZ.FRIIHH-JHQRPH.RU).
JLJXUH 62 3KrORJHQHVLF DQDQVLYR RI &F3< SURWHLQ VIHTXHFRHV ZLWK RUWKRORJKRVV SURWHLQV RI S. WKDOLDQD ($557$), & VLQHQVLV ($576$), 2. VDWLYD (25<6$)$, 6. OFRSUHULFSP (62$/$), 6. WXEHURVXP (62/78) DQG 9. YLQIHUD (9.79$)$, 7KH SKOR $+$00 DSSURDFK ZDV EDVHG RQ 11, (1HDUHVW 1HLIKERU.QWHFKDQJH) WRSRORJL 6XEILQLOLHV, (JUHHQ$)$, (EOXH) DQG (UHG) DUH LQGLFDWHG. 0DLQ ERRWVWUDV YDOXHV DUH LQGLFDWHG.
&+$37(5<1:7+(3<<\text{332}}&/615.2 \text{75,3857.7}<6\text{67}(0.1&. \&.$1(3+258

\text{\textbf{JLXUH 63 KORJHQHVLQ DQDO\textasciitildeHV RI 3328 SURWHQLQ VIHTXHQHVLQ ZLWK RUWKRORJRQV SURWHQLQV RI 8. WXDOLQD (557+),}}
\&. \text{VLQHQVLQ} \&.76,). 2. \text{VDWLYD (25<6),}} 6. \text{OFRSHULFXV (62вая) \&. 6. WXEHURXRXV (62/78) DQG 9. YLQJHVDL (9/79),}} \text{7KH SK\textasciitilde}>00 DSSURDFK ZDQ EDVHFRQG RQ 11. (1HHDUHVVV HLJIKERU WHUHFKDKQH) WRS\textasciitildeRJ. 6XEDPLOHV DUH LQGLFDW\textasciitildeH WKH GL\textasciitildeHUHQW FRORXUV. 0DLQ ERR\textasciitildeWDSV YDO\textasciitildeHV DUH LQGLFDW\textasciitildeH.
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 CcUBQ10F and R primer pair was used for the ubiquitin (UBI) as reference gene.

<table>
<thead>
<tr>
<th>Gene Family</th>
<th>Gene Name</th>
<th>Gene ID</th>
<th>Primer sequence 5'-3'</th>
</tr>
</thead>
</table>
| PYR/PYL/RCAR | CcPYR1 | Cc08_g02750 | F: CGGTGACGACTGTCCATGAG  
R: TCCGGCACGTCAGAAAGATATA |
|  | CcPYL2 | Cc08_g10450 | F: AAAAGTGGTGGCAGCCATTCG  
R: CCTCCATCCCTCTGCAAGTTG |
|  | CcPYL4 | Cc02_g05990 | F1: CCTATGCTTCTGGCTCCCTCA  
R1: CGCGGAACTGGTGGTTTAG  
F2: TACCATTGTGCGGTGCGACA  
R2: TCTCTCTGCGGCTCCCATGA |
|  | CcPYL7a | Cc00_g17400 | F: GAGGCGCTCGAGAAGCTGAC  
R: GCGGCTGGACATGCAAGAC |
|  | CcPYL7b | Cc00_g23730 | F: GCGGCTGGAGATTCTTGATC  
R: GCGCTGCAATCTGATGTC |
|  | CcPYL8a | Cc02_g01800 | F1: CCTATGCTTCTGGCTCCCTCA  
R1: CGCGGAACTGGTGGTTTAG  
F2: TACCATTGTGCGGTGCGACA  
R2: TCTCTCTGCGGCTCCCATGA |
|  | CcPYL8b | Cc08_g15960 | F: GCCAGAGGGAATAACACAGGA  
R: CAGCTAGGCGCTCAGGACAA |
|  | CcPYL9 | Cc02_g39180 | F: CACCCGTGCTCTTCCTCTCAG  
R: TCCTCACCAGTGACACAAAGC |
|  | CcPYL13 | Cc02_g15060 | F: TCCAAACCGATGACCTCACA  
R: TTGTGAAATTGCAGGACCAA |
|  | CcABI1 | Cc08_g11010 | F: TCTGCAAGGTGTAGGAAAGAAA  
R: CGAAACAAAGGCAAGCAACA |
|  | CcABI2 | Cc06_g11740 | F: TACGGCTGGTGGTGGCATTA  
R: CGCTCGGACATGCAAGAC |
|  | CcAHG2 | Cc08_g16010 | F: AGAGGGTGTCCTCAGGGATT  
R: GCTCTGCTTCTGGCTCCCAACAA |
|  | CcAHG3 | Cc02_g07430 | F: ACCGGGAGGATGACGATAATCG  
R: CCCCAAAGCTTGCTCAGTGG |
|  | CcHAB | Cc04_g01620 | F: TGCTTGGTGGATGCTCAGA  
R: CTTCTCCTCTGGCTCCCAACAA |
|  | CcHAL | Cc01_g13400 | F: CATCAGAGCCTGGTGGTCAAT  
R: CCACCGGGCTCTCTCCATACCT |
| PP2C | CcSnRK2.1 | Cc00_g19320 | F: TAGCCCCCGAGGTTCTCTCTCCT  
R: TCATCCCGCAAGGAGACACA |
|  | CcSnRK2.2 | Cc07_g05710 | F: CGAGGGTGGGCTGCTGGTTT  
R: GCTGGGCTCTCGGCTACAAACA |
|  | CcSnRK2.6 | Cc02_g18420 | F: GCATATATGTGCGCCGGAAGT  
R: AAGGGTGATGCGCCGGAAGT |
|  | CcSnRK2.7 | Cc07_g14700 | F: AAGCCCGAAGACACAGCTCA  
R: GATTGGGTGGGGAATGCAA |
|  | CcSnRK2.8 | Cc10_g06790 | F1: AACATGCGAGCGGGAGATT  
R1: CTCTGCGGCACTACCCATT  
F2: CGCTCTCAAAGGAGGCTTGCT  
R2: TCTCTCCTCCGCGGCACTACCT |
|  | CcSnRK2.10 | Cc02_g22790 | F: TCGATTCAAGGAGGTGGTGT  
R: TTCCCCTCCGAGCAGCATTACCT |
|  | CcSnRK2.11 | Cc08_g11200 | F: AGGACCTGACGCTCAGCCAA  
R: CCTGGGAATTTTTTCTGCTCCTT |
|  | CcSnRK2.12 | Cc00_g35430 | F: ACTTGAAGCTGACCCTCACCAA  
R: CCTGGGAATTTTTTCTGCTCCTT |
|  | CcSnRK2.13 | Cc00_g07830 | F: GTCAAGGAAGAGGATGATGGGAAG  
R: ACTTGAAGCTGAGAAAACACACTA |
|  | CcUBQ10 | Cc05_g13290 | F: AAGACAGCTCCACACGAGTACACGAT  
R: GCCAGGACCTTGGGCTGACTATA |
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 Phytozone 10.3 (http://phytozome.jgi.doe.gov/pz/portal.html).

<table>
<thead>
<tr>
<th>Coffee ID</th>
<th>NCBI ID</th>
<th>Gene ID</th>
<th>Sp</th>
<th>I</th>
<th>aa</th>
<th>Ident.</th>
<th>Align.</th>
<th>QC</th>
<th>e-value</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cc00_g17440</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAD25950.1</td>
<td>AT2G40350.1</td>
<td>At</td>
<td>0</td>
<td>175</td>
<td>47%</td>
<td>69%</td>
<td>91%</td>
<td>2.21E-47</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>KDO68852.1</td>
<td>orangel.1g038201m</td>
<td>Cs</td>
<td>0</td>
<td>201</td>
<td>49%</td>
<td>69%</td>
<td>97%</td>
<td>2.53E-49</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>EEE62745.1</td>
<td>LOC_Os05g12260.1</td>
<td>Os</td>
<td>0</td>
<td>196</td>
<td>48%</td>
<td>69%</td>
<td>84%</td>
<td>9.3E-42</td>
<td>hypothetical protein OsJ_17548</td>
<td></td>
</tr>
<tr>
<td>XP_004249065.1</td>
<td>Solyc10g076410.1.1</td>
<td>Sl</td>
<td>0</td>
<td>203</td>
<td>52%</td>
<td>71%</td>
<td>92%</td>
<td>4.02E-50</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>XP_006359557.1</td>
<td>PGSC0003DMG400029194</td>
<td>St</td>
<td>0</td>
<td>214</td>
<td>52%</td>
<td>72%</td>
<td>92%</td>
<td>7.21E-51</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>XP_002264158.1</td>
<td>GSVIVG01032747001</td>
<td>Vv</td>
<td>0</td>
<td>227</td>
<td>54%</td>
<td>72%</td>
<td>95%</td>
<td>7.95E-54</td>
<td>PREDICTED: abscisic acid receptor PYL4</td>
<td></td>
</tr>
<tr>
<td><strong>Cc00_g23730</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAD25950.1</td>
<td>AT2G40350.1</td>
<td>At</td>
<td>0</td>
<td>175</td>
<td>46%</td>
<td>68%</td>
<td>93%</td>
<td>6.88E-48</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>KDO68852.1</td>
<td>orangel.1g038201m</td>
<td>Cs</td>
<td>0</td>
<td>201</td>
<td>48%</td>
<td>68%</td>
<td>97%</td>
<td>3.42E-48</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>EEE62745.1</td>
<td>LOC_Os05g12260.1</td>
<td>Os</td>
<td>0</td>
<td>196</td>
<td>47%</td>
<td>68%</td>
<td>84%</td>
<td>1.57E-39</td>
<td>hypothetical protein OsJ_17548</td>
<td></td>
</tr>
<tr>
<td>XP_004249671.1</td>
<td>Solyc10g085310.1</td>
<td>Sl</td>
<td>0</td>
<td>213</td>
<td>50%</td>
<td>71%</td>
<td>93%</td>
<td>3.9E-49</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>XP_006359557.1</td>
<td>PGSC0003DMG400029194</td>
<td>St</td>
<td>0</td>
<td>214</td>
<td>51%</td>
<td>71%</td>
<td>92%</td>
<td>5.34E-50</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>XP_002264158.1</td>
<td>GSVIVG01032747001</td>
<td>Vv</td>
<td>0</td>
<td>227</td>
<td>53%</td>
<td>71%</td>
<td>94%</td>
<td>3.01E-52</td>
<td>PREDICTED: abscisic acid receptor PYL4</td>
<td></td>
</tr>
<tr>
<td><strong>Cc02_g01800</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAFO0266.1</td>
<td>AT5G53160.2</td>
<td>At</td>
<td>2</td>
<td>188</td>
<td>82%</td>
<td>92%</td>
<td>90%</td>
<td>5E-100</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>XP_006473961.1</td>
<td>orange1.1g028067m</td>
<td>Cs</td>
<td>2</td>
<td>197</td>
<td>78%</td>
<td>90%</td>
<td>98%</td>
<td>1.2E-103</td>
<td>PREDICTED: abscisic acid receptor PYL8-like</td>
<td></td>
</tr>
<tr>
<td>NP_001046464.1</td>
<td>LOC_Os05g12440.1</td>
<td>Os</td>
<td>2</td>
<td>204</td>
<td>79%</td>
<td>92%</td>
<td>94%</td>
<td>2.71E-98</td>
<td>Os02g0255500</td>
<td></td>
</tr>
<tr>
<td>XP_002270037.3</td>
<td>GSVIVG01028704001</td>
<td>Vv</td>
<td>2</td>
<td>185</td>
<td>85%</td>
<td>94%</td>
<td>96%</td>
<td>2.1E-109</td>
<td>PREDICTED: abscisic acid receptor PYL8</td>
<td></td>
</tr>
<tr>
<td><strong>Cc02_g05990</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_565928.1</td>
<td>AT2G40350.1</td>
<td>At</td>
<td>0</td>
<td>215</td>
<td>57%</td>
<td>71%</td>
<td>91%</td>
<td>9.06E-74</td>
<td>abscisic acid receptor PYL6</td>
<td></td>
</tr>
<tr>
<td>KDO68852.1</td>
<td>orangel.1g038201m</td>
<td>Cs</td>
<td>0</td>
<td>201</td>
<td>77%</td>
<td>87%</td>
<td>72%</td>
<td>5.75E-86</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>NP_001055819.1</td>
<td>LOC_Os05g39580.1</td>
<td>Os</td>
<td>0</td>
<td>216</td>
<td>65%</td>
<td>75%</td>
<td>71%</td>
<td>4.62E-62</td>
<td>Os05g0473000</td>
<td></td>
</tr>
<tr>
<td>XP_004235232.1</td>
<td>Solyc03g095780.1.1</td>
<td>Sl</td>
<td>0</td>
<td>201</td>
<td>75%</td>
<td>82%</td>
<td>88%</td>
<td>1.4E-99</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>XP_006353422.1</td>
<td>PGSC0003DMG400023949</td>
<td>St</td>
<td>0</td>
<td>218</td>
<td>74%</td>
<td>81%</td>
<td>91%</td>
<td>1.3E-104</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>XP_002264158.1</td>
<td>GSVIVG01032747001</td>
<td>Vv</td>
<td>0</td>
<td>227</td>
<td>68%</td>
<td>77%</td>
<td>90%</td>
<td>2.59E-91</td>
<td>PREDICTED: abscisic acid receptor PYL4</td>
<td></td>
</tr>
<tr>
<td><strong>Cc02_g15060</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAD25950.1</td>
<td>AT2G40350.1</td>
<td>At</td>
<td>0</td>
<td>175</td>
<td>56%</td>
<td>79%</td>
<td>87%</td>
<td>5.89E-54</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>KDO68852.1</td>
<td>orangel.1g038201m</td>
<td>Cs</td>
<td>0</td>
<td>201</td>
<td>60%</td>
<td>75%</td>
<td>93%</td>
<td>1.17E-60</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>NP_001049383.1</td>
<td>LOC_Os03g18600.1</td>
<td>Os</td>
<td>0</td>
<td>229</td>
<td>55%</td>
<td>77%</td>
<td>93%</td>
<td>5.45E-52</td>
<td>Os03g0297600</td>
<td></td>
</tr>
<tr>
<td>XP_002249671.1</td>
<td>Solyc10g085310.1.1</td>
<td>Sl</td>
<td>0</td>
<td>213</td>
<td>53%</td>
<td>79%</td>
<td>94%</td>
<td>3.56E-61</td>
<td>PREDICTED: abscisic acid receptor PYL4</td>
<td></td>
</tr>
<tr>
<td>XP_006359557.1</td>
<td>PGSC0003DMG400029194</td>
<td>St</td>
<td>0</td>
<td>214</td>
<td>57%</td>
<td>79%</td>
<td>93%</td>
<td>1.09E-61</td>
<td>PREDICTED: abscisic acid receptor PYL4-like</td>
<td></td>
</tr>
<tr>
<td>CAN72620.1</td>
<td>GSVIVG01013161001</td>
<td>Vv</td>
<td>0</td>
<td>172</td>
<td>62%</td>
<td>76%</td>
<td>97%</td>
<td>4.64E-70</td>
<td>hypothetical protein VITISV_004947</td>
<td></td>
</tr>
</tbody>
</table>
Table S1

<table>
<thead>
<tr>
<th>Coffee ID</th>
<th>NCBII ID</th>
<th>Gene ID</th>
<th>Sp</th>
<th>I</th>
<th>aa</th>
<th>Ident.</th>
<th>Align.</th>
<th>QC</th>
<th>e-value</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>3OQU</td>
<td>AT1G01360.1</td>
<td>At</td>
<td>2</td>
<td>205</td>
<td>75%</td>
<td>87%</td>
<td>96%</td>
<td>5.82E-87</td>
<td>Abscisic Acid Receptor Pyl9</td>
<td></td>
</tr>
<tr>
<td>XP_006476396.1</td>
<td>orange1.1g029200m</td>
<td>Cs</td>
<td>2</td>
<td>197</td>
<td>75%</td>
<td>88%</td>
<td>96%</td>
<td>8.42E-90</td>
<td>PREDICTED: abscisic acid receptor PYL8-like</td>
<td></td>
</tr>
<tr>
<td>NP_001054923.1</td>
<td>LOC_Os05g12260.1</td>
<td>Os</td>
<td>2</td>
<td>209</td>
<td>76%</td>
<td>88%</td>
<td>90%</td>
<td>5.84E-84</td>
<td>Os05g0213500</td>
<td></td>
</tr>
<tr>
<td>XP_004231210.1</td>
<td>Solye01g085700.2.1</td>
<td>St</td>
<td>2</td>
<td>186</td>
<td>83%</td>
<td>95%</td>
<td>91%</td>
<td>4.15E-98</td>
<td>PREDICTED: abscisic acid receptor PYL8</td>
<td></td>
</tr>
<tr>
<td>NP_001284557.1</td>
<td>PGSC0003DMG40000215</td>
<td>St</td>
<td>2</td>
<td>186</td>
<td>81%</td>
<td>95%</td>
<td>91%</td>
<td>5.95E-96</td>
<td>abscisic acid receptor PYL8-like</td>
<td></td>
</tr>
<tr>
<td>XP_010659134.1</td>
<td>GSVIVG01019517001</td>
<td>Vv</td>
<td>2</td>
<td>189</td>
<td>76%</td>
<td>89%</td>
<td>96%</td>
<td>6.96E-90</td>
<td>PREDICTED: abscisic acid receptor PYL8-like</td>
<td></td>
</tr>
<tr>
<td>NP_193521.1</td>
<td>AT4G17870.1</td>
<td>At</td>
<td>0</td>
<td>191</td>
<td>68%</td>
<td>82%</td>
<td>83%</td>
<td>2.37E-87</td>
<td>abscisic acid receptor PYR1</td>
<td></td>
</tr>
<tr>
<td>XP_006491739</td>
<td>orange1.1g046151m</td>
<td>Cs</td>
<td>0</td>
<td>187</td>
<td>55%</td>
<td>72%</td>
<td>78%</td>
<td>4.97E-66</td>
<td>PREDICTED: abscisic acid receptor PYL1-like</td>
<td></td>
</tr>
<tr>
<td>NP_001065470.1</td>
<td>LOC_Os10g42280.1</td>
<td>Os</td>
<td>0</td>
<td>212</td>
<td>62%</td>
<td>73%</td>
<td>75%</td>
<td>2.29E-64</td>
<td>Os10g0573400</td>
<td></td>
</tr>
<tr>
<td>XP_004245893.1</td>
<td>Solyc08g076960.1.1</td>
<td>Sl</td>
<td>0</td>
<td>231</td>
<td>72%</td>
<td>80%</td>
<td>97%</td>
<td>3.7E-110</td>
<td>PREDICTED: abscisic acid receptor PYR1-like</td>
<td></td>
</tr>
<tr>
<td>NP_001284559.1</td>
<td>PGSC0003DMG4000017514</td>
<td>St</td>
<td>0</td>
<td>231</td>
<td>70%</td>
<td>80%</td>
<td>97%</td>
<td>1.3E-107</td>
<td>abscisic acid receptor PYL1-like</td>
<td></td>
</tr>
<tr>
<td>XP_002280361.1</td>
<td>GSVIVG01013161001</td>
<td>Vv</td>
<td>1</td>
<td>214</td>
<td>76%</td>
<td>85%</td>
<td>78%</td>
<td>7.3E-99</td>
<td>PREDICTED: abscisic acid receptor PYR1-like</td>
<td></td>
</tr>
<tr>
<td>NP_180174.1</td>
<td>AT2G26040.1</td>
<td>At</td>
<td>0</td>
<td>190</td>
<td>72%</td>
<td>82%</td>
<td>95%</td>
<td>7.97E-85</td>
<td>abscisic acid receptor PYL2</td>
<td></td>
</tr>
<tr>
<td>KDO80051.1</td>
<td>orange1.1g046697m</td>
<td>Cs</td>
<td>0</td>
<td>187</td>
<td>82%</td>
<td>88%</td>
<td>99%</td>
<td>1.4E-106</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>NP_001172865.1</td>
<td>LOC_Os02g13330.1</td>
<td>Os</td>
<td>0</td>
<td>207</td>
<td>61%</td>
<td>74%</td>
<td>90%</td>
<td>2.27E-60</td>
<td>Os02g0226801</td>
<td></td>
</tr>
<tr>
<td>XP_004253195.1</td>
<td>Solyc12g095700.1.1</td>
<td>Sl</td>
<td>1</td>
<td>190</td>
<td>80%</td>
<td>88%</td>
<td>99%</td>
<td>1.1E-103</td>
<td>PREDICTED: abscisic acid receptor PYL2-like</td>
<td></td>
</tr>
<tr>
<td>NP_006360983.1</td>
<td>PGSC0003DMG400029952</td>
<td>St</td>
<td>1</td>
<td>188</td>
<td>84%</td>
<td>89%</td>
<td>98%</td>
<td>9.3E-107</td>
<td>PREDICTED: abscisic acid receptor PYL2-like</td>
<td></td>
</tr>
<tr>
<td>XP_006483333.1</td>
<td>GSVIVG01033562001</td>
<td>Vv</td>
<td>0</td>
<td>185</td>
<td>82%</td>
<td>88%</td>
<td>95%</td>
<td>2.8E-99</td>
<td>PREDICTED: abscisic acid receptor PYL2</td>
<td></td>
</tr>
<tr>
<td>NP_200128.1</td>
<td>AT5G53160.2</td>
<td>At</td>
<td>2</td>
<td>188</td>
<td>80%</td>
<td>91%</td>
<td>96%</td>
<td>2.32E-95</td>
<td>regulatory component of ABA receptor 3</td>
<td></td>
</tr>
<tr>
<td>XP_006476396.1</td>
<td>orange1.1g029200m</td>
<td>Cs</td>
<td>2</td>
<td>197</td>
<td>87%</td>
<td>95%</td>
<td>97%</td>
<td>2.9E-106</td>
<td>PREDICTED: abscisic acid receptor PYL8-like</td>
<td></td>
</tr>
<tr>
<td>NP_001046464.1</td>
<td>LOC_Os02g13540.1</td>
<td>Os</td>
<td>2</td>
<td>204</td>
<td>80%</td>
<td>90%</td>
<td>97%</td>
<td>1.92E-95</td>
<td>Os02g0255500</td>
<td></td>
</tr>
<tr>
<td>XP_004245823.1</td>
<td>Solyc08g082170.0.2.1</td>
<td>Sl</td>
<td>2</td>
<td>189</td>
<td>85%</td>
<td>95%</td>
<td>97%</td>
<td>3E-105</td>
<td>PREDICTED: abscisic acid receptor PYL9</td>
<td></td>
</tr>
<tr>
<td>NP_006343869.1</td>
<td>PGSC0003DMG400012155</td>
<td>St</td>
<td>2</td>
<td>189</td>
<td>85%</td>
<td>95%</td>
<td>97%</td>
<td>3E-105</td>
<td>PREDICTED: abscisic acid receptor PYL9-like</td>
<td></td>
</tr>
<tr>
<td>XP_010659134.1</td>
<td>GSVIVG01019517001</td>
<td>Vv</td>
<td>3</td>
<td>189</td>
<td>90%</td>
<td>95%</td>
<td>97%</td>
<td>7.7E-111</td>
<td>PREDICTED: abscisic acid receptor PYL8-like</td>
<td></td>
</tr>
</tbody>
</table>
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).

<table>
<thead>
<tr>
<th>Coffee ID</th>
<th>NCBI ID</th>
<th>Gene ID</th>
<th>Sp</th>
<th>I</th>
<th>Len.aa</th>
<th>Ident.</th>
<th>Align.</th>
<th>Query cov.</th>
<th>e-value</th>
<th>Function induced by BLAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ce01_g13400</td>
<td>NP_180499.1</td>
<td>AT2G29380.1</td>
<td>At</td>
<td>2</td>
<td>362</td>
<td>71%</td>
<td>83%</td>
<td>69%</td>
<td>2.3E-139</td>
<td>highly ABA-induced PP2C protein</td>
</tr>
<tr>
<td>XP_006488392.1</td>
<td>orange1.1g036852m</td>
<td>LOC_Os01g62760.1</td>
<td>Os</td>
<td>3</td>
<td>414</td>
<td>55%</td>
<td>68%</td>
<td>70%</td>
<td>8.7E-108</td>
<td>PREDICTED: protein phosphatase 2C 78-like</td>
</tr>
<tr>
<td>XP_004241211.1</td>
<td>Solyc08g076400.2.1</td>
<td>St</td>
<td>3</td>
<td>410</td>
<td>67%</td>
<td>76%</td>
<td>98%</td>
<td>7.2E-171</td>
<td>PREDICTED: protein phosphatase 2C 78-like</td>
<td></td>
</tr>
<tr>
<td>XP_003507089.1</td>
<td>PGSC0003DMG40003032</td>
<td>St</td>
<td>3</td>
<td>410</td>
<td>67%</td>
<td>76%</td>
<td>98%</td>
<td>7.2E-171</td>
<td>PREDICTED: protein phosphatase 2C 78-like</td>
<td></td>
</tr>
<tr>
<td>XP_002282608.1</td>
<td>GSVIT01024875001</td>
<td>Vv</td>
<td>3</td>
<td>408</td>
<td>65%</td>
<td>78%</td>
<td>98%</td>
<td>2.3E-158</td>
<td>PREDICTED: probable protein phosphatase 2C 78-like</td>
<td></td>
</tr>
<tr>
<td>Ce02_g07430</td>
<td>NP_172223.1</td>
<td>AT1G04730.1</td>
<td>At</td>
<td>2</td>
<td>442</td>
<td>63%</td>
<td>77%</td>
<td>69%</td>
<td>1.3E-134</td>
<td>protein phosphatase 2C 3</td>
</tr>
<tr>
<td>XP_006488392.1</td>
<td>orange1.1g036852m</td>
<td>LOC_Os01g62760.1</td>
<td>Os</td>
<td>3</td>
<td>414</td>
<td>58%</td>
<td>71%</td>
<td>73%</td>
<td>1.6E-121</td>
<td>PREDICTED: probable protein phosphatase 2C 78-like</td>
</tr>
<tr>
<td>XP_004239911.1</td>
<td>Solyc05g052980.2.1</td>
<td>Sl</td>
<td>3</td>
<td>409</td>
<td>68%</td>
<td>78%</td>
<td>100%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 37-like</td>
<td></td>
</tr>
<tr>
<td>XP_006355694.1</td>
<td>PGSC0003DMG400027196</td>
<td>St</td>
<td>3</td>
<td>418</td>
<td>67%</td>
<td>77%</td>
<td>98%</td>
<td>7.9E-163</td>
<td>PREDICTED: protein phosphatase 2C 37-like</td>
<td></td>
</tr>
<tr>
<td>XP_002282703.1</td>
<td>GSVIT01016485001</td>
<td>Vv</td>
<td>3</td>
<td>400</td>
<td>63%</td>
<td>73%</td>
<td>99%</td>
<td>7.9E-163</td>
<td>PREDICTED: protein phosphatase 2C 37-like</td>
<td></td>
</tr>
<tr>
<td>Ce04_g01620</td>
<td>NP_177421.1</td>
<td>AT1G17770.1</td>
<td>At</td>
<td>4</td>
<td>511</td>
<td>56%</td>
<td>74%</td>
<td>98%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 16</td>
</tr>
<tr>
<td>XP_006465975.1</td>
<td>orange1.1g069094m</td>
<td>LOC_Os01g40099.1</td>
<td>Os</td>
<td>3</td>
<td>352</td>
<td>66%</td>
<td>77%</td>
<td>64%</td>
<td>4.8E-161</td>
<td>PREDICTED: protein phosphatase 2C 16-like</td>
</tr>
<tr>
<td>EEE54872.1</td>
<td>Solyc09g121880.2.1</td>
<td>Sl</td>
<td>4</td>
<td>544</td>
<td>73%</td>
<td>82%</td>
<td>98%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C AB12 homolog</td>
<td></td>
</tr>
<tr>
<td>BA39959.1</td>
<td>Solyc09g121880.2.1</td>
<td>Sl</td>
<td>4</td>
<td>544</td>
<td>73%</td>
<td>82%</td>
<td>98%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 16-like</td>
<td></td>
</tr>
<tr>
<td>XP_006432955.1</td>
<td>PGSC0003DMG400002573</td>
<td>St</td>
<td>4</td>
<td>545</td>
<td>73%</td>
<td>83%</td>
<td>98%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 16</td>
<td></td>
</tr>
<tr>
<td>XP_002278167.2</td>
<td>GSVIT01016186001</td>
<td>Vv</td>
<td>3</td>
<td>548</td>
<td>62%</td>
<td>74%</td>
<td>98%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 16</td>
<td></td>
</tr>
<tr>
<td>Ce06_g11740</td>
<td>NP_177421.1</td>
<td>AT1G17770.1</td>
<td>At</td>
<td>4</td>
<td>511</td>
<td>63%</td>
<td>74%</td>
<td>61%</td>
<td>1.1E-142</td>
<td>protein phosphatase 2C 16</td>
</tr>
<tr>
<td>KDO73536.1</td>
<td>orange1.1g008890m</td>
<td>LOC_Os01g40099.1</td>
<td>Os</td>
<td>3</td>
<td>396</td>
<td>63%</td>
<td>75%</td>
<td>61%</td>
<td>6.9E-137</td>
<td>hypothetical protein Os01g02363</td>
</tr>
<tr>
<td>XP_001046464.1</td>
<td>LOC_Os01g40099.1</td>
<td>Os</td>
<td>3</td>
<td>396</td>
<td>63%</td>
<td>75%</td>
<td>61%</td>
<td>6.9E-137</td>
<td>hypothetical protein Os01g02363</td>
<td></td>
</tr>
<tr>
<td>XP_00423737.1</td>
<td>SOLYCH07g04990.2.1</td>
<td>Sl</td>
<td>3</td>
<td>543</td>
<td>52%</td>
<td>68%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 50</td>
<td></td>
</tr>
<tr>
<td>XP_006342333.1</td>
<td>PGSC0003DMG400018004</td>
<td>St</td>
<td>4</td>
<td>543</td>
<td>52%</td>
<td>68%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 6-like</td>
<td></td>
</tr>
<tr>
<td>XP_002279140.1</td>
<td>GSVIT01015156001</td>
<td>Vv</td>
<td>3</td>
<td>550</td>
<td>50%</td>
<td>66%</td>
<td>100%</td>
<td>5E-17</td>
<td>PREDICTED: protein phosphatase 2C 77</td>
<td></td>
</tr>
<tr>
<td>Ce08_g111010</td>
<td>NP_177421.1</td>
<td>AT1G17770.1</td>
<td>At</td>
<td>4</td>
<td>511</td>
<td>49%</td>
<td>63%</td>
<td>100%</td>
<td>2.6E-158</td>
<td>protein phosphatase 2C 16</td>
</tr>
<tr>
<td>KDO73536.1</td>
<td>orange1.1g008890m</td>
<td>LOC_Os01g40099.1</td>
<td>Os</td>
<td>3</td>
<td>396</td>
<td>69%</td>
<td>80%</td>
<td>59%</td>
<td>2E-155</td>
<td>hypothetical protein Cis1g008880mg</td>
</tr>
<tr>
<td>XP_001065470.1</td>
<td>LOC_Os01g40099.1</td>
<td>Os</td>
<td>3</td>
<td>396</td>
<td>69%</td>
<td>80%</td>
<td>59%</td>
<td>2E-155</td>
<td>hypothetical protein Cis1g008880mg</td>
<td></td>
</tr>
<tr>
<td>XP_004253091.1</td>
<td>Solyc12g096020.1</td>
<td>Sl</td>
<td>3</td>
<td>540</td>
<td>66%</td>
<td>79%</td>
<td>100%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 53</td>
<td></td>
</tr>
<tr>
<td>XP_006342498.1</td>
<td>PGSC0003DMG400029297</td>
<td>St</td>
<td>3</td>
<td>536</td>
<td>67%</td>
<td>79%</td>
<td>100%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 6-like</td>
<td></td>
</tr>
<tr>
<td>XP_010648365.1</td>
<td>GSVIT01035420001</td>
<td>Vv</td>
<td>4</td>
<td>551</td>
<td>66%</td>
<td>78%</td>
<td>100%</td>
<td>0</td>
<td>PREDICTED: protein phosphatase 2C 53</td>
<td></td>
</tr>
<tr>
<td>Ce08_g16010</td>
<td>NP_172223.1</td>
<td>AT1G04730.1</td>
<td>At</td>
<td>2</td>
<td>442</td>
<td>47%</td>
<td>62%</td>
<td>64%</td>
<td>1.1E-80</td>
<td>protein phosphatase 2C 3</td>
</tr>
<tr>
<td>KDO76517.1</td>
<td>orange1.1g023178m</td>
<td>LOC_Os01g46760.1</td>
<td>Os</td>
<td>2</td>
<td>403</td>
<td>55%</td>
<td>70%</td>
<td>64%</td>
<td>2.7E-98</td>
<td>hypothetical protein Os01g0656200</td>
</tr>
<tr>
<td>XP_004240955.1</td>
<td>Solyc06g051940.2.1</td>
<td>Sl</td>
<td>2</td>
<td>442</td>
<td>57%</td>
<td>73%</td>
<td>67%</td>
<td>4.6E-103</td>
<td>PREDICTED: probable protein phosphatase 2C 51</td>
<td></td>
</tr>
<tr>
<td>XP_006350681.1</td>
<td>PGSC0003DMG400099112</td>
<td>St</td>
<td>2</td>
<td>399</td>
<td>54%</td>
<td>70%</td>
<td>64%</td>
<td>1.8E-94</td>
<td>PREDICTED: probable protein phosphatase 2C 51-like</td>
<td></td>
</tr>
<tr>
<td>XP_002266149.1</td>
<td>GSVIT01019525001</td>
<td>Vv</td>
<td>2</td>
<td>393</td>
<td>57%</td>
<td>75%</td>
<td>70%</td>
<td>2E-107</td>
<td>PREDICTED: probable protein phosphatase 2C 51</td>
<td></td>
</tr>
</tbody>
</table>
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](http://phytozome.jgi.doe.gov/pz/portal.html)).

<table>
<thead>
<tr>
<th>Coffee ID</th>
<th>NCBI ID</th>
<th>Gene ID</th>
<th>Sp</th>
<th>I</th>
<th>Len.aa</th>
<th>Ident.</th>
<th>Align.</th>
<th>Query cov.</th>
<th>e-value</th>
<th>Function inferred by BLAST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cc00_g07830</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_172563.1</td>
<td>AT1G10940.1</td>
<td>At</td>
<td>9</td>
<td>363</td>
<td>80%</td>
<td>88%</td>
<td>70%</td>
<td>4.32E-76</td>
<td>serine/threonine-protein kinase SRK2A</td>
<td></td>
</tr>
<tr>
<td>XP_006477070.1</td>
<td>orange1.1g019433m</td>
<td>Cs</td>
<td>8</td>
<td>341</td>
<td>84%</td>
<td>90%</td>
<td>70%</td>
<td>2.11E-77</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3-like</td>
<td></td>
</tr>
<tr>
<td>NP_001050274.1</td>
<td>LOC_Os03g27280.1</td>
<td>Os</td>
<td>8</td>
<td>342</td>
<td>82%</td>
<td>88%</td>
<td>70%</td>
<td>2.78E-75</td>
<td>Os03g0390200</td>
<td></td>
</tr>
<tr>
<td>XP_005425833.1</td>
<td>Solyco08g07780.2.1</td>
<td>Sl</td>
<td>8</td>
<td>339</td>
<td>82%</td>
<td>90%</td>
<td>70%</td>
<td>9.46E-77</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td>XP_006359207.1</td>
<td>PGSC0003DMG400026211</td>
<td>St</td>
<td>8</td>
<td>339</td>
<td>82%</td>
<td>90%</td>
<td>70%</td>
<td>9.46E-77</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3-like</td>
<td></td>
</tr>
<tr>
<td>XP_002262726.1</td>
<td>GSIVT01004839001</td>
<td>Vv</td>
<td>8</td>
<td>340</td>
<td>83%</td>
<td>88%</td>
<td>70%</td>
<td>5.14E-76</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td><strong>Cc00_g19320</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_196476.1</td>
<td>AT5G08590.1</td>
<td>At</td>
<td>9</td>
<td>353</td>
<td>85%</td>
<td>93%</td>
<td>90%</td>
<td>0</td>
<td>serine/threonine-protein kinase SRK2G</td>
<td></td>
</tr>
<tr>
<td>XP_006466196.1</td>
<td>orange1.1g018734m</td>
<td>Cs</td>
<td>8</td>
<td>354</td>
<td>84%</td>
<td>92%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK2A-like</td>
<td></td>
</tr>
<tr>
<td>NP_001052827.1</td>
<td>LOC_Os04g35240.1</td>
<td>Os</td>
<td>8</td>
<td>359</td>
<td>76%</td>
<td>85%</td>
<td>99%</td>
<td>0</td>
<td>Os04g0343200</td>
<td></td>
</tr>
<tr>
<td>XP_004396281.1</td>
<td>Solyco05g056550.2.1</td>
<td>Sl</td>
<td>6</td>
<td>356</td>
<td>85%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK7-like</td>
<td></td>
</tr>
<tr>
<td>NP_001274892.1</td>
<td>PGSC0003DMG400023803</td>
<td>St</td>
<td>9</td>
<td>360</td>
<td>82%</td>
<td>90%</td>
<td>98%</td>
<td>0</td>
<td>serine/threonine-protein kinase SRK2B-like</td>
<td></td>
</tr>
<tr>
<td>XP_002267922.1</td>
<td>GSIVT01022427001</td>
<td>Vv</td>
<td>8</td>
<td>355</td>
<td>85%</td>
<td>90%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK2A</td>
<td></td>
</tr>
<tr>
<td><strong>Cc00_g35430</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_172563.1</td>
<td>AT1G10940.1</td>
<td>At</td>
<td>9</td>
<td>363</td>
<td>87%</td>
<td>93%</td>
<td>94%</td>
<td>8.76E-71</td>
<td>serine/threonine-protein kinase SRK2A</td>
<td></td>
</tr>
<tr>
<td>KDO07502.1</td>
<td>orange1.1g0124336m</td>
<td>Cs</td>
<td>8</td>
<td>269</td>
<td>92%</td>
<td>95%</td>
<td>94%</td>
<td>7.14E-75</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>BAT12097.1</td>
<td>LOC_Os10g44190.1</td>
<td>Os</td>
<td>8</td>
<td>289</td>
<td>88%</td>
<td>94%</td>
<td>94%</td>
<td>1.8E-72</td>
<td>Os10g0564500</td>
<td></td>
</tr>
<tr>
<td>XP_004248833.1</td>
<td>Solyco08g07780.2.1</td>
<td>Sl</td>
<td>8</td>
<td>339</td>
<td>92%</td>
<td>95%</td>
<td>94%</td>
<td>5.72E-74</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td>XP_006359207.1</td>
<td>PGSC0003DMG400026211</td>
<td>St</td>
<td>8</td>
<td>339</td>
<td>92%</td>
<td>95%</td>
<td>94%</td>
<td>7.26E-74</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3-like</td>
<td></td>
</tr>
<tr>
<td>XP_002262726.1</td>
<td>GSIVT01004839001</td>
<td>Vv</td>
<td>8</td>
<td>340</td>
<td>92%</td>
<td>94%</td>
<td>94%</td>
<td>3.02E-73</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td><strong>Cc2_g18420</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_567945.1</td>
<td>AT4G33950.1</td>
<td>At</td>
<td>9</td>
<td>362</td>
<td>87%</td>
<td>93%</td>
<td>94%</td>
<td>0</td>
<td>calcium-independent ABA-activated protein kinase</td>
<td></td>
</tr>
<tr>
<td>KDO49166.1</td>
<td>orange1.1g017933m</td>
<td>Cs</td>
<td>7</td>
<td>363</td>
<td>91%</td>
<td>97%</td>
<td>99%</td>
<td>0</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>NP_001050653.1</td>
<td>LOC_Os03g44160.1</td>
<td>Os</td>
<td>6</td>
<td>362</td>
<td>87%</td>
<td>92%</td>
<td>99%</td>
<td>0</td>
<td>Os03g0610900</td>
<td></td>
</tr>
<tr>
<td>XP_004302794.1</td>
<td>Solyco01g010280.2.1</td>
<td>Sl</td>
<td>9</td>
<td>362</td>
<td>93%</td>
<td>98%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK2E</td>
<td></td>
</tr>
<tr>
<td>NP_001275318.1</td>
<td>PGSC0003DMG400025895</td>
<td>St</td>
<td>10</td>
<td>362</td>
<td>93%</td>
<td>97%</td>
<td>99%</td>
<td>0</td>
<td>serine/threonine-protein kinase SRK2E-like</td>
<td></td>
</tr>
<tr>
<td>XP_002284959.1</td>
<td>GSIVT01031806001</td>
<td>Vv</td>
<td>8</td>
<td>363</td>
<td>90%</td>
<td>96%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK10</td>
<td></td>
</tr>
</tbody>
</table>
### Table S3

Continue...for legend see the previous page.

<table>
<thead>
<tr>
<th>Coffee ID</th>
<th>NCBI ID</th>
<th>Gene ID</th>
<th>Sp</th>
<th>I</th>
<th>Len.aa</th>
<th>Ident.</th>
<th>Align.</th>
<th>Query cov.</th>
<th>e-value</th>
<th>Function inferred by BLAST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cc02_g22790</strong></td>
<td>AAM57112.1</td>
<td>AT1G60940.1</td>
<td>At</td>
<td>8</td>
<td>361</td>
<td>84%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>putative serine/threonine-protein kinase</td>
</tr>
<tr>
<td>XP_006471015.1</td>
<td>orange1.1g18734m</td>
<td>Cs</td>
<td>8</td>
<td>351</td>
<td>87%</td>
<td>92%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK2A-like</td>
<td></td>
</tr>
<tr>
<td>NP_001052827.1</td>
<td>LOC_Os04g35240.1</td>
<td>Os</td>
<td>8</td>
<td>359</td>
<td>83%</td>
<td>93%</td>
<td>99%</td>
<td>0</td>
<td>Os04g032000</td>
<td></td>
</tr>
<tr>
<td>XP_004230475.1</td>
<td>Soly01g103940.2.1</td>
<td>Sl</td>
<td>9</td>
<td>361</td>
<td>88%</td>
<td>94%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK2B</td>
<td></td>
</tr>
<tr>
<td>NP_001274892.1</td>
<td>PGSC0003DMG400023803</td>
<td>St</td>
<td>9</td>
<td>360</td>
<td>91%</td>
<td>96%</td>
<td>99%</td>
<td>0</td>
<td>serine/threonine-protein kinase SRK2B-like</td>
<td></td>
</tr>
<tr>
<td>XP_002269221.1</td>
<td>GSVIVT0102333901</td>
<td>Vv</td>
<td>8</td>
<td>356</td>
<td>91%</td>
<td>95%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK2A</td>
<td></td>
</tr>
<tr>
<td><strong>Cc07_g05710</strong></td>
<td>NP_201489.1</td>
<td>AT5G66880.1</td>
<td>At</td>
<td>8</td>
<td>361</td>
<td>83%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>serine/threonine-protein kinase SRK2I</td>
</tr>
<tr>
<td>KDO49166.1</td>
<td>orange1.1g107933m</td>
<td>Cs</td>
<td>7</td>
<td>363</td>
<td>84%</td>
<td>93%</td>
<td>99%</td>
<td>0</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>NP_001050653.1</td>
<td>LOC_Os03g41460.1</td>
<td>Os</td>
<td>6</td>
<td>362</td>
<td>81%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>Os03g0610900</td>
<td></td>
</tr>
<tr>
<td>XP_004232055.1</td>
<td>Soly02g090390.2.1</td>
<td>Sl</td>
<td>8</td>
<td>352</td>
<td>86%</td>
<td>95%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK21</td>
<td></td>
</tr>
<tr>
<td>XP_006338224.1</td>
<td>PGSC0003DMG400025895</td>
<td>St</td>
<td>8</td>
<td>352</td>
<td>86%</td>
<td>95%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SRK21-like</td>
<td></td>
</tr>
<tr>
<td>XP_002284959.1</td>
<td>GSVIVT0101810600</td>
<td>Vv</td>
<td>8</td>
<td>363</td>
<td>82%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK10</td>
<td></td>
</tr>
<tr>
<td><strong>Cc07_g14700</strong></td>
<td>NP_567945.1</td>
<td>AT4G33950.1</td>
<td>At</td>
<td>9</td>
<td>362</td>
<td>73%</td>
<td>87%</td>
<td>93%</td>
<td>1.2E-173</td>
<td>calcium-independent ABA-activated protein kinase</td>
</tr>
<tr>
<td>XP_006466260.1</td>
<td>orange1.1g19628m</td>
<td>Cs</td>
<td>8</td>
<td>341</td>
<td>82%</td>
<td>90%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK2-like</td>
<td></td>
</tr>
<tr>
<td>NP_001060312.1</td>
<td>LOC_Os07g2940.1</td>
<td>Os</td>
<td>8</td>
<td>339</td>
<td>79%</td>
<td>88%</td>
<td>99%</td>
<td>0</td>
<td>Os07g0622000</td>
<td></td>
</tr>
<tr>
<td>XP_0010312635.1</td>
<td>Soly04g012160.2.1</td>
<td>Sl</td>
<td>8</td>
<td>345</td>
<td>80%</td>
<td>99%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK2</td>
<td></td>
</tr>
<tr>
<td>NP_001274912.1</td>
<td>PGSC0003DMG400023636</td>
<td>St</td>
<td>8</td>
<td>344</td>
<td>79%</td>
<td>87%</td>
<td>99%</td>
<td>0</td>
<td>serine/threonine-protein kinase SAPK2-like</td>
<td></td>
</tr>
<tr>
<td>XP_003652469.1</td>
<td>GSVIVT0100341900</td>
<td>Vv</td>
<td>8</td>
<td>338</td>
<td>83%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK2</td>
<td></td>
</tr>
<tr>
<td><strong>Cc08_g11200</strong></td>
<td>NP_567945.1</td>
<td>AT4G33950.1</td>
<td>At</td>
<td>9</td>
<td>362</td>
<td>72%</td>
<td>87%</td>
<td>95%</td>
<td>9.9E-180</td>
<td>calcium-independent ABA-activated protein kinase</td>
</tr>
<tr>
<td>XP_006470707.1</td>
<td>orange1.1g19433m</td>
<td>Cs</td>
<td>8</td>
<td>341</td>
<td>84%</td>
<td>91%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3-like</td>
<td></td>
</tr>
<tr>
<td>BAD17999.1</td>
<td>LOC_Os10g41490.1</td>
<td>Os</td>
<td>9</td>
<td>334</td>
<td>85%</td>
<td>93%</td>
<td>90%</td>
<td>0</td>
<td>serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td>XP_004245833.1</td>
<td>Soly08g07780.2.1</td>
<td>Sl</td>
<td>8</td>
<td>339</td>
<td>83%</td>
<td>91%</td>
<td>100%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td>XP_006359207.1</td>
<td>PGSC0003DMG400026211</td>
<td>St</td>
<td>8</td>
<td>339</td>
<td>82%</td>
<td>91%</td>
<td>100%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td>XP_002262726.1</td>
<td>GSVIVT0100489900</td>
<td>Vv</td>
<td>8</td>
<td>340</td>
<td>84%</td>
<td>90%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK3</td>
<td></td>
</tr>
<tr>
<td><strong>Cc10_g06790</strong></td>
<td>NP_974170.1</td>
<td>AT1G78290.3</td>
<td>At</td>
<td>5</td>
<td>343</td>
<td>83%</td>
<td>91%</td>
<td>88%</td>
<td>0</td>
<td>serine/threonine-protein kinase SRK2C</td>
</tr>
<tr>
<td>KDO81023.1</td>
<td>orange1.1g19628m</td>
<td>Cs</td>
<td>8</td>
<td>338</td>
<td>83%</td>
<td>90%</td>
<td>99%</td>
<td>0</td>
<td>hypothetical protein CISIN</td>
<td></td>
</tr>
<tr>
<td>NP_001050274.1</td>
<td>LOC_Os03g27280.1</td>
<td>Os</td>
<td>8</td>
<td>342</td>
<td>77%</td>
<td>87%</td>
<td>99%</td>
<td>0</td>
<td>Os03g0390200</td>
<td></td>
</tr>
<tr>
<td>XP_004237936.1</td>
<td>Soly04g074500.2.1</td>
<td>Sl</td>
<td>8</td>
<td>336</td>
<td>80%</td>
<td>90%</td>
<td>98%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK2</td>
<td></td>
</tr>
<tr>
<td>NP_001275016.1</td>
<td>PGSC0003DMG400030830</td>
<td>St</td>
<td>8</td>
<td>335</td>
<td>80%</td>
<td>90%</td>
<td>98%</td>
<td>0</td>
<td>serine/threonine-protein kinase SAPK2-like</td>
<td></td>
</tr>
<tr>
<td>XP_003634478.1</td>
<td>GSVIVT0100974001</td>
<td>Vv</td>
<td>8</td>
<td>335</td>
<td>81%</td>
<td>90%</td>
<td>99%</td>
<td>0</td>
<td>PREDICTED: serine/threonine-protein kinase SAPK2</td>
<td></td>
</tr>
</tbody>
</table>
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; Nilson and Assmann, 2007; Sirichandra et al., 2009). For example, applications of exogenous ABA was show to stimulate stomatal closure in the wiltly 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 CO₂, for example (Merilo et al., 2013). Under drought, some plant species maintain leaf water potential (isohydric behavior) while other favor stomatal conductance to maintain CO₂ assimilation (anisohydric behaviour). The first mechanism results of the enhancement of the ABA effect on stomatal conductance (gₛ) 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 D⁷ and D⁸ 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 DT and DS 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 DT and DS 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

DT (14, 73 and 120) and DS (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 µmol·m⁻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 DT (IAPAR59) and DS (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 DT and DS 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/m²/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)−ΔΔCt, where ΔCt target = Ct target gene – Ct reference gene and ΔΔCt = ΔCt target - ΔCt 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.
5(68/76)

([SUHVVLRQ SURILOHV RI &F3<\]332&-6Q5.2 JHQHV LQ KGURSRQLF-JURZQ SODQWV WUHDWHG ZLWK H[RJHQRXV S%$ 7KH H[SUHVVLRLQ SURILOHV RI WKH WULSDUWLWH VWPH &F3<\]332&-6Q5.2 JHQHV ZHUH DQDO\|H]G LQ FRIIHH SODQWV JURZLQI XQGHU KGURSRQLF FRQGGLWRQ DQG VXEPLWLW\HG WR S%$ WUHDPHQW ([LI\UH\(1 .IRU WKLV SXUSRVH, SODQWV RI &. FDQHSHKRUD DQG &. DUDELFD ZHUH LQFXEDWHG GXULQJ WKUHH GD\V LQ QXWULWH\HG VROXWLRQ FRQWDLQLQJ 500 0 RI S%$. /\DI VPDSOHV ZHUH FROO\FWHG IRU DUO SODQWV ( &. FDQHSHKRUD, 7: FORQHV 14 DQG 6: FORQHV 22; &. DUDELFD, 59 DQG 6: 5XEL\E) EHIRUH DVVD\ (FRQWURO, ZLWKRXW S%$) DQG DIWHU RQH (24 KRXUV) DQG WKUHH GD\V (72 KRXUV) XQGHU S%$ WHUHPHQW. 7K\HV VDP\SOHV ZHUH XVHG IRU T3&5 ([LI\UH 3 DQG \(L)\UH 4) [SULPHQWQ DOPLURFURSHQVDQDO\\UHVVHV ([LI\UH 5).

([SUHVVLRQ RI 3</ JHQHV $PRQJ WKH QLQH &F3</ JHQHV SUHYLRXVO\ LGHQWLILHG (FKDSWHU )\, 3<<2, \&F3<\7D, 3<\E DQG 3<\B ZHUH QRV H[SUHVWHG LQ OHDYHV RI &. FDQHSHKRUD DQG &. DUDELFD JHQ\RVSHV HLWKHU XQGHU FRQWURO RU S%$ WHUHPHQW. +RZHYHU, WKH 3<\51, 3<\4, 3<\8D, 3<\9) DQG 3<\9 JHQHV ZHUH H[SUHVWHG LQ ORZ O\YHO LQ OH\DYHV RI DOO FRIIHH JHQ\RVSHV JURZQ XQGHU KGURSRQLF FRQGGLWRQV ZLWKRXW S%$ ([LI\UH 2).

Q &. FDQHSHKRUD, XS-UHJXODWHHG[SUHVVLRLQ RI &F3<5$DQG &F3<\8E JHQHV ZDV FOHDO\|REVHUYHG DIWHU 24 KRXUV RI S%$ WHUHPHQW VSHFLILFDQO\ LQ OHDYHV RI WKH 7: FORQHV 14 EXW QRW LQ WKR\VH RI 6: FORQH
22. For both clones, leaf expression of \textit{PYR1}, \textit{PYL4}, \textit{PYL8a}, \textit{PYL8b} and \textit{PYL9} genes decrease hereafter to be undetectable at 72h of ABA treatment.

    In \textit{C. arabica}, \textit{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 \textit{CaPYR1}, \textit{CaPYL4}, \textit{CaPYL8a} and \textit{CaPYL9} genes was no more detected in I59 but observed in Rubi. In both genotypes, expression of \textit{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 \textit{PYL} expression profiles between D\textsuperscript{T} and D\textsuperscript{S} clones in each coffee species but also between \textit{C. canephora} and \textit{C. arabica} plantlets, mainly regarding the time-course of \textit{PYL} expression upon ABA treatments.
Figure 2 Expression profiles of PYL genes in leaves of C. canephora D³ (clone 14) and D⁸ (clone 22) and C. arabica D³ (I59) and D⁸ (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 chosen as preferential internal calibrator (RE=1).
Figure 3 Expression profiles of PP2C genes in leaves of C. canephora D5 (clone 14) and D5 (clone 22) and C. arabica D5 (I59) and D5 (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^7^ (clone 14) and D^8^ (clone 22) and *C. arabica* D^I59^ and D^Rubi^ (clone 19) 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 chosen 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\textsuperscript{T} and D\textsuperscript{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\textsuperscript{T} clone 14 than D\textsuperscript{S} clone 22 mainly for CcSnR2.7 and CcSnR2.11 genes, and to a lesser extent for CcSnRK2.8, and CcSnRK2.10. The contrary was observed for CcSnRK2.6 that had higher expression in D\textsuperscript{S} clone 22 than in D\textsuperscript{T} clone 14. Expression of CcSnRK2.7 and CcSnRK2.11 genes appeared greatly up-regulated in D\textsuperscript{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\textsuperscript{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\textsuperscript{S} cultivar IAPAR59. In the D\textsuperscript{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\textsuperscript{T} and D\textsuperscript{S} clones of C. canephora.

In leaves, the D\textsuperscript{T} and D\textsuperscript{S} plants of C. canephora presented differences in stomatal cell responses under ABA treatment (Figure 5).
\[ 1 \pm \frac{37 \times 5.2}{100} \times 100\% = \frac{195.9}{100} \times 100\% = 195.9\% \]

\[ \text{S influx} (\mu m^2) \]

\[ \text{ABA Treatment (Hours)} \]

<table>
<thead>
<tr>
<th>ABA Treatment (Hours)</th>
<th>S influx (\mu m^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>24</td>
<td>0.8</td>
</tr>
<tr>
<td>72</td>
<td>0.8</td>
</tr>
</tbody>
</table>

\[ \text{JXUH} (\text{YDODXLWRQ RI S\%S HIIHFW LQ PRGXODWLQJ WKH JXUDG FHOOV VWRPDWDO DSHUWXH LQ FRQHHOHGYH. 7KH VWRPDWDO} \]

\[ \text{DSHUWXH OHQWK ZDV PDHVXUHV LQ WKH JXUDG FHOOV RI \& FDQHJSKRU \}^2 \times 14 \text{ (FLUFOHV DQG \}^6 \text{ 22 FORQHV (WULDQJOHV) DQG \& DUDELFD }^7 \times 0.59 \text{ (VTXDUHV) DQG }^4 \times 5 \text{ (FURVVHV) FXOWLYDULQ LQ 24 DQG }^7 \text{ KRUXV DIWHU} \]

\[ \text{DSOSLDWLQRI } \text{HRIHJIKV} \text{ S\%S } \text{SOO } \text{ 0 VROXWLRQ DQG } \text{XQGHU FRQWURO \text{FRQGLWWLQ}Q} \text{(ZLWKXW S\%S, OK). 7KH} \]

\[ \text{VWRPDWDO DSHUWXH YDOXHV DUH JLYHQ DV DQ DLYHUJH RI D } \text{KXQGURG FHOOV PDHVXUHPHQW } \text{IRU HDERF } \text{FORQH/UEWLYDU.} \]

\[ \text{SOGHU FRQWURO \text{FRQGLWWLQ}Q, WKH } \text{\&. FDQHJSKRU \}^7 \times 0.59 \text{ FORQHV 14 VKRZHG LQYHUDJH ODUHJU VWRPDWDO} \]

\[ \text{DSHUWXH WKDQ }^6 \text{ FORQHV 22 (JXUH) S\%S } \text{SIWHU 24K DQG 72K RI S\%S WUHDPHQW, QR VLIQLILFQDW GLIIHIIQFHI} \]

\[ \text{ZDV REVHUYHG EHWZHHQ }^7 \text{ DQG }^6 \text{ FORQHV }^7 \text{ KHQ DQDO\}H \text{VIHSDUDWHO, HLWKHU FORQHV VKRZHG VLIQLILFQDW} \]

\[ \text{UHVSRQHQVVR \text{WR S\%S WUHDPHQWQWV. SIWHU WKH ILUWV 24K, WKH FORQHV }^7 \text{ 14 DQG }^6 \text{ 22 VKRZHG DQ LQFUHDHV} \text{RI} \]

\[ \text{VWRPDWDO FORVXUH (GHUHDFQ LQ JXUDG FHOOV VWRPDWDO DSHUWXH). SIWHU 72K, WKLV SURFHFW ZDV IXOO UHYHUVHG} \]

\[ \text{LQ }^6 \text{ FORQHV 22, ZKHUH QRQH VLIQLILFQDW GLIIHIIQFHI ZDV REVHUYHG LQ VWRPDWDO DSHUWXH EHWZHHQ OK DQG} \]

\[ \text{72K, DQG SUDWLDO }^0 \text{ UHYHUVHG LQ }^7 \text{ FORQHV 14, ZKHUH GHVSLWH RI VWRPDWDO DSHUWXH LQFUHDHV VLIQLILFQDW} \]

\[ \text{GLIIHIIQFHI ZVHVLWLO REVHUYHG EHWZHHQ OK DQG }^7 \text{22QR }^7 \text{ FORQHV DQG LQFUHDHV RI} \]

\[ \text{VWRPDWDO FORVXUH (GHUHDFQ LQ JXUDG FORQHV S\%S UHYHUVHG LQ }^6 \text{ WR S\%S QDVHQFH} \]

\[ \text{ZDV REVHUYHG EHWZHHQ S\%S WUHDPHQW, QR VLIQLILFQDW GLIIHIIQFHI ZDV REVHUYHG LQ VWRPDWDO DSHUWXH EHWZHHQ OK DQG} \]

\[ \text{72K, DQG SUDWLDO }^0 \text{ UHYHUVHG LQ }^7 \text{ FORQHV 14, ZKHUH GHVSLWH RI VWRPDWDO DSHUWXH LQFUHDHV VLIQLILFQDW} \]

\[ \text{GLIIHIIQFHI ZVHVLWLO REVHUYHG EHWZHHQ OK DQG }^7 \text{22QR }^7 \text{ FORQHV DQG LQFUHDHV RI} \]

\[ \text{VWRPDWDO FORVXUH (GHUHDFQ LQ JXUDG FORQHV S\%S UHYHUVHG LQ }^6 \text{ WR S\%S QDVHQFH} \]

\[ \text{ZDV REVHUYHG EHWZHHQ S\%S WUHDPHQW, QR VLIQLILFQDW GLIIHIIQFHI ZDV REVHUYHG LQ VWRPDWDO DSHUWXH EHWZHHQ OK DQG} \]

\[ \text{72K, DQG SUDWLDO }^0 \text{ UHYHUVHG LQ }^7 \text{ FORQHV 14, ZKHUH GHVSLWH RI VWRPDWDO DSHUWXH LQFUHDHV VLIQLILFQDW} \]

\[ \text{GLIIHIIQFHI ZVHVLWLO REVHUYHG EHWZHHQ OK DQG }^7 \text{22QR }^7 \text{ FORQHV DQG LQFUHDHV RI} \]

\[ \text{VWRPDWDO FORVXUH (GHUHDFQ LQ JXUDG FORQHV S\%S UHYHUVHG LQ }^6 \text{ WR S\%S QDVHQFH} \]

\[ \text{ZDV REVHUYHG EHWZHHQ S\%S WUHDPHQW, QR VLIQLILFQDW GLIIHIIQFHI ZDV REVHUYHG LQ VWRPDWDO DSHUWXH EHWZHHQ OK DQG} \]

\[ \text{72K, DQG SUDWLDO }^0 \text{ UHYHUVHG LQ }^7 \text{ FORQHV 14, ZKHUH GHVSLWH RI VWRPDWDO DSHUWXH LQFUHDHV VLIQLILFQDW} \]

\[ \text{GLIIHIIQFHI ZVHVLWLO REVHUYHG EHWZHHQ OK DQG }^7 \text{22QR }^7 \text{ FORQHV DQG LQFUHDHV RI} \]

\[ \text{VWRPDWDO FORVXUH (GHUHDFQ LQ JXUDG FORQHV S\%S UHYHUVHG LQ }^6 \text{ WR S\%S QDVHQFH} \]

\[ \text{ZDV REVHUYHG EHWZHHQ S\%S WUHDPHQW, QR VLIQLILFQDW GLIIHIIQFHI ZDV REVHUYHG LQ VWRPDWDO DSHUWXH EHWZHHQ OK DQG} \]

\[ \text{72K, DQG SUDWLDO }^0 \text{ UHYHUVHG LQ }^7 \text{ FORQHV 14, ZKHUH GHVSLWH RI VWRPDWDO DSHUWXH LQFUHDHV VLIQLILFQDW} \]

\[ \text{GLIIHIIQFHI ZVHVLWLO REVHUYHG EHWZHHQ OK DQG }^7 \text{22QR }^7 \text{ FORQHV DQG LQFUHDHV RI} \]

\[ \text{VWRPDWDO FORVXUH (GHUHDFQ LQ JXUDG FORQHV S\%S UHYHUVHG LQ }^6 \text{ WR S\%S QDVHQFH} \]
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 Dᵀ 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* D⁵ Rubi, the HAB gene was upregulated after 24 h ABA treatment and the expression levels continue to increase at 72 h. The D⁵ 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 D⁵ 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* Dᵀ clone 14 but not in D⁵ clone 22. In the first 24h guard cells of *C. arabica* Dᵀ 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 Dᵀ clone 14 and I59 has been more efficient in the stomatal regulation under ABA exogenous treatment than D⁵ 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 Dᵀ 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 DT clone 14 presented higher expression for the AHG2, AHG3, HAI (PP2Cs) and SnRK2.7, SnRK2.11 (SnRK2s) compared to the DS clone 22 under control conditions which suggest the existence of the tripartite system (mode off) ready to be activated in the DT 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 DS 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 DT clone 14 and the DS 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 DT I59 had a faster response to ABA stimuli compared to the DS Rubi. With 24h it was observed that the DT 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 DT (clone 14 and I59) compared to DS (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


CHAPTER 2: GENE EXPRESSION PROFILES IN COFFEA ARABICA AND COFFEA CANEPHORA LEAVES REVEALED TRANScriptionAL REGULATIONS OF KEY GENES INVOLVED IN ABA SIGNALING


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 \( \Psi_{pd} \) 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 \( \text{D}^7 \) and \( \text{D}^5 \) clones, as has recently been observed for \( \text{CcDREB1D} \) gene of \( \text{C. canephora} \) (Alves et al., submitted).

Finally, those genes with higher correlated drought-induced expression identified during this work (e.g. \( \text{CcAHG2} \) and \( \text{CcSnRK2.2} \)) could be tested in other \( \text{C. canephora} \) 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
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
© Springer Science+Business Media Dordrecht 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

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 2000), 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 infl 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 Fxalt 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 (alloetraploid, 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 diterpenes (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. 2010). 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′S–47°43′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 (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 (24:1:1) was then added before removal of cell debris by centrifugation (10 min, 15,000g). 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 (10 min, 15,000g). 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, 15,000g) 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-32P-dCTP as described by the supplier (GE Healthcare). Total RNAs were transferred to Hybond N+ membranes which
Table 1 List of primers

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSP1</td>
<td>5'-CAGATCCACCAGCAACAGTACAACC-3'</td>
</tr>
<tr>
<td>GSP2</td>
<td>5'-CAGTGCAACCCCCAGATGATTCTC-3'</td>
</tr>
<tr>
<td>LTP-F1</td>
<td>5'-G GCCCTTTTGGTTTCTCATAAAGAT-3'</td>
</tr>
<tr>
<td>LTP-F2</td>
<td>5'-GACCTTTTTGTTTCTCATAATGATG-3'</td>
</tr>
<tr>
<td>LTP-F3</td>
<td>5'-GAAATCCTGGGGTGTTGCG-3'</td>
</tr>
<tr>
<td>LTP-F4</td>
<td>5'-AAGGATGCTCAATGCT-3'</td>
</tr>
<tr>
<td>LTP-R1</td>
<td>5'-ATTCAACACCATTACTAGTTTTCGAGC-3'</td>
</tr>
<tr>
<td>LTP-R2</td>
<td>5'-CACCATTAATGGGAAGTGTCG-3'</td>
</tr>
<tr>
<td>LTP-F1</td>
<td>5'-CTTGGTCGCTAAATGCGCAACT-3'</td>
</tr>
<tr>
<td>BUBI-F</td>
<td>5'-AAGGATGCTCAACAGATGACAT-3'</td>
</tr>
<tr>
<td>BUBI-R</td>
<td>5'-GGGACAGCCTTGCTGACTATA-3'</td>
</tr>
<tr>
<td>F1-pBI</td>
<td>5'-CGGAAGCTTGCATCTAAATACATTCG-3'</td>
</tr>
<tr>
<td>F2-pBI</td>
<td>5'-CGGAAGCTTCATGAAATGCAATCC-3'</td>
</tr>
<tr>
<td>F3-pBI</td>
<td>5'-CGGAAGCTTCCAAGACATTATATAGATG-3'</td>
</tr>
<tr>
<td>F4-pBI</td>
<td>5'-CGGAAGCTTCTCCACACTTTCACAAATCTGG-3'</td>
</tr>
<tr>
<td>R1-pBI</td>
<td>5'-CGGGATCCGAAAAGCAAATGCGAAGAGAG-3'</td>
</tr>
<tr>
<td>FORmax</td>
<td>5'-GCCAGGGTTTTCCCAGTCACGACGTTGTAA-3'</td>
</tr>
<tr>
<td>REVmax</td>
<td>5'-CACAACAGGAAACAGATGACATGATTTCTTC-3'</td>
</tr>
</tbody>
</table>

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 (dT15) 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 (SYBR Green qPCR Mix-UDG/ROX, Invitrogen). The reaction mixture was incubated for 2 min at 50 °C (Uracil DNA-Glycosylase 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(1−1/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)−ΔΔCt where ΔΔCt = Ct target gene − Ct CcUBQ10 and ΔΔCt = ΔCt target − ΔCt 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 CalLTP1a, CalLTP2 and CalLTP3a 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 Tag 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
SGNCaU607388  -aagaacacgttgaacatcaccatcagaaaattttctctctctctcctctcttgc 59
SGNCcU613906  cagcagg--gtcgaacatcaccatcagaaaattttctctctctctcctctcttgc 58
SGNCaU610393  -aagaacacgtaaacatcaccatcagaaaattttctctctctctcctctcttgc 59
Contig22413  ggccatctgtaaacatcaccatcagaaaattttctctctctctcctctcttgc 45

SGNCaU607388  ttttcataaag
SGNCcU613906  ttttcataaat
SGNCaU610393  ttttcataaat

Contig22413  ATG
SGNCaU610393  ATG
SGNCcU613906  ATG

SGNCaU607388  GTG
SGNCcU613906  GCT
SGNCaU610393  GCT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU613906  GGCGTTGCTGGCCTGACCGCGCTGGCTGCAATGGCAACCATCAG
SGNCcU613906  GGCGTTGCTGGCCTGACCGCGCTGGCTGCAATGGCAACCATCAG
SGNCaU610393  GGCGTTGCTGGCCTGACCGCGCTGGCTGCAATGGCAACCATCAG

Contig22413  TGGCGTGGCTGGACCGCGCTGGCTGCAATGGCAACCATCAG
SGNCaU610393  TGGCGTGGCTGGACCGCGCTGGCTGCAATGGCAACCATCAG
SGNCcU613906  TGGCGTGGCTGGACCGCGCTGGCTGCAATGGCAACCATCAG

Contig22413  TTATAAAAAATCCCAATACCAAAAAACGCTTTGAAGCCTGTGGTCTGAAATGGCCAACTT
SGNCaU610393  TTATAAAAAATCCCAATACCAAAAAACGCTTTGAAGCCTGTGGTCTGAAATGGCCAACTT
SGNCcU613906  TTATAAAAAATCCCAATACCAAAAAACGCTTTGAAGCCTGTGGTCTGAAATGGCCAACTT

Contig22413  ATGCCAAGGAACAGGAGCCGTGCTTCTGCAATTTTATCAAAGATCCAGCATACCGGCAAAA
SGNCaU610393  ATGCCAAGGAACAGGAGCCGTGCTTCTGCAATTTTATCAAAGATCCAGCATACCGGCAAAA
SGNCcU613906  ATGCCAAGGAACAGGAGCCGTGCTTCTGCAATTTTATCAAAGATCCAGCATACCGGCAAAA

Contig22413  GAT
SGNCaU610393  GAT
SGNCcU613906  GAT

SGNCaU613906  TCCTGTCATCTGCCTCCTCCTCCTCAAAAATCGCTCTTCACTTTTATCAAAGATCCAGCAG
SGNCcU613906  TCCTGTCATCTGCCTCCTCCTCCTCAAAAATCGCTCTTCACTTTTATCAAAGATCCAGCAG
SGNCaU610393  TCCTGTCATCTGCCTCCTCCTCCTCAAAAATCGCTCTTCACTTTTATCAAAGATCCAGCAG

Contig22413  GCC
SGNCaU610393  GCC
SGNCcU613906  GCC

SGNCaU610393  GCT
SGNCcU613906  GCT
SGNCaU607388  GCT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  TGAT
SGNCcU613906  TGAT
SGNCaU610393  TGAT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  TGAT
SGNCcU613906  TGAT
SGNCaU610393  TGAT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  GCT
SGNCcU613906  GCT
SGNCaU610393  GCT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  AAT
SGNCcU613906  AAT
SGNCaU610393  AAT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  AG
SGNCcU613906  AG
SGNCaU610393  AG

Contig22413  GCT
SGNCaU610393  GCT
SGNCcU613906  GCT

SGNCaU613906  TAA
SGNCcU613906  TAA
SGNCaU610393  TAA

Contig22413  ATG
SGNCaU610393  ATG
SGNCcU613906  ATG

SGNCaU607388  TGAT
SGNCcU613906  TGAT
SGNCaU610393  TGAT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  TGAT
SGNCcU613906  TGAT
SGNCaU610393  TGAT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

SGNCaU607388  TGAT
SGNCcU613906  TGAT
SGNCaU610393  TGAT

Contig22413  TAA
SGNCaU610393  TAA
SGNCcU613906  TAA

**Fig. 1** Alignment of nsLTP-encoding nucleic sequences found in public databases. Contig22413 was found in the database of the Coffea Genome Project (http://bioinfo04.ibi.unicamp.br). Contigs SGN-CaU607388, SGN-CaU613906 and SGN-CaU610393 were found in the SOL genomic database (http://solgenomics.net). The nsLTP-encoding sequences are 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. Catuai 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-chloro-3-indolyl-β-D-glucuronide) solution (100 mM NaH₂PO₄, 500 mM EDTA, 0.1 % Triton X-100, 0.5 mM K₃Fe(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 Na₄EDTA, 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
Characterization of nsLTP-encoding cDNA and gene sequences

Electronic Northern and Fisher’s statistical tests based on the Coffee Genome Project data (http://bioinfofo04.ibi.uni-camp.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 *Coffeea 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 clustering 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. Mundo Novo and *C. canephora*. From *C. arabica*, four sequences were isolated: *CalTP1a*...
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 (pl) 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
they contained a putative signal peptide with a cleavage site between the amino acid residues A29 and V30. As pre-proteins, the CaLTP1a (deduced from CaLTP1a cDNA and gene) and CaLTP2 (deduced from CaLTP2cDNA 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 CaLTP3 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: 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 (C1 to C8), the C5XC6 motif (characterized by the presence of hydrophobic residues such as leucine or phenylalanine at position 5 of the full-length proteins) and followed the protein pattern CX4LX2CX9-11P[S,T] X2CCX5QX2-4[C,L,F]CX2[A,L,I]X[D,N]PX10-12[K,R] X4-5CX3-4PX0-2C.

Phylogenetic analysis reveals that all coffee 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 hydrophilic. 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 properties as well as the localization in protein structure. The replaced amino acids conserved the physico-chemical properties 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 LTP3a 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, the expression of both the LTP3 and LTP1–LTP2 genes was high at 90 and 120 DAF and undetectable in the latest stages of maturation.

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
Fig. 5 nsLTP Alignments and 3D prediction. A Alignment of the 7 coffee nsLTP together with 1TUK (the wheat nsLTP used for modelsation). 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 (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 (TATAAT) 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).
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 (GRWAAA) 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. 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. Catuai 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.
Analysis of LTP promoter in transgenic tobacco plants

A deletion analysis was carried out to precisely define 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

![Diagrammatic representation of the nsLTP promoter. The fragments amplified from the genomic DNA of C. arabica cv. Catuai 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](image-url)
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.
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 a 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 pl. 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 *Coffeea* 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 (103-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 diterpenes (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, caeleosin and stereooleosin) (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 TGAC 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 RF 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 erdl (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 CNPITA) 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 macroproject program. 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 Sv73812). The 3′ end fragment (411 bp) of the *LTP* promoter was patented (BR1020120081628) by INPI-Brazil.

References


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


Identification of candidate genes for drought tolerance in coffee by high-throughput sequencing in the shoot apex of different Coffea arabica cultivars

Luciana Souto Mofatto1†, Femanda de Araújo Carneiro2†, Natalia Gomes Vieira2†, Karoline Estefani Duarte2, Ramon Oliveira Vidal1, Jean Carlos Alekevetch2, Michelle Guighton Cotta2, Jean-Luc Verdeil3, Fabienne Lapeyre-Montes3, Marc Lartaud3, Thierry Leroy3, Fabien De Bellis3, David Pot3, Gustavo Costa Rodrigues4, Marcelo Falsarella Carazzolle1, Gonçalo Amarante Guimarães Pereira1, Alan Carvalho Andrade2,5 and Pierre Marraccini2,3*

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)
(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, *Coffeea 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 *Coffeea 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 *Coffeea sp*. In a previous study analyzing 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 *Coffeea 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 *Coffeea 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 HT83/2 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$_2$O cm$^{-1}$.

Sampling
For both cultivars and experiments, leaf predawn water potentials ($\Psi_{pd}$) 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 $\Psi_{pd}$ measurements were also collected from the same plants. At the end of the 2011 dry season, $\Psi_{pd}$ 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).

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 80 (20 for each library) candidate genes (CGs) in plagiotropic buds of Rubi and IAPAR59. The best primer pairs were selected based on (1) differentially expressed genes in plagiotropic buds of Rubi and IAPAR59 (RUB-C/I59-C) and (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.

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) x 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)^∆∆Ct where ∆Ct target = Ct target gene – Ct reference gene and ∆∆Ct = ∆Ct target – ∆Ct 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 (Qiimaging 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

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein name</th>
<th>C. canephora</th>
<th>GB</th>
<th>ATP</th>
<th>SGN</th>
<th>Primer Primer sequences</th>
<th>bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaUBQ10</td>
<td>Ubiquitin</td>
<td>Cc02_031600</td>
<td>GW488515</td>
<td>32782</td>
<td>U637098</td>
<td>BUBI-FBUBI-R</td>
<td>104</td>
</tr>
<tr>
<td>CaAEP1a</td>
<td>Putative aldose 1-epimerase</td>
<td>Cc07_003170</td>
<td>GT005185</td>
<td>716</td>
<td>U637659</td>
<td>716-1F716-1R</td>
<td>75</td>
</tr>
<tr>
<td>CaCAB2a</td>
<td>Chlorophyll a/b-binding protein</td>
<td>Cc09_009030</td>
<td>GT003492</td>
<td>33540</td>
<td>U629601</td>
<td>48565-F48565-R</td>
<td>100</td>
</tr>
<tr>
<td>CaCHI1a</td>
<td>Class III chitinase</td>
<td>Cc11_00410</td>
<td>GT012279</td>
<td>32745</td>
<td>U637166</td>
<td>50103-F50103-R</td>
<td>70</td>
</tr>
<tr>
<td>CaCHI2a</td>
<td>Putative chitinase</td>
<td>Cc00_014300</td>
<td>GT011845</td>
<td>32737</td>
<td>U638035</td>
<td>53058-F53058-R</td>
<td>70</td>
</tr>
<tr>
<td>CaCHI3a</td>
<td>Chitinase-like protein</td>
<td>Cc03_013720</td>
<td>GW491433</td>
<td>32875</td>
<td>U645893</td>
<td>23638-F23638-R</td>
<td>130</td>
</tr>
<tr>
<td>CaDLP1a</td>
<td>Dirigent-like protein</td>
<td>Cc00_027410</td>
<td>GW477731</td>
<td>35149</td>
<td>nfd</td>
<td>39577-F39577-R</td>
<td>70</td>
</tr>
<tr>
<td>CaELIP3a</td>
<td>Early light-induced protein (ELIP)</td>
<td>Cc03_004300</td>
<td>GR985685</td>
<td>32771</td>
<td>U631550</td>
<td>32771-F32771-R</td>
<td>100</td>
</tr>
<tr>
<td>CaGAS2a</td>
<td>Glucosyltransferase arbutin synthase</td>
<td>Cc02_039100</td>
<td>GT697284</td>
<td>3945</td>
<td>U632419</td>
<td>632419-F632419-R</td>
<td>101</td>
</tr>
<tr>
<td>CaGRP2a</td>
<td>Glycin-rich protein</td>
<td>Cc00_016260</td>
<td>GW430980</td>
<td>32799</td>
<td>U635030</td>
<td>53139-1 F53139-1R</td>
<td>100</td>
</tr>
<tr>
<td>CaH2Aa</td>
<td>Putative histone H2A</td>
<td>Cc01_012440</td>
<td>GT723387</td>
<td>33557</td>
<td>U630412</td>
<td>53417-F53417-R</td>
<td>80</td>
</tr>
<tr>
<td>CaHSP3a</td>
<td>Heat shock protein (HSP) 70 kDa</td>
<td>Cc02_008040</td>
<td>GR982512</td>
<td>33197</td>
<td>U636531</td>
<td>33197-1 F33197-1R</td>
<td>100</td>
</tr>
<tr>
<td>CaIPS1a</td>
<td>Myo-inositol 1-phosphate synthase</td>
<td>Cc07_015530</td>
<td>GT003538</td>
<td>10496</td>
<td>U632517</td>
<td>10496-1 F10496-1R</td>
<td>100</td>
</tr>
<tr>
<td>CaJAMT1a</td>
<td>Jasmonate-O-methyltransferase</td>
<td>Cc03_007330</td>
<td>GR989151</td>
<td>33008</td>
<td>U631389</td>
<td>47327-F47327-R</td>
<td>100</td>
</tr>
<tr>
<td>CaMAS1a</td>
<td>Methylamine-A synthase</td>
<td>Cc00_013640</td>
<td>GW479615</td>
<td>33413</td>
<td>nfd</td>
<td>33413-F33413-R</td>
<td>60</td>
</tr>
<tr>
<td>CaPP2a</td>
<td>Putative phloem protein 2 (PP2)</td>
<td>Cc03_013000</td>
<td>GR995691</td>
<td>33207</td>
<td>U633544</td>
<td>33207-1 F33207-1R</td>
<td>90</td>
</tr>
<tr>
<td>CaPSB8a</td>
<td>Photosystem II CP47 (psbB)-like protein</td>
<td>Cc03_001300</td>
<td>GR995691</td>
<td>33207</td>
<td>U633544</td>
<td>33207-1 F33207-1R</td>
<td>80</td>
</tr>
<tr>
<td>CaSAMT2a</td>
<td>S-adenosyl-methionine-methyltransferase</td>
<td>Cc03_005630</td>
<td>DV672716</td>
<td>754</td>
<td>U629783</td>
<td>34318-F34318-R</td>
<td>80</td>
</tr>
<tr>
<td>CaSDC1a</td>
<td>S-adenosyl-L-methionine decarboxylase</td>
<td>Cc11_011130</td>
<td>GT002431</td>
<td>8508</td>
<td>U629687</td>
<td>8508-1F8508-1R</td>
<td>100</td>
</tr>
<tr>
<td>CaSLP1i</td>
<td>Subtilisin-like protein</td>
<td>Cc00_019100</td>
<td>GW430663</td>
<td>1620</td>
<td>nfd</td>
<td>7961-F7961-R</td>
<td>80</td>
</tr>
<tr>
<td>CaSTK1i</td>
<td>Hypothetical S/T protein kinase</td>
<td>Cc00_018670</td>
<td>GT687049</td>
<td>6301</td>
<td>U631794</td>
<td>6301-1F6301-1R</td>
<td>80</td>
</tr>
<tr>
<td>CaUNK1i</td>
<td>Unknown protein 1</td>
<td>Cc03_008880</td>
<td>DV689820</td>
<td>33062</td>
<td>U614843</td>
<td>182052-F182052-1R</td>
<td>79</td>
</tr>
</tbody>
</table>
Table 1 Candidate genes and corresponding primers used for qPCR experiments (Continued)

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Accession</th>
<th>Description</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaUNK2(^a)</td>
<td>Unknown protein 2</td>
<td>CC07_g01940 DV708962</td>
<td>31492 U637447 33353-33353-R</td>
</tr>
<tr>
<td>CaUNK3(^b)</td>
<td>Unknown protein 3</td>
<td>nf nf</td>
<td>22823 nf 22823-F22823-R</td>
</tr>
<tr>
<td>CaUNK4(^b)</td>
<td>Unknown protein 4</td>
<td>CC06_g11210 GW465088</td>
<td>39984 nf 55677-F55677-R</td>
</tr>
<tr>
<td>CaUNK5(^b)</td>
<td>Unknown protein 5</td>
<td>CC08_g09510 GW474926</td>
<td>4578 nf 4578-F4578-R</td>
</tr>
<tr>
<td>CaUNK6(^b)</td>
<td>Unknown protein 6</td>
<td>CC03_g06850 GT002178</td>
<td>34993 U632634 39984-F39984-R</td>
</tr>
<tr>
<td>CaUNK7(^b)</td>
<td>Unknown protein 7</td>
<td>CC03_g00560 GW444736</td>
<td>33613 U631416 25639-F25639-R</td>
</tr>
<tr>
<td>CaUNK8(^b)</td>
<td>Unknown protein 8</td>
<td>CC00_g04970 DV695331</td>
<td>33190 U640780 55689-F55689-R</td>
</tr>
<tr>
<td>CaUNK9(^b)</td>
<td>Unknown protein 9</td>
<td>CC03_g08920 GT649500</td>
<td>32762 U636808 30926-F30926-R</td>
</tr>
<tr>
<td>CaUNK10(^b)</td>
<td>Unknown protein 10</td>
<td>nf GT648004</td>
<td>14813 U645073 D18240-FLP18100-R</td>
</tr>
<tr>
<td>CaUNK11(^b)</td>
<td>Unknown protein 11</td>
<td>CC03_g14330 GR991912</td>
<td>8598 U637116 55677-F55677-R</td>
</tr>
<tr>
<td>CaUNK12(^b)</td>
<td>Unknown protein 12</td>
<td>CC10_g12840 nf</td>
<td>53029 nf 53029-F53029-R</td>
</tr>
<tr>
<td>CaUNK13(^b)</td>
<td>Unknown protein 13</td>
<td>CC00_g17760 GT673421</td>
<td>14198 U639484 33980-F33980-R</td>
</tr>
<tr>
<td>CaUNK14(^b)</td>
<td>Unknown protein 14</td>
<td>CC00_g16260 GT672564</td>
<td>48325 U635030 42747-F42747-R</td>
</tr>
<tr>
<td>CaUNK15(^b)</td>
<td>Unknown protein 15</td>
<td>CC00_g04970 GR983286</td>
<td>33190 U636790 30926-F30926-R</td>
</tr>
<tr>
<td>CaUNK16(^b)</td>
<td>Unknown protein 16</td>
<td>nf GW464209</td>
<td>9761 U639049 18112-F18112-R</td>
</tr>
<tr>
<td>CaUNK17(^b)</td>
<td>Unknown protein 17</td>
<td>CC03_g08920 GT685623</td>
<td>32762 U636808 42747-F42747-R</td>
</tr>
<tr>
<td>ColTP1(^a), ColTP2(^b)</td>
<td>Non-specific lipid transfer protein (nsLTP)</td>
<td>CC11_g09700 HG008739HG008740</td>
<td>46897 U632702 LTP-R1LTP-FT</td>
</tr>
<tr>
<td>ColTP3(^d)</td>
<td>Non-specific lipid transfer protein (nsLTP)</td>
<td>CC04_g06890 HG008741</td>
<td>3368 U632702 LTP-R1LTP-FT</td>
</tr>
<tr>
<td>LTP(^d)</td>
<td>nf</td>
<td>- U632702 LTP-F100LTP-R100</td>
<td>5(^\prime) TGCAATTTCATCAAGATCCGCAACG 3(^\prime) 5(^\prime) AGTTGGCCATTGCAGACCAAC 3(^\prime) 93</td>
</tr>
</tbody>
</table>

\(^a\) 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 SG\(N\) [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 (SG\(N\): tools/blast/SG\(N\) Clusters [current version] / Coffee species Clusters, GB: BLASTn/Nucleotide collection [nr/nt]). The size of amplicons is based on the unigene. \(^b\): candidate genes (n = 20) identified during this study. \(^c\): orphan genes (n = 14) previously described [35] and analysed in this study. \(^d\): 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 ≤ 0.05 was considered significant for all the statistical analyses.

Results

Monitoring drought under field conditions

In 2009, leaf predawn water potential (Ψpd) 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 Ψpd 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 Ψpd values in IAPAR59 indicated that it had better access to soil water. The first rains then occurred and the Ψpd values of drought-stressed plants increased almost to those measured in irrigated plants, illustrating the complete recovery of stressed plants. In 2011, Ψpd was measured at the peak of the drought (end of dry season). Under drought conditions, both Rubi and IAPAR59 had similar Ψpd values that were more negative than those measured in 2009, indicating more severe drought stress in 2011 (Fig. 2b).

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

<table>
<thead>
<tr>
<th>Libraries</th>
<th>Total reads</th>
<th>Trimmed reads</th>
<th>Average length of reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Sanger</td>
<td>195,110</td>
<td>151,403</td>
<td>518</td>
</tr>
<tr>
<td>IS9-C</td>
<td>135,304</td>
<td>66,641</td>
<td>325</td>
</tr>
<tr>
<td>IS9-D</td>
<td>282,213</td>
<td>112,518</td>
<td>351</td>
</tr>
<tr>
<td>RUB-C</td>
<td>230,064</td>
<td>101,394</td>
<td>360</td>
</tr>
<tr>
<td>RUB-D</td>
<td>345,751</td>
<td>153,572</td>
<td>342</td>
</tr>
<tr>
<td>Total</td>
<td>1,188,442</td>
<td>585,528</td>
<td>379.2</td>
</tr>
</tbody>
</table>

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 IS9: 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 (0.74 % of total clusters), 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\textsuperscript{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\textsuperscript{-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\textsuperscript{-121}, 66 % identity, 82 %, similarity) with monomilacone 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

<table>
<thead>
<tr>
<th>Libraries</th>
<th>EdgeR DEG (% of total clusters)</th>
<th>DEseq DEG (% of total clusters)</th>
<th>Total DEG (% of total clusters)</th>
<th>Up-regulated clusters (% of total clusters)</th>
<th>Down-regulated clusters (% of total clusters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I59-D/I59-C</td>
<td>209 (0.49%)</td>
<td>176 (0.41%)</td>
<td>226 (0.52%)</td>
<td>165 (0.38%)</td>
<td>61 (0.14%)</td>
</tr>
<tr>
<td>RUB-D/RUB-C</td>
<td>323 (0.75%)</td>
<td>306 (0.71%)</td>
<td>343 (0.80%)</td>
<td>251 (0.58%)</td>
<td>92 (0.21%)</td>
</tr>
<tr>
<td>RUB-C/I59-C</td>
<td>173 (0.40%)</td>
<td>169 (0.39%)</td>
<td>184 (0.43%)</td>
<td>104 (0.24%)</td>
<td>80 (0.19%)</td>
</tr>
<tr>
<td>RUB-D/I59-D</td>
<td>392 (0.91%)</td>
<td>433 (1.00%)</td>
<td>490 (1.14%)</td>
<td>320 (0.74%)</td>
<td>170 (0.39%)</td>
</tr>
</tbody>
</table>

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.
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 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 (A1310148), 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).
Drought influences leaf cuticle thickness

Values of leaf cuticle thickness are given in Table 4. Bars = 20 μm

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

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...
Table 4 Influence of drought on leaf cuticle thickness

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Cuticle thickness (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>IAPAR59: 1.49±0.19&lt;sup&gt;(a)&lt;/sup&gt; Rubi: 1.75±0.15&lt;sup&gt;(b)&lt;/sup&gt;</td>
</tr>
<tr>
<td>Drought</td>
<td>IAPAR59: 1.98±0.19&lt;sup&gt;(c)&lt;/sup&gt; Rubi: 1.73±0.28&lt;sup&gt;(b)&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

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
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; Σ 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 (Ψp) 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 Sadenosyl-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 determination of drought tolerance in coffee.

Another interesting response concerned the differential expression of the CalMAS1 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 CalMAS1 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 CalAMT1 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 *TeSAMDC* 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, D1 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*Es* (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 CaLPS1 (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 D7 (IAPAR59) and D5 (Rubi) cultivars grown under control and drought conditions. Major differences between these plants concerned their phenotypic behaviour (e.g. predawn leaf water potential, \(\Psi_{pd}\)) 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 epigenetic 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 (I99: 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 Aymibér 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 Fivet-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 (KEI, NGV), Brazilian Consortium of Coffee Research (FAC, JCA) and CAPES-COFECUB Project 5v738-12 (MGCC).

Author details
1Laboratório de Genômica e Expressão (LGE), Departamento de Genética e Evolução, Instituto de Biologia/UNICAMP, Cidade Universitária Zeferino Vaz, 13083-970 Campinas, SP, Brazil. 2Embrapa Recursos Genéticos e Biotecnologia (LGM-NTBio), Parque Estação Biológica, CP 02372, 70770-917, Brasília, DF, Brazil. 3CIRAD UMR RAGAP, F-34398 Montpellier, France. 4Embrapa Informática Agropecuária, UNICAMP, Av. André Tosello nº 209, CP 6041, 13083-886 Campinas, SP, Brazil. 5Institute of Biotechnology, Parque Estação Biológica, CP 02372, 70770-917, Brasília, DF, Brazil. 6Institute of Botany, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, 30123-970, Belo Horizonte, MG, Brazil. 7Instituto de Biologia/UNICAMP, Cidade Universitária Zeferino Vaz, 13083-970 Campinas, SP, Brazil. 8Department of Genetics, School of Agriculture, University of São Paulo, Jaboticabal, SP, Brazil. 9present address: Embrapa Café, INOVACAFÉ, Campus UFPA, 37200-000 Lavras, MG, Brazil.

Received: 14 October 2015 Accepted: 13 April 2016

Published online: 19 April 2016

References


84. Cameron KD, Moskal WA, Smart LB. A second member of the Nicotiana glauca lipid transfer protein gene family, NgTP2, encodes a divergent and differentially expressed protein. Funct Plant Biol. 2006;33:141–52.
