THÈSE

PRÉSENTÉE A

L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE SCIENCES ET ENVIRONNEMENTS

Par **Emilie VILLAR**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPECIALITE : ECOLOGIE EVOLUTIVE, FONCTIONNELLE ET DES COMMUNAUTES

Plasticité phénotypique et moléculaire de deux clones d'eucalyptus sous contrainte hydrique au champ.

Devant la commission d'examen formée de :

Jean-Louis JULIEN	Directeur de recherche, INRA, Clermont-Ferrand	Rapporteur
Bertrand MULLER	Directeur de recherche, INRA, Montpellier	Rapporteur
Franck BRIGNOLAS	Professeur des Universités, Orléans	Examinateur
Pascal MONTORO	Chercheur, CIRAD, Montpellier	Examinateur
Eric GOMES	Professeur des Universités, Bordeaux I	Examinateur
Christophe PLOMION	Directeur de recherche, INRA, Bordeaux	Directeur
Jean-Marc GION	Chercheur, CIRAD, Bordeaux	Co-directeur

Remerciements

Tout d'abord je tiens à remercier Jean Marc Bouvet de m'avoir permis de faire partie de l'UPR39 Génétique Forestière du CIRAD, ainsi qu'Antoine Kremer, puis Rémi Petit de m'avoir accueilli au sein de l'UMR BioGeCo.

Je souhaite également remercier Messieurs Bertrand Muller et Jean Louis Jullien, d'avoir accepté d'être les deux rapporteurs de ce travail, ainsi que Messieurs Franck Brignolas, Pascal Montoro et Eric Gomes d'avoir bien voulu faire partie de mon jury de thèse.

Mes deux encadrants de thèse, Christophe et Jean-Marc, ont su me faire partager leur vision de la recherche, m'ont toujours appuyé pour aller en formation, même à l'autre bout du monde. Votre complémentarité a été un grand atout dans cette thèse, et je vous en remercie vivement. J'adresse un remerciement particulier à Jean-Marc, qui m'a fait confiance pour la mise en place du dispositif expérimental, et puis qui m'a ensuite donné ma chance pour réaliser cette thèse.

Je voudrais remercier chaleureusement toute l'équipe du CRDPI (ex UR2PI (ex UAIC)) pour les deux ans que j'y ai passé en pré-thèse, et qui m'ont permis non seulement de commencer ma thèse avec une idée du programme d'amélioration « eucalyptus », mais aussi avec le dispositif expérimental de ma thèse en place. Mes premières pensées vont à Jean Noël Marien pour m'avoir recrutée en tant que Volontaire Internationale, m'avoir soutenu, aidé, et formé tout au long de ces deux ans, et surtout, d'avoir eu confiance en moi quand je lui ai parlé du dispositif expérimental de l'ATP ! Je ne peux évidemment pas oublier Aubin R. Saya, toujours prêt pour une petite blague, qui a su temporiser quand la pression montait, et à tout moment me donner de judicieux conseils. Je pense également aux équipes de terrains, spécialement à Andréas Ndeko, l'homme de l'ATP, sans qui j'aurais vraisemblablement fait ma thèse sur le radis. Ce projet a mobilisé énormément de personnel de l'UR2PI, en tout cas André Mabiala, mon vieux Mabios, qui fait encore résonner son rire dans la savane de Yanika. Et puis il y a Dame Toto, Mélanie Toto qui s'est occupé des boutures, Juste Akana (qui ne rate pas le manioc au bord de la route), pour la préparation de terrain, Jissé, Jean Claude Mazoumbou (Le Mc Gyver congolais), pour les sondes TDR et les flux de sève, les chauffeurs Daniel Ossibi (qui m'aura fait danser jusqu'à la chute) et Maurice Kenga (et ses fameuses histoires de démons cornus), qui nous ont toujours filé un petit coup de patte pour boucler la semaine. Et pendant ce temps là, mon papa Charles se démenait pour me sortir des traquenards administratifs et autres. Et Maman Julienne… faisait tout le reste : négociations, motivation, le petit café qui va bien, un coup de balais, et beaucoup de rires ! Merci aussi à Joël Polidori, qui a pris le relais sur l'essai quand je suis retournée à mes pipettes. A l'UR2PI, il y avait aussi : mes yayas François le pape et Hugo, Gerda, Armel, Agnès, Joseph, Fanny, Sylvain, Célestin, David, Arsène, Séraphin et Gabriel, et tant d'autres, je pense à vous, vous me manquez trop même…

Ma thèse s'appuie sur la participation et la collaboration d'un bon nombre « d'agents du CIRAD », également : Christophe Jourdan, (Couscous pour les participants du P4), co-responsable du projet, nkumbi émérite et reconnu, le plus grand approvisionneur en caisse de jus et en solutions à tout problème ; Philippe Vigneron qui m'a tout appris sur l'eucalyptus, des pigeons verts, et aussi du pâté de tête ; Philippe Deleporte, le roi de la manipulation de tableaux de données et de bilans hydriques ; « les écophys » : Agnès de Grandcourt, Yann Nouvellon, Olivier Roupsard et Jean Paul Laclau, pour tous les conseils, et l'aide à l'interprétation des données ; Laurent St-André pour déchiffrer ses fichues allocations de biomasse ; Louis Mareschal pour le coup de main sur les prélèvements de sol ; Jean-Luc Verdeil et son équipe pour les analyses d'anatomie, Christian Mertz pour l'analyse des

phénols. Enfin, j'adresse un remerciement particulier aux membres du feu UPR39, nouvellement GF&P: Bénédicte Favreau, Gilles Chaix, Alexandre Vaillant, et surtout merci pour tout Roselyne !!!

Je tiens également à remercier les membres de mes comités de thèse, qui m'ont aidé à m'orienter au milieu de ce flot de données, insisté pour trouver le fil qui reliait le tout, et eu la patience de s'intéresser aux différentes disciplines du projet. Ce sont Jacqueline Grima-Pettenati, Irène Hummel et Erwin Dreyer. Merci aussi à l'équipe Génotoul pour l'assemblage 454 : Céline Noirot et Christophe Klopp, ainsi qu'à Matias Kirst et Evandro Novaes, pour m'avoir accueilli à l'Université de Floride, pour débrouiller ces données de séquençage. Enfin je classe dans cette catégorie hétérogène, le grand, le seul, l'unique, Frank Bedon, plus qu'un collègue de bureau, plus qu'un très bon peintre en bâtiment, plus que MON post-doc attitré, en tout cas un très bon conseiller qui m'a mis sur les bons rails en partant.

Je n'aurais pas pu mener cette thèse sans tous les membres de Pierroton : d'abord l'équipe de transcriptome, avec Grégoire, Philou, Valérie et Céline, qui m'ont formé et aidé pour les analyses d'expression ; il y a aussi la PGT Crew, avec Franck, Christophe et Sarah, qui m'ont permis d'accélérer tout ça avec le fameux Ken dit La Press associé au cycleur de lumière; Jean Marc et Thierry de m'avoir initié aux joies du cd.. et autres nohup avec un big-up pour Isa, qui arrivait à trouver un peu de temps aux "eucas"; Annabel d'avoir bien voulu se frotter aux données de ces satanés génomistes ; Alain d'avoir pris le temps de m'expliquer les analyses multivariées ; merci également à Henri de m'avoir dépanné pour la journée moustache ; et puis merci à tous de m'avoir accueillie presque comme si j'étais pas du Cirad, avec chaleur, et toujours un peu de temps autour d'un café, pour parler science ou autres! J'envoie une pensée particulière à tous les collègues thésards : François, Erwan, Lélia, Jérômes, même les pathos et les entomos, c'était chouette ! Et encore un Big-Up pour la Klanik production : Kutulu, Peio, Flo-Flo La Broutille, Sarah le pti Chat, et LaKarN aka JeJeune aka GranGran aka BooBoo. And I would like to thank Eric Mandrou to give me the opportunity to do my PhD in an excellent mood.

Tout cela n'aurait pas non plus été possible sans les encouragements des gens de l'autre côté de la barrière, pour qui « je vois pas pourquoi tu bosses 3 ans là-dessus, si on met plus d'eau c'est sûr que ça pousse plus ». D'abord les potes « congolais » : Matthieu, Solenn & Gilles, Renaud, Mat Oo, Julie & Fab. Et puis, il y a les potes de toujours, Marie et Bertrand, toujours prêts pour une petite virée au Bassin, Pipif et Mathilde, pour les vacances relax en Ardèche, LouLou, Manu et les charentais, Julien mon seul copain bordelais, Dame Marion et sa denrée, Ben, Jick et Popoye ainsi que le patron, les carombais et les carombaises pour les vendanges, où il n'y a plus rien à penser qu'à « couper du raisin »…

Merci à la familia, mes parents qui ont toujours fait le maximum pour mes études, mes frères à qui j'ai voulu damer le pion (ça y est Moun, 3/3!) et qui se sont inquiétés pendant 3 ans de l'avancement de ma thèse, et aussi à Marc, qui m'a fait tomber dedans quand j'étais petite.

Enfin, celui qui a du faire le plus de sacrifices au cours de cette thèse, au moins pour supporter l'humidité et les girondins de Bordeaux, celui qui a su me faire garder les pieds sur terre et en même temps me faire m'évader un peu, qui a su comprendre que c'était important pour moi. C'est peu quelques lignes pour te remercier, Jean, d'avoir été à mes côtés pendant ces trois années.

Memento

Comparaison de 2 génotypes:

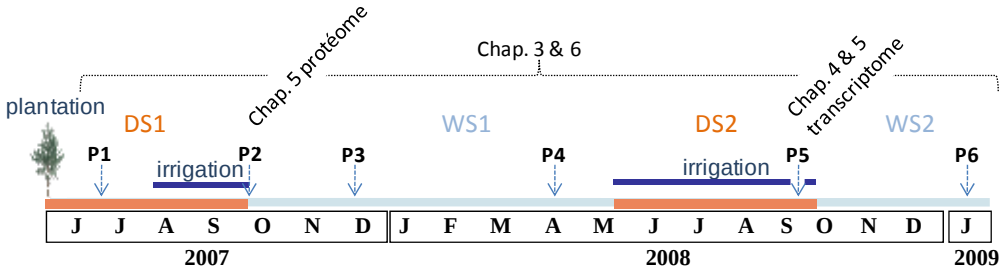
1-41 (moins productif: 25,1 m^3/ha/an) 18-50 (plus productif: 28,5 m^3/ha/an)

Soumis à aux traitements :

IR (irrigués pendant les saisons sèches) NI (non-irrigués pendant les saisons sèches)

IR2 (irrigués uniquement la seconde saison sèche)

Prélèvements de P1 à P6:



DS1	1ère saison sèche	DS2	2nde saison sèche
WS1	1ère saison des pluies	WS2	2nde saison des pluies

Abréviations utilisées pour les variables mesurés:

HT	Hauteur	Mr	nervure centrale
C	Circonférence	LB	limbe foliaire
VOB	Volume sous écorce	Lep	épiderme inférieur
SDW	biomasse de bois	Uep	épiderme supérieur
RDW	biomasse racinaire	Lcu	cuticule inférieure
LDW	biomasse foliaire	Ucu	cuticule supérieure
SLA	surface spécifique	SP	parenchyme spongieux
MLA	surface moyenne foliaire	PP	parenchyme palissadique
δ13C	composition isotopique du carbone	VB	faisceau central
SD	densité stomatique	XV	vaisseaux du xylème
ratio S/R	ratio de biomasse aérienne / racinaire	Uco	collenchyme supérieur
SWC	contenu en eau du sol	Lco	collenchyme inférieur

Liste des abréviations

4CL	4-coumarate:CoA ligase
ABA	Acide Abscissique
ADNc	Acide DésoxyriboNucléique complémentaires
ARN	Acide RiboNucléique
ATP	Adénosine triphosphate
BP	Processus Biologique
C3H	p-coumarate 3-hydroxylase
C4H	cinnamate 4-hydroxylase
CAD	cinnamyl alcool dehydrogenase
CC	Composant Cellulaire
CCoAOMT	Caffeoyl coenzyme A O-methyltransferase
CCR	cinnamoyl-CoA reductase
COMT	acide caféique O-méthytransférase
CRDPI	Centre de Recherche sur la Durabilité et la Productivité des Plantations Industrielles
DE-contigs	Contigs différentiellement exprimés
EST	Marqueur de Séquences Exprimées
EUE_{GPP}	Efficience d'utilisation (calculée par rapport à la production primaire brute)
F5H	ferulate 5-hydroxylase
FAO	Food and Agriculture Organization
FDR	False Discovery Rate
GES	Gaz à Effet de Serre
GIEC	Groupe Intergouvernemental d'Experts sur le Climat
GxE	Interaction génotype-environnement
GxT	Interaction génotype-traitement
HCT	p-hydroxycinnamoyl-CoA:quinate shikimate p-hydroxycinnamoyltransférase
HPLC	Chromatographie liquide à haute pression
IPCC	Intergovernmental Panel on Climate Change
LEA	Late Embryogenesis Abundant
MF	Fonction moléculaire
Mpb	Méga (10^6) paires de base
NGS	Nouvelles générations de séquençage
PAL	Phénylalanine Ammonia-Lyase
pF	Potentiel hydrique (logarithme de la charge hydraulique)
QTL	Quantitative Trait Loci : régions génomiques à effet majeur
ROS	Espèces réactives d'oxygène
RT-qPCR	Reaction de polymérisation en chaine, quantitative en temps réel
SAGE	Analyse d'Expression de Gènes en Série
SNP	Single Nucleotide Polymorphism
SPIR	Spectrométrie Proche Infra-Rouge
UE	Elément de l'unigène
UTR	Régions transcrites non traduites

Table des matières

Liste des figures 4

Liste des Tableaux 6

Avant-propos 9

Chapitre I : Introduction bibliographique 11

1 Les changements globaux 12

2 Les forêts, facteur de développement 15

3 L'adaptation des arbres forestiers aux changements climatiques 18

4 Quelles réponses à la contrainte hydrique ? 28

5 Les NGS pour répondre à des questions de recherche en génomique environnementale 46

Références bibliographiques 63

Chapitre II: Contexte et objectifs 77

1 Contexte de l'étude 78

2 Objectifs de la thèse 79

3 Moyens mis en œuvre 81

Références bibliographiques 88

Chapitre III: Plasticité phénotypique de l'*Eucalyptus* en réponse à un déficit hydrique saisonnier au champ 89

Phenotypic plasticity in response to seasonal water deficit on field of Eucalyptus 89

1 Introduction 91

2 Materials and Methods 93

3 Results 100

4 Discussion 112

5 Conclusions 116

Supplemental File 1 : Biomass predictions 118

1 Methods 118

2 Results 119

References 122

Chapitre 4: Analyse de la plasticité foliaire par une approche intégrative 127

The multi-scale plasticity of leaves seems to influence the growth of two *Eucalyptus* genotypes subjected to water deficit 128

1 Background 129

2 Materials and Methods 131

3 Results 135

4 Discussion ..... - 146 -

5 Conclusion ..... - 150 -

References ..... - 153 -

Chapitre 5: Analyse de la plasticité moléculaire ..... - 159 -

RNA-Seq reveals genotype-specific molecular responses during seasonal water deficit in eucalyptus ..... - 159 -

1 Abstract ..... - 161 -

2 Background ..... - 162 -

3 Methods ..... - 164 -

4 Results ..... - 169 -

6 Conclusions 186

Supporting information 189

1 Methods 189

2 Results 189

References 199

Proteomic plasticity of two *Eucalyptus* genotypes under contrasted water regimes in the field 206

1 Abstract 207

2 Introduction 208

3 Materials and Methods 210

4 Results 215
5 Discussion 222
References 228
Supplementary data for qRT-PCR (S1). 243
1 Methods 243
2 Results 244
References 246
Chapitre VI- Régulation de l'expression des gènes de la voie de biosynthèse des lignines .. 249
Régulation de l'expression des gènes de la voie de biosynthèse des lignines 250
1 Introduction 250
2 Matériel et méthode 253
3 Résultats 259
4 Discussion 267
5 Conclusion 269
Références bibliographiques 270
Discussion générale et perspectives 273
1 Mise en place d'un dispositif expérimental pour étudier la plasticité phénotypique et moléculaire de l'eucalyptus 274
2 Une vision intégrative de la réponse à la contrainte hydrique pour deux clones d'eucalyptus 278
3 Perspectives 283
Références bibliographiques 287

Liste des figures

Chapitres 1 et 2 – Introduction et contexte

Figure 1 – Ecarts de température pour la période 1951-1980. 12
Figure 2 – Limites potentielles à la production nette primaire. 14
Figure 3 – Tendance des stocks de carbone dans la biomasse forestière. 16
Figure 4 – Surfaces forestières plantées en 2005, et projetées en 2030. 17
Figure 5 – Le rôle de la plasticité phénotypique et de la diversité génétique. 20
Figure 6 – Normes de réactions. 22
Figure 7 – Paramètres utilisés pour décrire la plasticité. 23
Figure 8 – Trajectoires ontogéniques de *Pinus canariensis*. 24
Figure 9 – Différents types de plasticité. 25
Figure 10 – Le potentiel hydrique du sol. 28
Figure 11 – Le transport de l'eau dans les arbres. 30
Figure 12 – Les réponses possibles de la plante. 32
Figure 13 – Les voies de transduction du signal. 36
Figure 14 – Les effets de la sécheresse. 43
Figure 15 – Exemple de modélisation du continuum sol-plante-atmosphère. 44
Figure 16 – Les différents types d'ARN. 54
Figure 17 – L'immuno-précipitation de chromatine. 55
Figure 18 – Le séquençage bisulfite. 56
Figure 19 – Applications possibles des NGS. 57
Figure 20 – Nombre d'EST d'Eucalyptus publiés. 62
Figure 21 – Photographie du dispositif expérimental. 82
Figure 22 – Plan du dispositif expérimental. 83
Figure 23 – Déroulement de l'expérimentation. 84
Figure 24 – Disposition des sondes TDR. 85

Chapitre 3 – Plasticité phénotypique

Figure 1 – Field trial map. 94
Figure 2 – Soil NIR spectra analysis. 100
Figure 3 – Meteorological records. 101
Figure 4 – Evolution of water reserve. 102
Figure 5 –Increment of growth traits. 105
Figure 6 – Increments of biomasses. 108
Figure 7 – S/R ratio repartition. 108
Figure 8 – SLA and $d^{13}C$. 111
Figure 9 – Plasticity indices. 115
Figure S1 – Biomass equations and measured sampled. 121

Chapitre 4 – Plasticité foliaire

Figure 1 – Cross-section of eucalyptus leaf. 133
Figure 2 – Soil water content. 136
Figure 3 – Scatter plots for growth, morphological and physiological traits. 138
Figure 4 – Principal component analysis of anatomic measurements. 141
Figure 5 – Comparison of gene expression. 144

Chapitre 5 – Analyse du transcriptome

Figure 1 – Procedure used, from tissue sampling to sequencing. 166
Figure 2 – Pipeline used for sequence analysis. 167
Figure 3 – Profile of soil water content (SWC). 170
Figure 4 - Above-ground biomass. 171
Figure 5 – Pearson correlation coefficient for read frequencies. 173
Figure 6 – Differentially expressed contigs. 175
Figure 7 – Fold-change (FC) distribution. 176
Figure 8 – Distribution of Gene Ontology categories. 178
Figure 9 – Distribution of "GxT" contigs between metabolic pathways. 180
Additional file 2 – Distribution of read length for the three half-runs. 193
Additional file 3 – Increasing coverage with successive runs. 193
Additional file 4 – Comparison of the distribution of Gene Ontology (GO). 194
Additional file 6 – Distribution of “G” contigs. 195
Additional file 8 – Distribution of “T” contigs. 197

Chapitre 5 – Analyse du proteome

Figure 1 – Water supply. 215
Figure 2 – Significant spots following two-way ANOVA. 217
Figure 3 – Fold-change ratios. 219
Figure 4 – Hierarchical clustering of protein abundance. 221
Figure S1 – Representative 2-D gels. 238
Figure S2 – Fold-change distribution of significant spots. 239
Figure S3 – Fold-change distribution of spots normalized volumes from leaves (A) and stems (B). 240
Figure S4 – Reaction norm diagrams of transcript and protein abundance. 241

Chapitre 6 – Voie de biosynthèse des lignines

Figure 1 – Voie de biosynthèse des phénylpropanoïdes et des monolignols. 251
Figure 2 – Déroulement de l’expérimentation de terrain. 254
Figure 3 – Illustration des organes et tissus prélevés. 256
Figure 4 – Exemples de courbes de fusion. 257
Figure 5 – Expression moyenne des différents gènes dans les différents tissus étudiés. 263
Figure 6 – Expression moyenne des différents gènes pour chaque point de prélèvement. 264

Liste des Tableaux

Chapitres 1 et 2 – Introduction et contexte

Tableau 1 – Comparaison des caractéristiques des NGS. 48
Tableau 2 – Projets d'assemblage *de novo*. 50
Tableau 3 – Etudes récentes réalisées par RNA-Seq. 51
Tableau 4 – Caractéristiques des deux clones utilisés. 82

Chapitre 3 – Plasticité phénotypique

Table 1– Irrigation performed. 96
Table 2 – Sampling numbers. 97
Table 3 – Summary statistics of growth traits. 103
Table 4 – Summary statistics of dry weights. 106
Table 5 – Summary statistics of physiologic traits. 110
Table S1 – Final weighted equations. 120

Chapitre 4 – Plasticité foliaire

Table 1 – Summary statistics of growth, morphological and ecophysiogical traits. 137
Table 2 – Summary statistics of anatomic traits. 139
Table 3 – Summary statistics of polyphenols accumulation analysis. 142
Table 4 – Summary statistics of gene expression analysis. 145
Table S1 – Primers used. 152

Chapitre 5 – Analyse du transcriptome

Table 1 - Summary statistics for the three 454-sequencing half-runs. 172
Table 2 – Assembly statistics. 172
Table 3 – Annotation results. 174
Table 4 – Summary of phenotypic and molecular plasticity. 187
Supporting file 1 – Primers used for validation by RT-qPCR. 191
Supporting file 2 – Comparison of fold-change (FC) ratio between RT-qPCR and digital analysis. 192
Additional file 1 – Results of the ANOVA for above-ground biomass. 193
Additional file 5 – Differentially expressed contigs. 194
Additional file 7 – Significant categories for the Wilcoxon rank sum test. 196
Additional file 9 – "G" contigs related to secondary metabolism. 198

Chapitre 5 – Analyse du protéome

Table 1 – Phenotypic variables. 216
Table 2 – Protein identity. 235
Table S1 – Two-way ANOVA results. 242
Table A – Primers pairs for PCR amplification and qRT-PCR. 247

Table B – DNA sequences obtained by PCR amplification in eucalyptus. 247

Chapitre 6 – Voie de biosynthèse des lignines

Tableau 1 – Amorces utilisées pour l'amplification par RT-qPCR. 257
Tableau 2 – Résultats de la recherche de 10 gènes de la voie de biosynthèse des lignines dans les banques de séquences 454. 260
Tableau 3 – Résultats de l'ANOVA effectuée pour chaque point de prélèvement. 265
Tableau 4 – Niveaux relatifs d'expression par génotype et par traitement. 266

Chapitre 7 – Conclusion

Tableau 1 – Récapitulatif des principaux résultats obtenus dans cette thèse. 276

Avant-propos

Ce projet de thèse a émergé en 2007, alors que j'étais volontaire internationale à l'Unité de Recherche sur la Productivité des Plantations Industrielles, à Pointe Noire, en République du Congo. En charge d'une partie du programme d'amélioration de l'eucalyptus en collaboration avec le CIRAD, Jean-Marc Gion et Christophe Jourdan m'ont proposé de coordonner la mise en place de l'audacieux dispositif expérimental qu'ils avaient imaginé pour étudier la réponse à la contrainte hydrique chez l'eucalyptus. Cet essai s'inscrivait parmi un ensemble de 3 dispositifs installés au champ et se proposait de décrire la plasticité (et son contrôle génétique) de l'expression du génome jusqu'à des caractères très intégrés comme la biomasse, chez l'eucalyptus soumis à une contrainte hydrique lors de la saison sèche en République du Congo. Les mêmes questions étaient posées pour le caféier au Brésil et le pin maritime en Aquitaine. A la fin de mon volontariat international (Décembre 2007), j'ai pu poursuivre le travail de terrain commencé au Congo grâce à l'attribution d'une allocation de recherche du département BIOS du CIRAD, en m'installant à l'UMR 1202 (BioGeCo) de Bordeaux en Janvier 2008. Je présenterai donc dans cette thèse une partie des résultats obtenus sur le dispositif expérimental « eucalyptus », principalement en axant son contenu sur l'analyse des déterminants moléculaires de la réponse à la contrainte hydrique.

Le manuscrit s'articulera autour de 5 chapitres.

Le **chapitre introductif** se propose de faire un tour d'horizon **bibliographique** du contexte et des concepts utilisés dans la thèse. Il présente synthétiquement la problématique des changements globaux et de leur impact sur les systèmes forestiers plantés ainsi que l'importance des forêts pour le développement socio-économique. Il aborde les différents devenirs possibles pour les forêts soumises aux changements climatiques, et s'attarde sur la relation entre l'eau et la plante, en décrivant la réponse à la contrainte hydrique à différentes échelles. Enfin, il synthétise l'apport des nouvelles technologies de séquençage à des questions de recherche de génomique environnementale.

Le **second chapitre** présente les **objectifs** de mon travail de thèse, en spécifiant le contexte de l'étude, les hypothèses et questions de recherche associées, ainsi que les moyens mis en œuvre pour y répondre.

Le **troisième chapitre** présente le dispositif expérimental, et décrit l'analyse de la **plasticité de caractères écophysiologiques** (discrimination isotopique du carbone, densité stomatique, surface foliaire spécifique) en lien avec celle observable pour la **croissance** (hauteur, circonférence, allocation de biomasse). Cette analyse est effectuée pour les deux clones d'eucalyptus étudiés soumis à deux régimes hydriques différents (avec ou sans contrainte hydrique), au cours des 2 premières années de plantation. Il fait l'objet d'un projet d'article rédigé en anglais.

Le **quatrième chapitre** porte sur l'étude de la **plasticité phénotypique au niveau foliaire** observée à la fin de la seconde saison sèche. Il intègre **différentes échelles d'analyse** (morphologie, anatomie, chimie foliaire, accumulation de transcrits). Il s'inscrit donc dans une démarche de biologie intégrative, et fait également l'objet d'un projet d'article, rédigé en anglais.

Le **cinquième chapitre** s'intéresse à la plasticité moléculaire, via **l'analyse du transcriptome d'apex caulinaires et du protéome de feuilles et de tiges,** échantillonnés à la fin de la saison sèche. Ces analyses ont été valorisées par l'écriture de deux articles. Celui portant sur le transcriptome a été soumis à BMC Genomics, et celui sur le protéome à Plant, Cell & Environment.

Le **sixième chapitre** se focalise sur la **voie de biosynthèse des phényl-propanoïdes** : Il s'agit d'un suivi de l'accumulation des transcrits liés à cette voie de biosynthèse dans différents tissus au cours des 2 premières années de croissance. Ce chapitre a été rédigé en français et fera l'objet d'un article en anglais après la thèse.

Enfin, le **chapitre 7** synthétise les résultats obtenus, discute des caractères présentant de l'interaction génotype-environnement, la composante génétique de la plasticité phénotypique, et des mécanismes moléculaires sous-jacents. Les perspectives issues de ce travail à court et plus long terme y seront également exposées.

Chapitre I - Introduction bibliographique

1 Les changements globaux

Le changement global désigne l'ensemble des changements induits dans la dynamique de la biosphère par les activités humaines, directement ou non (Trometter et Weber, 2005). Cela correspond aux changements climatiques, mais aussi aux autres changements d'origine anthropique ayant des conséquences globales. En tant que modification de la dynamique de la biosphère, le changement global existe depuis longtemps ; ce qui est nouveau, à l'échelle des deux derniers siècles, c'est son amplification et l'accélération de son rythme, à l'image de l'évolution des températures moyennes ou du dioxyde de carbone dans l'atmosphère.

1.1 Le bouleversement climatique

Différentes variations de paramètres du climat ont été constatées ces dernières années principalement en conséquence de l'augmentation des gaz à effet de serre (GES) dans l'atmosphère (figure 1).

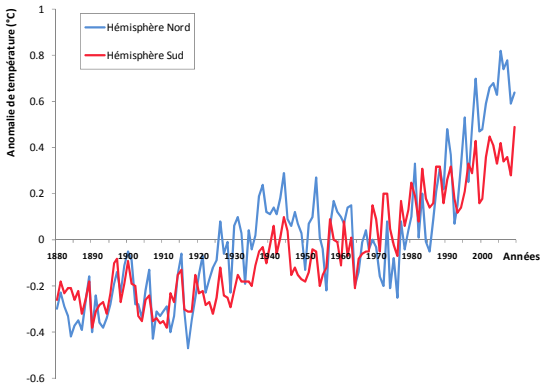


Figure 1 : Ecarts de température par rapport aux normales calculées pour la période 1951-1980, pour les hémisphères Nord et Sud. Source : http://data.giss.nasa.gov/

D'après les prévisions du groupe intergouvernemental d'experts sur le climat (GIEC), il y a peu de chance pour que ces tendances s'inversent, et on peut surtout s'attendre à une intensification des modifications du climat (IPCC, 2007).

Le réchauffement de la planète est sans doute la plus connue de ces modifications. En effet, les hypothèses les plus optimistes (stabilisation des émissions de GES au niveau de 1990) révèlent tout de même une augmentation d'environ 1,1°C d'ici à 2100. Vu le développement continu des activités industrielles, le scénario le plus communément envisagé pour 2100 est celui où la concentration en CO_2 aura doublé par rapport à l'ère pré-industrielle ; la température de la Terre s'élèvera alors de 3°C en moyenne d'ici à 2100.

Une augmentation de la température moyenne à la surface de la terre devrait selon toute vraisemblance entraîner une augmentation de l'évaporation et des précipitations. Il est toutefois difficile de prédire et de mesurer avec précision les changements dans le cycle hydrologique, vu les processus complexes de l'évaporation, des courants atmosphériques et des précipitations. Néanmoins, les scientifiques du GIEC prédisent une augmentation très probable des précipitations aux latitudes élevées et, au contraire, une diminution probable sur la plupart des terres émergées subtropicales, conformément aux tendances relevées récemment (IPCC, 2007).

Une autre conséquence de cette perturbation du climat est la multiplication des phénomènes extrêmes et des anomalies climatiques. Divers événements climatiques exceptionnels ont marqués les dernières années : que ce soit les étés caniculaires (2003 et 2006), les tempêtes de 1999 et 2008 en Europe, la saison cyclonique de 2005, ou les fortes précipitations au cours de l'année 2000, les catastrophes naturelles se sont multipliées. On estime à 200 millions les personnes touchées chaque année par des phénomènes climatiques extrêmes.

1.2 Impact sur les rendements des productions végétales

La variation de divers paramètres climatiques, conséquence directe ou indirecte des changements globaux, risque d'affecter les rendements des cultures de façon hétérogène en fonction des zones géographiques considérées. Ainsi, l'augmentation de CO_2 atmosphérique et de la température devraient améliorer les rendements dans les zones de hautes latitudes, tandis que l'augmentation de température associée à la diminution des précipitations dans les faibles latitudes constitue une menace pour le développement durable de ces régions. Le fait que les jours froids soient moins nombreux et moins froids, et que les jours chauds soient plus nombreux et plus chauds sur la plupart des terres émergées va entrainer une hausse des rendements dans les régions froides, mais une baisse de rendement et une augmentation du risque d'incendies dans les régions chaudes. Les invasions de pathogènes risquent d'être également plus fréquentes. D'autre part, l'augmentation des fréquences de précipitation violente risque d'accroitre l'érosion des sols, la verse de certaines cultures, voire même de rendre impossible les cultures sur des terres détrempées. L'augmentation d'activité cyclonique intense risque d'entrainer des pertes de récoltes et la dégradation des forêts par déracinement. Enfin, l'augmentation de la fréquence des phénomènes d'élévation extrême du niveau de la mer pourra augmenter la salinité des eaux d'irrigation, des estuaires et des systèmes d'eau douce (IPCC, 2007).

Les stress abiotiques (sécheresse, salinité, fortes températures et stress oxydatifs…) constituent la première cause de perte de production végétale, réduisant les rendements moyens pour les espèces cultivées majeures jusqu'à 50% (Wang 2003). Ainsi, les variations environnementales engendrées par le réchauffement global risquent d'influencer les rendements de façon très importante (figure 2).

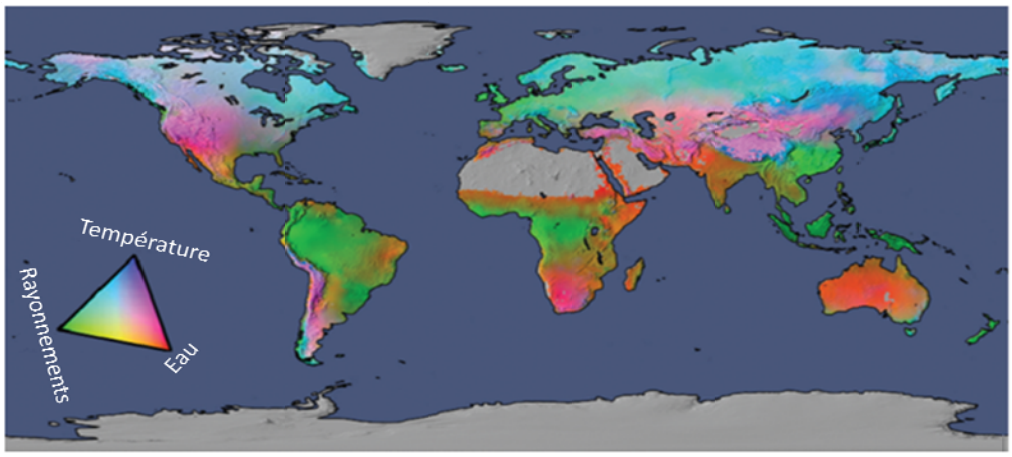


Figure 2: Limites potentielles à la production nette primaire de la végétation prenant en compte le déficit de pression de vapeur d'eau, le bilan hydrique et la température (issu de Boisvenue et Running, 2006). Le code couleur indique le facteur le plus limitant pour chaque zone du Monde.

Pour mieux prévoir les conséquences économiques des variations de rendements de culture, la FAO est en train de développer un modèle (http://www.fao.org/climatechange/mosaicc/en/) prenant en compte les scénarios climatiques envisagés par l'IPCC, les paramètres de la culture d'intérêt, et les caractéristiques hydrologiques de la région. Ainsi, le premier cas d'étude de ce modèle a permis de montrer qu'au Maroc, seules les cultures irriguées devraient accroître leur rendement (jusqu'à +20% en 2050 pour les légumineuses et fourrages), sous réserve que de l'eau d'irrigation soit disponible. Les rendements des cultures devraient peu varier jusqu'à 2030, mais à partir de 2050 les rendements des cultures pluviales devraient chuter de -10 à -30% selon les types de cultures.

1.3 Le cas particulier des arbres forestiers

La productivité d'espèces pérennes comme les arbres forestiers risque d'être particulièrement touchée par ces variations environnementales. En effet, ces espèces devraient subir ces changements au cours de quelques, voire d'une seule génération. De par leur longévité et leur taille, les arbres forestiers doivent faire face à ces contraintes de façon récurrente, ce qui à terme limitera la production de bois.

Dans les zones tempérées où la température, la lumière et les éléments nutritifs sont les ressources les plus limitantes, le réchauffement climatique aura des conséquences positives sur la productivité des forêts, à condition que les ressources hydriques ne soient pas limitées (Boisvenue et Running, 2006). Par contre, sur le tiers des terres émergées ou l'approvisionnement en eau est limité, le réchauffement climatique pourrait avoir de graves conséquences en terme de baisse de croissance, voire de mortalité. De plus, les épisodes climatiques extrêmes, tels que l'occurrence plus fréquente et plus intense de sécheresses, risquent de concerner la majorité des forêts mondiales, qui pourront voir leur croissance réduire, et leur santé menacée à cause de l'augmentation de la pression due aux pathogènes (Desprez-Loustau *et al.* 2006). Ainsi, en réalisant une méta-analyse, Allen *et al.* (2010) ont relevé un nombre croissant d'études liant la mortalité d'espèces forestières à des stress hydriques ou de chaleur. Si de nombreux bémols peuvent être émis, principalement concernant le manque de connaissances sur les interactions entre climat et mortalité des arbres, il semble tout de même que ces dernières années montrent une hausse de la mortalité des arbres en liaison avec le réchauffement climatique. D'autre part, la diversité des réponses en fonction des espèces rend difficile la prédiction d'évolution des forêts dans les années à venir. Breda *et al.* (2006) ont montré que la vulnérabilité à la cavitation ainsi que la marge de potentiel hydrique entre l'ouverture complète des stomates et la cavitation pouvaient être très différentes d'une espèce à l'autre, chez les espèces de forêts tempérées. Ricker *et al.* (2006) ont simulé la croissance de deux espèces d'arbres du Mexique sous changement climatique. La simulation a prédit que l'espèce montagneuse diminuera sa croissance relative de 10% au cours de sa durée de vie, tandis que l'espèce de forêt ombrophile tropicale l'augmentera de 25%. Ils concluent que la relation entre température et précipitations devrait avoir un rôle clé plutôt que l'augmentation des températures seules.

2 Les forêts, facteur de développement

Les forêts couvrent 30% des terres émergées, soit un peu plus de 4 milliards d'hectares. Si les forêts du monde suscitent aujourd'hui un intérêt sans précédent, c'est qu'elles jouent un rôle fondamental à différents niveaux.

2.1 Le rôle des forêts

Au niveau socio-économique, elles assurent la production de bois rond industriel et de bois énergie indispensable au développement. D'après les projections, la population mondiale devrait passer de 6.4 milliards d'êtres humains en 2005, à 7,5 milliards en 2020 et 8,2

milliards en 2030. Cette augmentation de population s'accompagnera donc de l'augmentation de la demande en matière premières et plus spécifiquement des besoins en bois. Ainsi la consommation de bois rond industriel devrait passer de 1682 à 2165 millions de m^3 de 2005 à 2020 (FAO, 2010). De plus les politiques environnementales poussent à accroître la part des énergies renouvelables telles que l'utilisation de bois énergie. Pour répondre à ces demandes croissantes, il faudra donc soit augmenter les surfaces plantées, soit augmenter la productivité des forêts et des processus de transformation du bois rond.

Au niveau écologique, le rôle des forêts dans le cycle mondial du carbone est indiscutable. La possibilité d'atténuer les effets des changements climatiques en réduisant les émissions de carbone dues au déboisement et à la dégradation des forêts et en augmentant l'absorption de carbone à travers le boisement et la gestion durable des forêts met en relief le rôle essentiel qu'elles jouent dans le maintien de la vie sur Terre. Les estimations réalisées pour l'évaluation des ressources forestières mondiales en 2010 (FRA 2010) montrent que les forêts du monde renferment 289 gigatonnes (Gt) de carbone dans leur seule biomasse. Au niveau mondial, le carbone stocké dans la biomasse forestière a diminué d'environ 0,5 Gt chaque année sur la période 2005-2010, à cause essentiellement de l'amenuisement de la superficie forestière mondiale (figure 3).

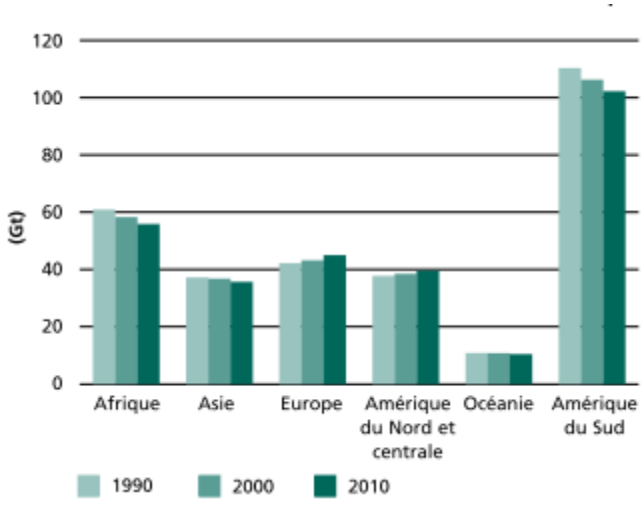


Figure 3: Tendance des stocks de carbone (en Gt) dans la biomasse forestière de 1990 à 2010. Source : FRA 2010.

De plus, les forêts jouent aussi un rôle dans la conservation des écosystèmes, comme par exemple la conservation des sols et des eaux, la maîtrise des avalanches, la fixation des dunes, la lutte contre la désertification ou la protection côtière. Par ailleurs, même si il est difficile de le quantifier, les forêts mondiales abriteraient plus d'un tiers de la biodiversité terrestre.

Enfin, au niveau social, si plus de 10 millions de personnes travaillent dans le secteur forestier, plus de 1,2 milliard de personnes vivant dans la pauvreté extrême dépendent d'une manière ou d'une autre de la forêt pour subsister. Elle est aussi un lieu privilégié permettant la conservation de patrimoines culturels et la mise en place d'activités récréatives ou éducatives (FRA 2010).

2.2 Les forêts plantées pour répondre aux besoins croissants

La perte nette de surface forestière entre 2000 et 2010 est évaluée à environ 5.2 millions d'hectares par an (FAO, 2010). Ces pertes se situent principalement dans les pays « à faible revenu » des tropiques (Brésil et en Indonésie), tandis que les pays plus développés tendent à augmenter leurs surfaces forestières (par exemple, en France on constate une augmentation de 1.7Mha des surfaces forestières depuis les 20 dernières années). Les forêts plantées constituant seulement 7% des surfaces forestières avec 271 millions d'hectares, c'est évidemment les surfaces de forêts naturelles qui reculent le plus (Carle and Holmgren, 2008). Cependant, les politiques environnementales mises en place ces dernières années permettent de réduire les surfaces naturelles exploitées en les classant en zone protégées: à l'heure actuelle, jusqu'à 13.5% des surfaces forestières mondiales sont gérées en tant que surfaces protégées (Schmitt *et al.*, 2008). De plus, les surfaces plantées tendent à augmenter afin de diminuer la pression exercée sur les forêts naturelles (figure 4) : en 2000, alors qu'elles représentaient seulement 3% des surfaces forestières, les forêts de plantations produisaient plus de 35% du bois commercialisé.

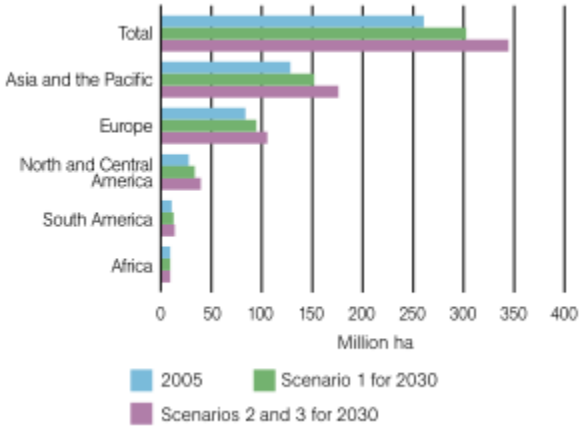


Figure 4: Surfaces forestières plantées en 2005, et projetées en 2030 selon 2 scénarios : le 1er, si l'accroissement des surfaces plantées diminue de moitié par rapport aux dernières estimations (en prenant en compte la diminution des espaces disponibles et de la demande), le second prenant en compte un accroissement continu (issu de Carle et Holmgren, 2008).

Les trois quart de ces forêts plantées font appel à des essences indigènes, alors qu'un quart sont constituées d'espèces introduites (FRA 2010). En 2005, les conifères constituaient 54% des surfaces plantées pour la production, avec le genre *Pinus* principalement (32% des surfaces), tandis que les feuillus représentaient 39%. Le genre *Eucalyptus* est le feuillu le plus

planté au monde, occupant 8% des surfaces plantées, devant les genres *Acacia*, *Tectona*, *Castanea* et *Populus*. Ce genre tend à occuper une place prépondérante (50% du marché en 2005) dans le commerce des fibres courtes. Il a été introduit à travers le monde sur des latitudes très variables : du sud de l'Europe jusqu'en Afrique du Sud. Plusieurs programmes d'amélioration génétique des eucalyptus sont conduits au Brésil, en Afrique, en Europe, ou en Indonésie dans le cadre de la production de pâte à papier, et dans certains cas, pour le bois d'énergie (charbon de bois destiné à des applications industrielles dans les hauts fourneaux par exemple).

L'importance des forêts dans les différents rôles qu'elles exercent pour l'équilibre socio-économique et écologique de la planète rend indispensable leur adaptation aux changements globaux. Compte-tenu des pressions anthropiques et climatiques inéluctables, il semble nécessaire d'étudier dès maintenant les possibilités d'adaptation des écosystèmes forestiers, afin de prévoir les conséquences et d'anticiper les changements à venir.

3 L'adaptation des arbres forestiers aux changements climatiques

Face aux changements globaux, il est possible d'envisager 3 situations possibles pour les populations d'arbres forestiers : extinction, migration ou adaptation (Aitken *et al.*, 2008).

3.1 La migration des espèces

Selon Thomas *et al.* (2004), un réchauffement global de 0,8°C à 2,2°C en 2050 devrait se solder à terme par l'extinction de 38 à 57% des espèces végétales dans le Cerrado brésilien, de 3 à 16% des plantes européennes, et de 24 à 40% des *Proteaceae* sud-africaines. L'ampleur de l'effet estimé varie non seulement avec l'intensité du réchauffement, mais aussi avec les possibilités de migration des espèces, puisque ces deux facteurs agissent en synergie : l'incapacité à migrer vers des sites distants multiplie environ par deux les taux d'extinction prévus. Seules les espèces très mobiles ou disposant de «couloirs de migration» pourront coloniser des habitats distants au climat plus favorable (Teyssedre, 2004). L'estimation des taux passés de migration des espèces grâce à la paléobotanique et les analyses génétiques sont fréquemment utilisées pour prédire le potentiel de migration contemporaine (Aitken *et al.*, 2008). D'après Aitken (2008), le taux de migration des espèces forestières (en général en altitude, et en direction des lisières nord de l'aire de répartition) est estimé à moins de 100 m par an grâce à la dispersion de graines. Une fois les populations fondatrices établies, la

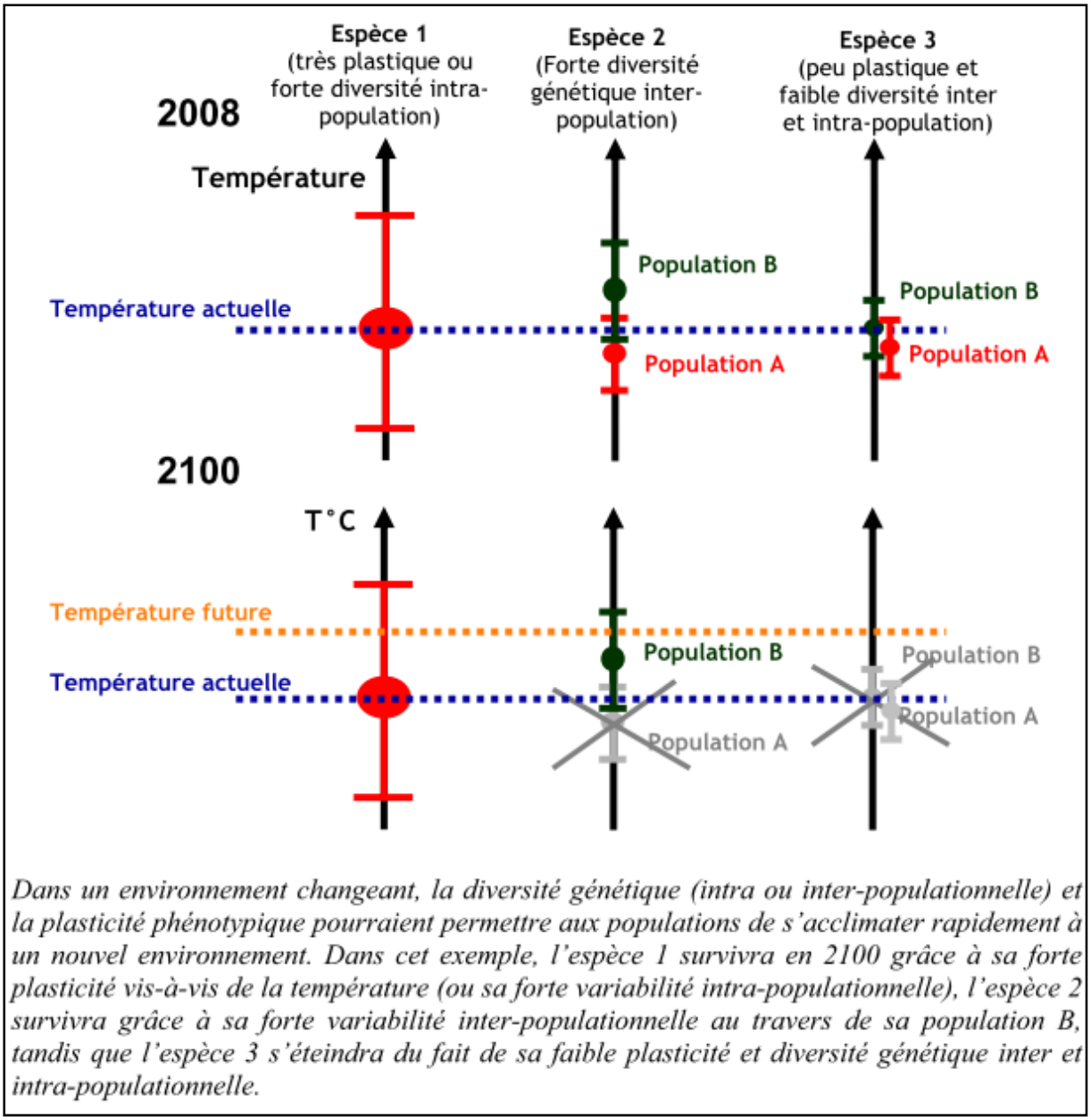
longévité et la plasticité phénotypique pourra permettre la persistance de ces populations. L'établissement de populations en effectif suffisamment important et génétiquement variables pour être viables et devenir adaptées localement est également possible via les flux de pollen à longue distance et l'hybridation avec des populations en place (Petit *et al.*, 2003).

De la même façon, pour persister dans un environnement alors qu'il est entrain de changer, les populations peuvent déployer deux mécanismes : l'adaptation génétique et la plasticité phénotypique (Chevin *et al.* 2010, figure 5).

3.2 L'adaptation génétique

L'adaptation génétique est le résultat de l'action de la sélection naturelle sur la diversité génétique, qui va favoriser les génotypes les mieux adaptés aux conditions environnementales (Krutovskii et Neale, 2001). Une population avec une base génétique large, donc un fort niveau de diversité génétique aura plus de chance de contenir des génotypes adaptés aux conditions climatiques de demain.

La diversité des populations a tout d'abord été analysée en utilisant des marqueurs moléculaires dits neutres, c'est-à-dire non affectés par la sélection, pour retracer l'évolution démographique des populations (Gonzalez-Martinez *et al.*, 2006). Ces marqueurs moléculaires (tels que les motifs microsatellites) sont disposés « au hasard » dans le génome. Dans ce cas, on émet l'hypothèse que les forces évolutives (dérive, mutation, migration) affectent l'ensemble du génome tandis que la sélection affectera des régions spécifiques du génome. Différentes études ont montré que le niveau de diversité génétique neutre chez les arbres forestiers était globalement assez élevé (revue par Petit et Hampe, 2006). Ceci peut être expliqué par leur régime de reproduction majoritairement allogame, l'importance des flux de gènes via la dispersion du pollen, un nombre élevé de chromosomes (lié à une possibilité de recombinaison plus élevée), une taille efficace des populations élevée, les effets majoritairement diversifiant de la sélection naturelle et leur âge tardif de maturité reproductive (évitant les effets de fondation). Néanmoins, si le niveau de différenciation génétique mesurée par des marqueurs moléculaires neutres entre populations est faible, le



Dans un environnement changeant, la diversité génétique (intra ou inter-populationnelle) et la plasticité phénotypique pourraient permettre aux populations de s'acclimater rapidement à un nouvel environnement. Dans cet exemple, l'espèce 1 survivra en 2100 grâce à sa forte plasticité vis-à-vis de la température (ou sa forte variabilité intra-populationnelle), l'espèce 2 survivra grâce à sa forte variabilité inter-populationnelle au travers de sa population B, tandis que l'espèce 3 s'éteindra du fait de sa faible plasticité et diversité génétique inter et intra-populationnelle.

Figure 5 : Représentation schématique du rôle de la plasticité phénotypique et de la diversité génétique dans le cadre des réchauffements climatiques (issu de Vitasse, 2009).

niveau de différenciation génétique des caractères quantitatifs est généralement élevé (Petit et Hampe, 2006). Par exemple, Tripiana *et al.* (2007) ont montré que si la différenciation entre populations d'*E. urophylla* des îles de la Sonde estimée à partir de marqueurs microsatellites était faible (F_{ST} de 0.04), la différenciation entre population estimée à partir des caractères de croissance et de survie est élevée (Q_{ST} entre 0.07 et 0.33). Si l'on est capable de différencier des populations au niveau de leur phénotype, il doit donc être possible de différencier des populations en fonction de leur génotype, en utilisant des marqueurs liés à des caractères fonctionnels.

La recherche de marqueurs moléculaires révélant une signature de la sélection naturelle permet d'identifier des gènes d'intérêt pour l'adaptation des populations, s'il est possible d'écarter un signal démographique qui aurait pu entraîner une signature similaire à la sélection. Pour les espèces dont le génome n'est pas séquencé comme la plupart des arbres forestiers ou dans le cas de grand génome (comme celui des conifères), l'approche la plus utilisée consiste à focaliser les analyses de diversité sur des régions spécifiques du génome, appelées gènes candidats. En effet les marqueurs idéaux pour l'étude de la variation adaptative doivent présenter les critères suivants : i/ être impliqués dans le contrôle génétique de caractères adaptatifs, ii/ avoir une séquence identifiée et une fonction connue, et iii/ avoir une variation allèlique facilement identifiable (Gonzalez-Martinez, 2006). Ces gènes candidats peuvent être choisis en fonction de leurs profils d'expression et/ou de leur fonction putative, et/ou de leur colocalisation avec des QTL (région génomique d'intérêt) pour tel ou tel caractère adaptatif d'intérêt.

Dans ce cadre, un des objectifs de ma thèse a consisté à rechercher des gènes candidats pour l'adaptation à la contrainte hydrique, sur la base de leur expression et de leur fonction.

3.3 Plasticité phénotypique

La plasticité phénotypique d'un génotype est sa capacité à produire plusieurs phénotypes en réponse à un changement de l'environnement (Bradshaw,1965; Schlichting, 1986). Comme nous l'avons vu précédemment, elle joue un rôle important pour l'adaptation des individus face aux changements climatiques car cette réponse est plus rapide que le déploiement de la diversité génétique, particulièrement chez les espèces à cycle long comme les arbres forestiers, mais également car c'est la seule réponse possible au niveau individuel (Hamrick, 2004 ; Jump et Penuelas, 2005).

3.3.1 Comment la décrire ?

La plasticité phénotypique est souvent décrite par la courbe de réponse du phénotype en fonction de l'environnement appelée norme de réaction (figure 6). On peut alors comparer les normes de réactions de différentes unités génétiques.

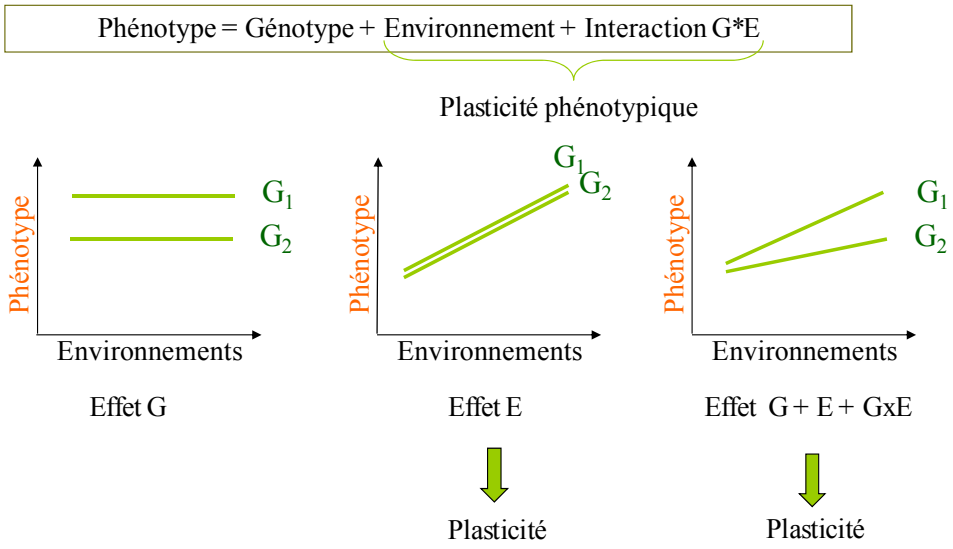


Figure 6 : Différentes normes de réactions représentant les variations phénotypiques de deux génotypes sur une gamme d'environnements. On peut décomposer la variation phénotypique en effet du génotype, de l'environnement, et de l'interaction génotype-environnement. Les deux composantes de la plasticité phénotypique sont les effets dus à l'environnement et à l'interaction GxE, cette dernière représentant la composante génétique de la plasticité phénotypique.

Dans le cas idéal, les unités génétiques devraient être représentées par des clones pour minimiser la variation génétique, mais le plus souvent, ce sont des familles voire des populations qui sont comparées.

La composante environnementale représente généralement des variations spatiales, soit par des sites d'étude représentant des conditions édapho-climatiques contrastées, soit par des milieux contrôlés dans lesquels on modifie un ou quelques paramètres environnementaux (température, alimentation hydrique). Il est également possible de décliner le paramètre environnemental comme une variation temporelle, ce qui est le cas typique des études portant sur la dendroplasticité (Martinez-Meier *et al.*, 2009).

Finalement, le phénotype étudié peut correspondre à des caractères les plus élémentaires (expression de gènes, paramètre anatomiques…) jusqu'à des caractères les plus complexes issus de l'intégration de nombreux caractères fonctionnels (croissance, taux de dissémination des graines… Violle *et al.*, 2009).

Statistiquement, la décomposition de la variance phénotypique est fréquemment utilisée pour expliquer les variations du caractère observé. Le phénotype est alors considéré comme la résultante d'effets génotypiques, environnementaux et de l'interaction génotype-environnement. La plasticité phénotypique fait intervenir les effets environnementaux (E) et l'interaction Génotype-Environnement (GxE) comme le montre la figure 6 (Pigliucci, 1996). De nombreux indices (revus par Valladares *et al.*, 2006) ont été développés pour tenter de quantifier la variation phénotypique induite par l'environnement, en fonction du but et du contexte de l'étude. Ainsi, la comparaison de la plasticité phénotypique entre espèces ou populations fera intervenir préférentiellement les composantes d'une analyse de variance à deux facteurs et interaction (G, E, et GxE), tandis que pour la comparaison de la plasticité de différents caractères, il sera plus pertinent d'utiliser des indices faisant intervenir des régressions, comme par exemple la pente de la norme de réaction.

3.3.2 Interprétations biologiques

La description des normes de réaction permet de mieux caractériser la réponse des génotypes aux variations environnementales. Ainsi, un certain nombre d'attributs peuvent être définis (figure 7).

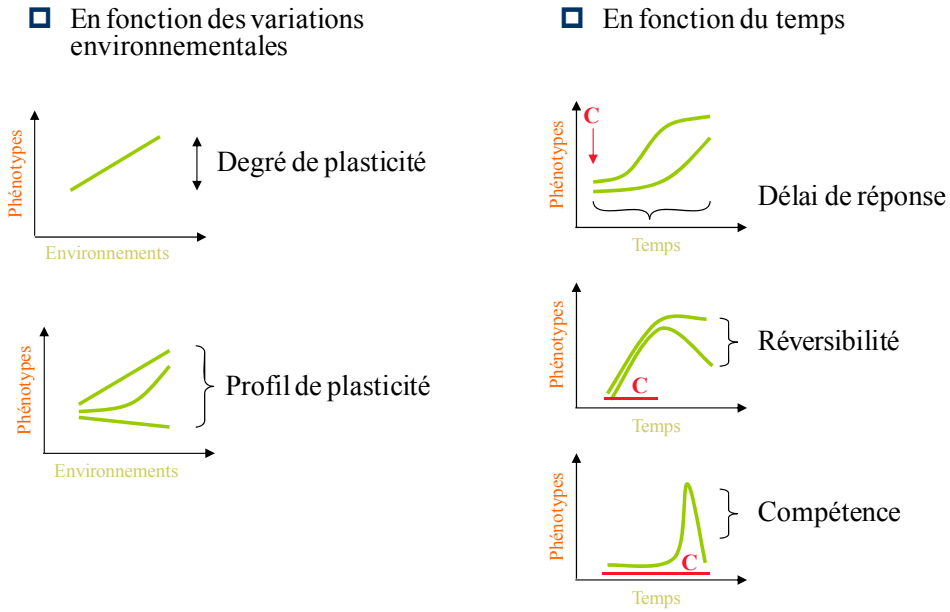


Figure 7 : Paramètres utilisés pour décrire la plasticité à partir de normes de réaction construites en fonction du temps, ou de l'environnement. C : contrainte environnementale.

Si l'on étudie des variations spatiales, on peut définir le degré de plasticité (amplitude de la réponse) et le profil de plasticité (la réponse est-elle linéaire ou plus complexe? De pente

positive ou négative?). Au niveau temporel, on peut définir un délai de réponse, la réversibilité potentielle de la réponse, et observer des phénomènes de compétence, qui correspondent à la capacité ou non de l'organisme à répondre à des stimuli environnementaux à tous les stades de son développement ontogénique.

Une des difficultés d'interprétations de la plasticité phénotypique est principalement liée à l'intégration de l'ontogénie de l'organisme étudié, c'est-à-dire la succession de stades de développement de l'individu depuis sa naissance jusqu'à sa mort. La plasticité passive (variation d'un caractère comme une conséquence de la variation de la croissance induite par l'environnement) se distingue de la plasticité ontogénétique (variation dans la trajectoire ontogénique d'un caractère induite par l'environnement). Les conclusions concernant la plasticité phénotypique peuvent ainsi varier considérablement si des changements ontogéniques dans l'expression phénotypique ne sont pas prises en compte (Wright et McConnaughay, 2002). Par exemple, Climent (publié dans Chambel *et al.*, 2005) a montré que *Pinus canariensis* ne montre aucune différence d'allocation de biomasse entre des environnements humides et secs lorsque l'on prend en compte la taille de la plante (figure 8). Pourtant, la biomasse sèche et le ratio biomasse racinaire / aérienne sont significativement différents entre les environnements, ceci étant dû uniquement aux effets de taille. Pour séparer les effets de l'ontogénie de la plasticité vraie, il semble donc important d'incorporer des covariables liées à la taille de l'arbre aux modèles, ou de comparer des relations allométriques (Poorter et Nagel, 2000).

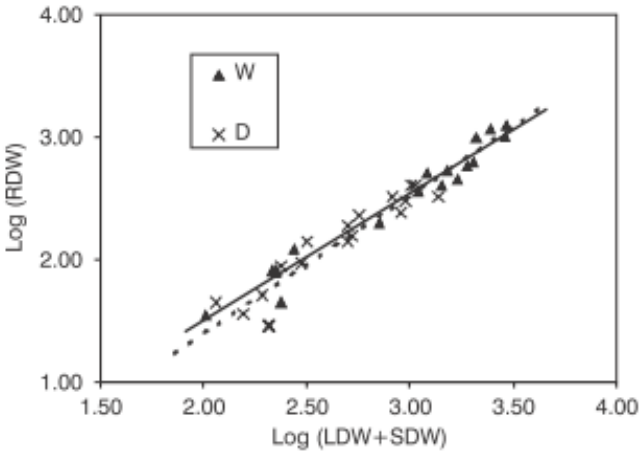


Figure 8 : Régressions linéaires entre trajectoires ontogéniques de compartiments de la biomasse (log transformée) pour des plants de *Pinus canariensis* irrigués (W) et en déficit hydrique (D). RDW : biomasse sèche racinaire, LDW : biomasse sèche foliaire, SDW : biomasse sèche des tiges. Lorsqu'on prend en considération la taille des plants (représentés ici par la somme des deux biomasses aériennes), l'allocation de biomasse entre système racinaire et aérien n'est pas différente entre les deux traitements. Pourtant, à la fois la biomasse sèche et le ratio biomasse racinaire sur biomasse aérienne sont significativement différentes entre les deux traitements (données issues de Climent *et al*., publiées dans Chambel *et al*., 2005).

3.3.3 La plasticité est-elle adaptative ?

Au final, la question clé pour les enjeux d'adaptation est de savoir quel rôle va jouer la plasticité phénotypique sur la valeur sélective des individus (aussi appelée fitness) (Alpert et Simms, 2002), c'est-à-dire si la plasticité a une valeur adaptative. La fitness peut être considérée au sens biologique (capacité d'un individu à transmettre ses allèles à la génération suivante, comme par exemple la quantité de graines ou de pollens produits) pour les espèces de forêts naturelles ou au sens agronomique (capacité d'un individu à correspondre aux attentes du sélectionneur, comme par exemple la croissance ou le rendement) pour les espèces de plantations.

Les réponses plastiques sont dites adaptatives si elles permettent aux individus de maintenir leur fitness par ajustement fonctionnel à travers de larges spectres environnementaux (figure 9). La sélection favorise alors la variation phénotypique. La plasticité phénotypique peut également être neutre si elle n'augmente pas la fitness du génotype porteur. Elle n'est alors pas soumise à sélection. Mais la plasticité phénotypique peut également être délétère, si elle diminue la fitness du génotype porteur, et elle est alors contre-sélectionnée (Alpert et Simms, 2002).

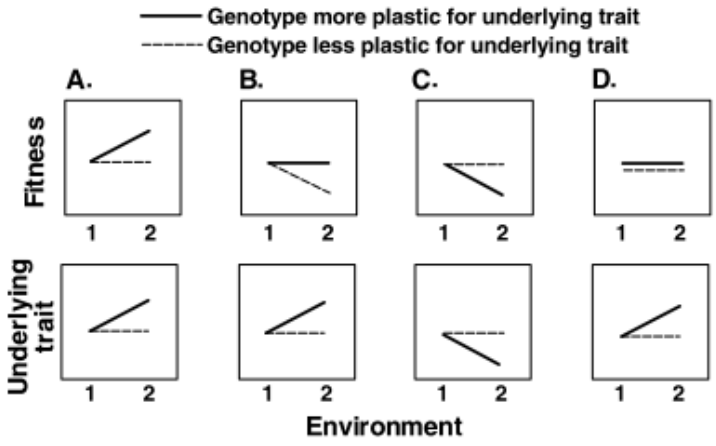


Figure 9 : Différents types de plasticité : (A) plasticité adaptative du trait sous-jacent avec une augmentation de la fitness ; (B) plasticité adaptative du caractère sous-jacent qui permet de maintenir la fitness ; (C) plasticité délétère du caractère sous-jacent qui diminue la fitness; (D) plasticité neutre du caractère sous-jacent, qui n'affecte pas la fitness (issu de Alpert et Simms, 2002).

3.3.4 Les contraintes s'opposant à la mise en place de la plasticité adaptative

Un certain nombre de contraintes peuvent s'opposer à la sélection de la plasticité phénotypique.

Tout d'abord, l'expression d'un phénotype plastique plutôt que celle d'un développement fixé peut réduire la fitness d'un génotype, car cela nécessite la mise en place de différents mécanismes qui engendrent des coûts, tels que l'acquisition de l'information environnementale (surtout chez les organismes non-sessiles), la maintenance de la machinerie sensorielle et régulatrice, la production de structures pour le développement plastique. L'instabilité développementale et les corrélations génétiques engendrent également des coûts inhérents à la plasticité phénotypique (van Kleunen et Fischer, 2005). L'instabilité développementale se manifeste comme des erreurs de développement reflétant des variations intra-individuelles. Si la plasticité est associée avec des imprécisions du développement, cela peut avoir des conséquences en terme de fitness (De Witt *et al.*, 1998). Finalement, les coûts génétiques sont causés par des corrélations négatives entre la plasticité d'un caractère et la fitness. Ils peuvent être la conséquence de la pléiotropie (le produit d'un gène influe sur plusieurs caractères), de la liaison génétique (deux allèles de deux gènes différents sont transmis ensemble d'un individu à sa descendance), ou de l'épistasie (interactions entre gènes), impliquant des gènes à la fois liés à la variation de fitness et au caractère plastique (van Kleunen et Fischer, 2005).

De plus, l'utilisation de la plasticité présente certaines limites (De Witt *et al.*, 1998). Les indices utilisés par la plante pour évaluer l'environnement peuvent manquer de fiabilité. Le délai de réponse peut être trop long pour répondre à temps à la variation environnementale. C'est le cas par exemple de modifications morphologiques qui ont un temps de mise en place assez long. De plus, un individu généraliste a plus de risques d'être moins performant qu'un individu spécialisé sur son environnement de prédilection. Enfin, le problème de l'épi-phénotype peut également limiter l'intérêt de la plasticité phénotypique : la variation du caractère, intégrée au cours du développement juvénile, est souvent plus performante que lorsqu'elle intervient de façon additive au phénotype original.

L'existence de telles contraintes demande donc de prendre en compte un certain nombre de paramètres pour pouvoir analyser ce phénomène si complexe de plasticité phénotypique et de présumer de son avantage adaptatif.

3.3.5 La sélection de la plasticité phénotypique

Dans le cadre de l'adaptation aux contraintes environnementales, identifier des réponses plastiques fonctionnellement adaptatives malgré leur coût est un enjeu majeur pour l'amélioration génétique des arbres forestiers. Il serait alors possible de décrire et d'intégrer, dans les processus de gestion des populations d'amélioration, la diversité génétique en présence pour ces réponses adaptatives, afin d'améliorer leur résilience.

A l'heure actuelle, les principaux caractères sélectionnés chez les espèces de plantation sont liés à l'augmentation de la productivité, comme les caractères de croissance (hauteur et circonférence), de rectitude et de branchaison, de résistance aux ravageurs ou de qualité du bois. L'amélioration génétique pour la croissance présente une contrainte majeure : ce caractère a une héritabilité faible à moyenne avec un fort effet de l'environnement sur la variation du phénotype notamment au jeune âge (Kube *et al.*, 2001 ; Grans *et al.*, 2009 ; Stackpole *et al.*, 2010).

On peut alors imaginer prendre en compte dans les programmes d'amélioration des critères qui renseignent sur la capacité de l'arbre à interagir avec son environnement de façon à favoriser sa croissance. Pour qu'un caractère soit pris en compte dans les programmes d'amélioration dans de telles circonstances, il faut i/ qu'il apporte un avantage en terme d'interaction de l'arbre avec son environnement, ii/ qu'il soit variable iii/ qu'il soit héritable iv/ qu'il soit positivement corrélé avec les caractères cibles de la sélection et v/ qu'il soit facilement mesurable à grande échelle. En général, il semble que les héritabilités calculées pour les caractères plastiques sont généralement plus faibles que pour des caractères classiques (Scheiner, 1993), mais de nombreuses études montrent que la variation génétique pour les caractères plastiques est relativement élevée (van Kleunen et Fischer, 2005).

Trouver le (ou les) caractère (s) pertinent(s) pour intégrer la plasticité phénotypique à un programme d'amélioration ou, dans une moindre mesure, pour analyser l'impact de la sélection de caractères d'intérêt (comme la croissance) sur la capacité d'adaptation des arbres aux changements climatiques n'est pas tâche facile. Ayant peu d'informations sur les caractères qui gouvernent la plasticité phénotypique des arbres, il a fallu, dans un premier temps, bien connaitre l'impact des contraintes environnementales sur l'espèce analysée et la variation des caractères qu'elles vont engendrer.

Dans le cadre de cette thèse, nous nous sommes intéressés à la contrainte hydrique car c'est une des contraintes environnementales qui semble le plus influencer la croissance, donc la productivité des arbres (Stape *et al.*, 2010).

4 Quelles réponses à la contrainte hydrique ?

4.1 L'eau dans le sol

Les qualités hydriques d'un sol ne se déclinent pas uniquement avec le pourcentage d'eau qu'il contient. En effet, différentes forces interviennent en fonction de la composition physique et chimique des sols, pour retenir ou repousser l'eau. Elles peuvent être des forces osmotiques (dues à l'attraction de l'eau par les ions) et/ou des forces matricielles (dues à la nature poreuse du sol et à la présence de minéraux argileux), composées de forces capillaires et de forces d'adsorption (Calvet, 2003).

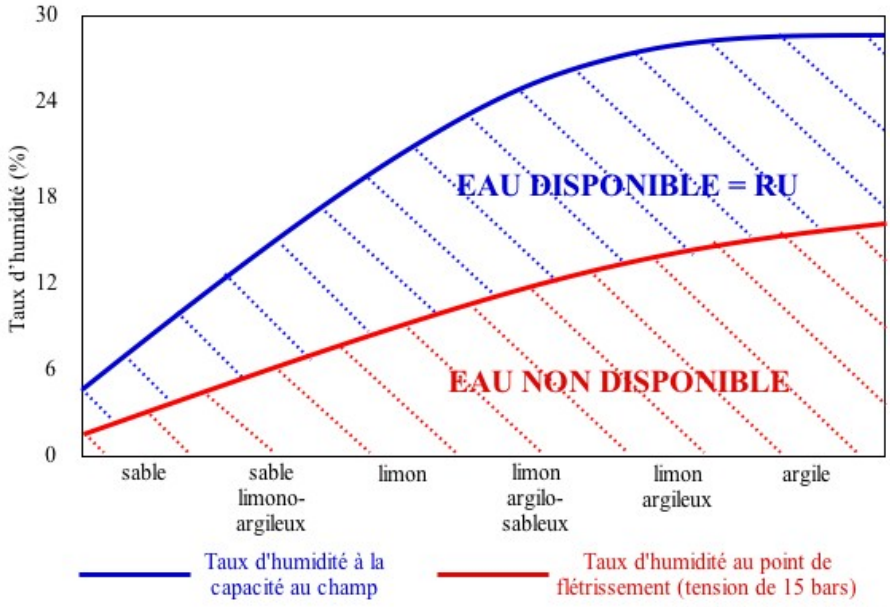


Figure 10 : En fonction de la composition des sols, le taux d'humidité minimal pour que l'eau du sol soit disponible pour la plante est variable. La capacité au champ est la quantité d'eau que le sol est capable de retenir contre les forces de gravitation. Le point de flétrissement permanent est la valeur du potentiel hydrique du sol à laquelle les plantes ne peuvent retrouver la pression de turgescence, même la nuit, en l'absence de transpiration. L'écart entre ces deux valeurs constitue la réserve utile (d'après Buckman et Brady, 1965).

On peut alors définir différents paramètres pour caractériser un sol et son utilisation agronomique. Le potentiel hydrique est la mesure de la pression requise pour extraire l'eau du sol. Par analogie au pH, les pédologues utilisent le concept du pF qui est le logarithme du potentiel hydrique, car celui-ci augmente très rapidement avec le dessèchement du sol.

L'échelle logarithmique permet alors de comparer aisément humidité du sol et potentiel hydrique. L'humidité équivalente correspond à la capacité de rétention maximale du sol. Au dessus de cette valeur, le sol est gorgé d'eau et il y a ruissellement. Le point de flétrissement donne l'humidité minimale du sol, en dessous de laquelle la plante flétrit de façon irréversible. La réserve utilisable est la différence entre ces deux humidités et indique la fourchette d'humidité du sol pour laquelle la plante pourra se développer et le sol ne dégorgera pas (figure 10).

4.2 L'eau dans la plante

Le transit de l'eau dans la plante s'effectue en fonction du potentiel hydrique des différentes régions traversées. En climat tempéré dans les régions humides, le potentiel hydrique d'un sol est très proche de zéro, tandis que le potentiel hydrique de l'atmosphère est de l'ordre de −100MPa. L'eau passe donc de manière purement physique du sol vers l'atmosphère en provoquant un flux physique continuel appelé continuum sol/plante/atmosphère (figure 11). L'eau et les nutriments que le sol contient sont absorbés par les racines (au niveau de poils absorbants) par un mécanisme osmotique, qui va conduire l'eau jusqu'au cylindre central par voie apoplastique (à travers la paroi), symplastique (à travers les vacuoles et le cytoplasme) et intercellulaire. C'est la poussée racinaire. La transpiration foliaire constitue alors le moteur principal pour conduire cette sève brute jusqu'aux feuilles à travers le xylème. L'eau arrivant du xylème diffuse par voie apoplasmique dans le mésophylle, et une grande partie est évaporée via les stomates. Cette eau joue de nombreux rôles dans la plante : donneur d'électrons lors de la photosynthèse, solvant pour les molécules hydrophiles. Elle constitue aussi le milieu et parfois le substrat de nombreuses réactions chimiques. De plus, cette eau permet l'organisation des structures cellulaires en assurant des fonctions mécaniques (maintien de la turgescence) ou physiologiques (l'eau apoplasmique relie différentes cellules d'un tissu). Finalement, elle sert également à véhiculer les produits de la photosynthèse sous forme de sève élaborée.

La teneur en eau des plantes est régulée par le maintien d'un équilibre entre absorption d'eau et transpiration. Si l'absorption d'eau est supérieure à la transpiration, on peut observer chez certaines espèces comme les graminées des phénomènes de guttation pour éviter une surpression hydrique dans les tissus foliaires. Dans le cas inverse, si l'absorption est insuffisante, la transpiration va être réduite, via le facteur essentiel de cette régulation : la fermeture des stomates. C'est souvent le cas lorsque la plante subit une contrainte hydrique.

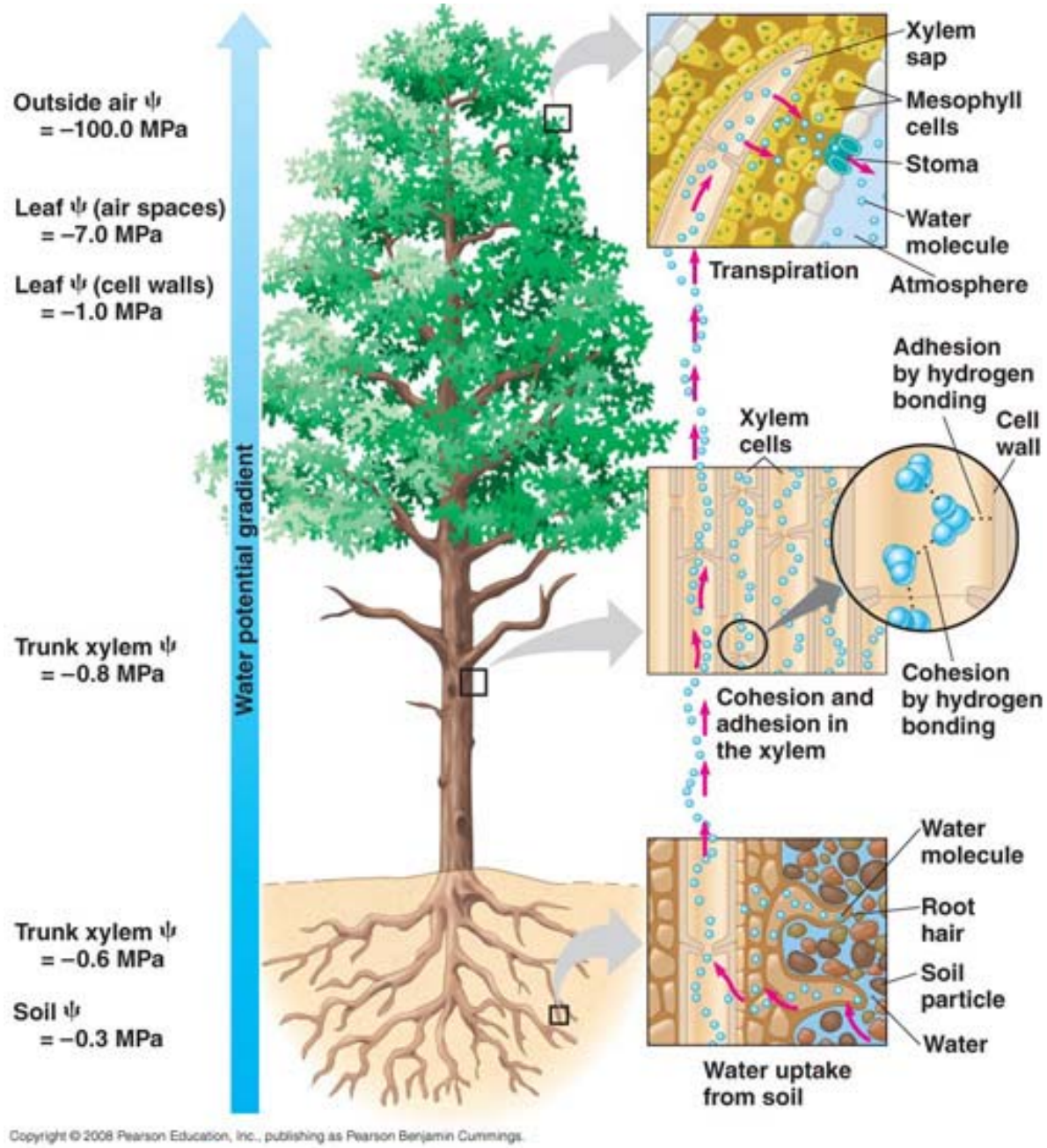


Figure 11 : Le transport de l'eau dans les arbres suit les potentiels hydriques (Ψ) décroissants. Après absorption de l'eau du sol au niveau des poils racinaires, l'eau est transportée à travers le xylème jusqu'au mésophylle. Une partie est évaporée à travers les stomates lors de la transpiration.

(http://www.bio.miami.edu/dana/226/226F09_10.html).

4.3 Les stratégies mises en place en réponse à la contrainte hydrique

Chez les plantes, différentes stratégies de réponse au stress hydrique ont été décrites (Levitt, 1972).

- L'échappement concerne les plantes vivant en régions arides, la plante effectue son cycle végétatif avant qu'il n'y ait un déficit en eau et effectue rapidement son cycle reproducteur après une pluie. Les graines produites par ces types de plantes sont capables de résister à de longues périodes de sécheresse. Les plantes longévives n'appliquent pas cette stratégie.

- L'évitement s'applique à un plus grand nombre d'espèces. Cette stratégie consiste à maintenir le potentiel hydrique et à limiter au maximum l'évapotranspiration. Pour limiter les pertes d'eau, certaines plantes ferment rapidement leurs stomates, d'autres réduisent leur surface foliaire (perte de feuilles et/ou production de feuilles modifiées généralement plus petites).

- Certaines espèces tolérantes sont dotées de mécanismes leur permettant de supporter le stress hydrique, lorsque le potentiel hydrique diminue, elles ajustent le potentiel osmotique de leurs cellules afin de maintenir leur turgescence et ainsi conserver leur rigidité.

En réalité, le déficit hydrique est composé de différentes formes de contraintes associées : la diminution de la disponibilité en eau dans le sol, une augmentation du stress osmotique, une diminution de l'humidité atmosphérique, une augmentation de la température, une augmentation de l'intensité lumineuse, etc… Les situations de stress hydrique sont fortement limitantes pour la productivité du fait que les plantes doivent avant tout assurer leur survie (Chaves *et al.*, 2003).

4.4 La réponse à l'échelle de la plante

Face à la contrainte hydrique, les plantes ont une gamme de réponses assez complexes, qui varient au cours du temps en fonction des espèces, néanmoins, un certain nombre de réponses sont communes à la plupart des végétaux. D'une manière générale, les mécanismes mis en place visent à conserver l'eau dans la plante en augmentant l'absorption de l'eau mais aussi en diminuant les pertes (figure 12, Chaves *et al.*, 2003).

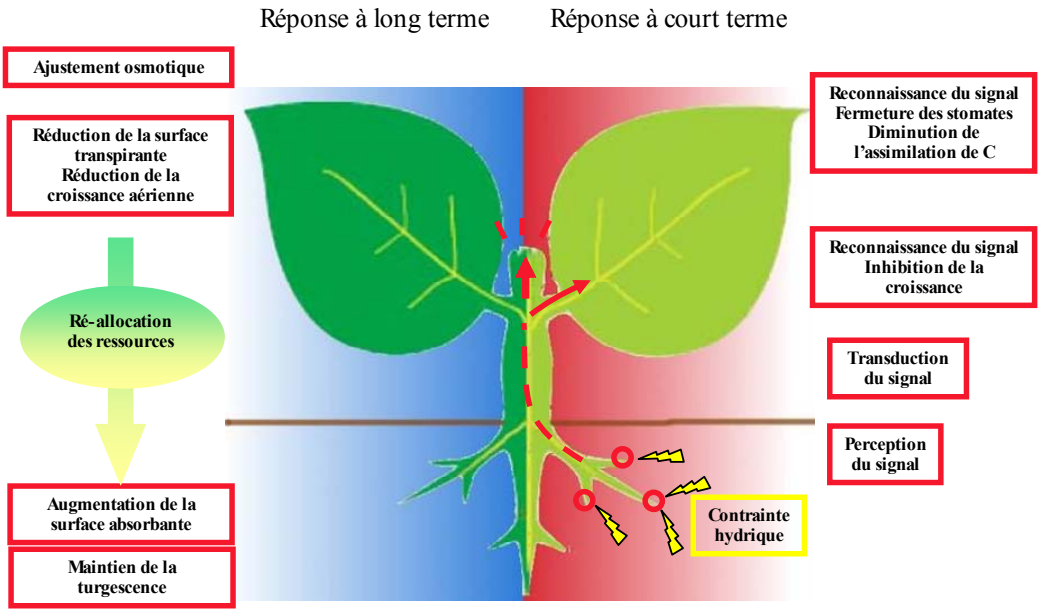


Figure 12 : Diagramme représentant les réponses à la contrainte hydrique possibles de la plante, à court ou à long terme, adapté de Chaves *et al.* 2003.

4.4.1 Les réponses communes

Différents ajustements permettent d'optimiser l'absorption d'eau : tout d'abord, la réallocation des ressources nutritives vers les racines afin de favoriser l'exploration racinaire. L'efficacité de l'absorption de l'eau dans le sol dépend à la fois de leur extension spatiale et de la densité du système racinaire (Levitt, 1980). Ces réallocations peuvent aller jusqu'à la sénescence des feuilles âgées pour recycler les nutriments qu'elles contiennent dans les feuilles et tiges juvéniles (Chaves *et al.*, 2003). Ce phénomène de sénescence permet également de réguler les pertes en eau, car en diminuant sa surface foliaire totale, la plante réduit sa transpiration. La diminution de la surface foliaire peut également prendre la forme d'un arrêt ou d'une diminution de la croissance des feuilles, voire la réduction du nombre et de la viabilité des bourgeons foliaires (Bréda *et al.*, 2006), ce qui a plus long terme limitera le développement du houppier. L'enroulement des feuilles et la variation des angles d'insertion des feuilles permettent également de limiter les pertes en eau en réduisant l'absorbance de la lumière (Ehleringer et Cooper, 1992). Néanmoins, le phénomène majeur permettant la limitation des pertes en eau réside en la fermeture des stomates, ce qui permet de réguler la vitesse de transpiration de l'eau par rapport à la vitesse d'absorption par les racines.

Un des premiers processus touché par la contrainte hydrique est l'activité photosynthétique. Cette diminution d'activité peut être liée à la fermeture stomatique ou à un affaiblissement métabolique. En effet, la fermeture des stomates engendre consécutivement une diminution de la diffusion de CO_2 à l'intérieur des chloroplastes (Lawlor, 1995). Cela a pour conséquence

directe l'interruption du cycle de Calvin, la RuBisCo n'ayant plus de substrat CO_2, et donc la production en composés organiques devient limitée. D'autre part l'augmentation de la concentration en ions dans le chloroplaste, particulièrement des ions Mg^{2+}, perturbe le flux de protons nécessaire à la synthèse de l'ATP.

La plante doit également se protéger contre les contraintes primaires qui entrainent des réponses spécifiques (variations de l'osmolarité cellulaire) ou secondaires qui engendrent des réponses non-spécifiques au stress hydrique (accumulation d'espèces activées d'oxygènes (ROS), augmentation de la température foliaire, carences minérales) (Beck *et al.*, 2007).

Pour maintenir la turgescence cellulaire et stabiliser les protéines et structures cellulaires, la concentration en ions est augmentée grâce à l'accumulation d'osmolytes. Ces osmolytes peuvent être des sucres comme le raffinose, le sucrose, le tréhalose ou le sorbitol, des alcools de sucre comme le mannitol, des acides aminés comme la proline, des amines comme la glycine-bétaïne ou des polyamines.

Les espèces activées d'oxygène (ROS) sont majoritairement formées au niveau des chloroplastes. En effet, au cours de la photosynthèse, lorsque la fixation de CO_2 est limitée par la fermeture des stomates, les électrons se fixent plus qu'usuellement à l'oxygène, ce qui conduit à une accumulation de ROS dans le chloroplaste (Dat, 2000). D'autres ROS peuvent également être produites au cours de la photorespiration quand la RuBisCo utilise l'O_2 à la place du CO_2 et forme des peroxydes d'oxygène, ou par l'auto-oxydation des composants réduits de la chaine respiratoire dans les mitochondries. Les ROS telles que les ions superoxydes, les peroxydes d'oxygène et les radicaux hydroxyles peuvent être des signaux induisant des mécanismes protecteurs (Cf § 4.5.2.2) aussi bien que des agents qui endommagent les cellules. En effet ces molécules hautement réactives inhibent la capacité de réparer les dommages causés au photosystème II, inactivent les enzymes par oxydation de leur groupement thiol, mais oxydent aussi diverses protéines, lipides et acides nucléiques (Halliwell et Gutteridge, 1999). Les mécanismes de détoxification des ROS peuvent être des mécanismes non enzymatiques ou enzymatiques. Les antioxydants non enzymatiques comprennent majoritairement : l'ascorbate (vitamine C), le glutathion, le tocophérol (vitamine E), les flavonoïdes, les alcaloïdes et les caroténoïdes. Les mécanismes enzymatiques comprennent les enzymes superoxyde-dismutases, peroxydases et catalases (Bartels et Sunkar, 2005).

4.4.2 La spécificité des arbres

Face à la contrainte hydrique, les arbres se distinguent des autres plantes supérieures par deux caractéristiques principales : i/ leur croissance bien plus importante implique un système vasculaire plus étendu et de plus grands besoins en eau, ii/ leur longévité peut les exposer à subir de façon répétée des variations de l'environnement au cours de leur cycle de vie.

Comme nous l'avons vu précédemment, la transpiration par les feuilles constitue le principal moteur du flux d'eau à travers la plante. Donc plus la surface foliaire d'une plante sera développée, plus ses besoins en eau seront conséquents. Or, chez les arbres à maturité la surface foliaire est nettement plus élevée que chez les autres plantes, ce qui conduit à une plus grande transpiration, et donc un plus grand besoin d'eau.

Une autre spécificité des arbres est leur taille. L'ascension de la sève jusqu'aux feuilles est réalisée grâce à la tension exercée par la transpiration. Celle-ci est assurée par capillarité et est donc conditionnée par une continuité hydraulique à travers toute la plante (Tyree et Ewers, 1991). Or, chez les arbres, la continuité hydraulique doit être assurée sur une grande distance. Lors du déficit hydrique, le potentiel hydrique au niveau des feuilles peut atteindre des valeurs très négatives, et la tension que subissent les colonnes d'eau où est conduite la sève brute gagne en intensité. Sous de telles tensions, il est possible que la colonne d'eau se rompe (on parle alors de cavitation) : le conduit se remplit d'air et ne peut plus conduire la sève.

Pour maintenir l'intégrité de leur système hydraulique, les arbres peuvent déployer différentes stratégies (Breda et al .,2006): i/ mettre en place des processus de régulation à court terme et réversibles tels qu'une réduction de la transpiration par fermeture des stomates, ou ii/ utiliser des réponses à plus long terme comme le renforcement du xylème grâce aux lignines pour résister à la cavitation, réduire leur surface transpirante soit par réallocation de biomasse ou sénescence, et améliorer la conductance hydraulique, l'augmentation de la biomasse racinaire semblant être alors la meilleure stratégie dans ce but.

4.5 La réponse au niveau moléculaire

Les recherches menées sur la réponse moléculaire à une contrainte abiotique ont essentiellement été menées chez *Arabidopsis thaliana* (Seki *et al.*, 2002, Bray 2002), et chez les plantes de grandes cultures (Zheng *et al.*, 2004 ; Yang *et al.*, 2004). Les résultats obtenus ont permis d'avancer l'hypothèse que la réponse à un stress n'est pas le fait de l'expression d'un seul gène mais de l'expression coordonnée d'un grand nombre d'entre eux. Les délais de réponses variés de ces gènes (de la seconde à la semaine) ont amené Sunkar et Bartels (2002)

à émettre l'hypothèse que les gènes de la réponse précoce fourniraient une protection initiale et une amplification du signal alors que ceux qui répondraient tardivement pourraient être impliqués dans la tolérance au stress hydrique. Les gènes induits par contrainte hydrique peuvent être classés en 2 groupes selon les protéines pour lesquels ils codent (Shinozaki, 2007):

- protéines de régulation qui interviennent dans la transduction du signal et l'expression de gènes de réponse au stress.
- protéines de structures et enzymes qui interviennent dans la tolérance aux stress abiotiques.

4.5.1 La perception du signal de stress

A priori, le changement de pression osmotique dans les cellules constitue le déclenchement majeur de la réponse moléculaire à la contrainte hydrique chez les plantes au niveau moléculaire (Shinozaki et Yamaguchi-Shinozaki, 1997). Comme aucune molécule végétale n'a été réellement identifiée comme osmosenseur, savoir comment la plante ressent le signal de stress reste encore une question non élucidée. Certains homologues d'osmosenseurs de levures ont été trouvé chez Arabidopsis, comme *AtHK1* (Urao *et al.*, 1999). D'autres candidats ont été proposés en prenant en compte leur profil expressionnel et leur fonction, tel que le gène NtC7 qui code pour une protéine membranaire réceptrice (Tamura *et al.*, 2003) ou Cre1qui code pour un récepteur de cytokinine (Reiser *et al.*, 2003).

4.5.2 Les gènes de régulation impliqués dans la transduction du signal

4.5.2.1 Transduction du signal ABA-dépendante et indépendante

L'utilisation de mutants insensibles à l'ABA et de mutants déficients pour l'ABA, a montré que le signal « stress hydrique » peut emprunter deux types de voies de transduction majeurs : l'un où l'expression des gènes induits dépend de l'ABA et l'autre où cette expression est indépendante de l'ABA (figure 13, Bray, 1997 ; Shinozaki et Yamaguchi-Shinozaki, 1997). Ces voies de transduction nécessitent la synthèse de facteurs de transcription qui se fixent sur une partie du promoteur des gènes cibles, appelé élément de réponse. Ces éléments de réponses fonctionnent alors comme des éléments *cis*, c'est-à-dire qu'ils régulent l'expression de gènes localisés sur la même molécule d'ADN (Shinozaki et Yamaguchi-Shinozaki, 2007). Parmi les gènes ayant une induction ABA-dépendante, des facteurs de transcription de type MYC et MYB (Abe *et al.*, 1997), NAC (Fujita *et al.*, 2004) et bZIP (Uno *et al.*, 2000) ont été mis en évidence. Parmi les gènes ayant une induction ABA-indépendante, les facteurs de

transcription appartenant à la famille ERF/AP2 (Yamaguchi-Shinozaki et Shinozaki, 2005) ainsi que NAC (Tran *et al.*, 2004) ont été décrits.

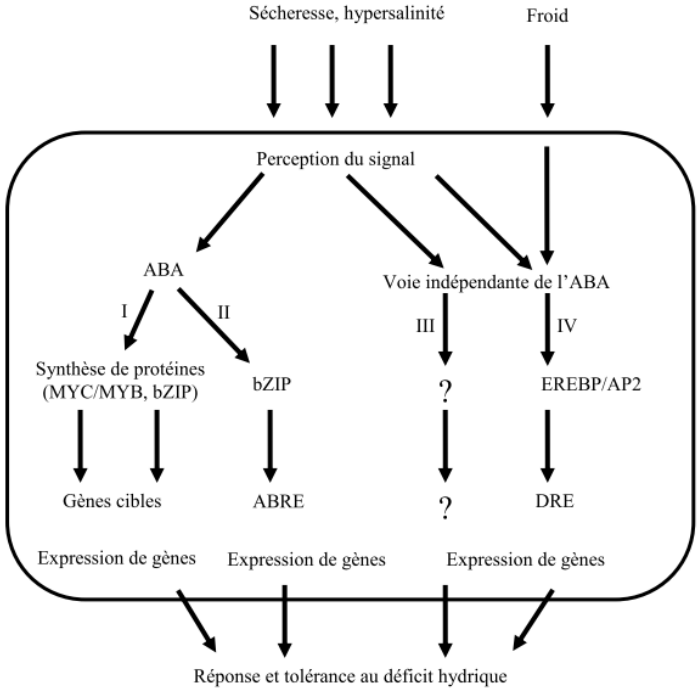


Figure 13 : Les voies de transduction du signal entre la perception du déficit hydrique et l'expression de gènes peuvent être dépendantes ou indépendantes de l'acide abscissique (ABA), et font intervenir des éléments de réponse, sur lesquels peuvent se fixer des facteurs de transcription (MYC/MYB, bZIP). D'après Shinozaki et Yamaguchi-Shinozaki, 1997.

Si le rôle de l'ABA comme messager du signal de stress hydrique est établi depuis longtemps (Jackson, 1997) et que son rôle dans le contrôle de la fermeture des stomates est avéré (Borel *et al.*, 2001), l'effet de l'ABA sur la croissance des organes reste soumis à controverse et semble dépendre des nombreux facteurs (stade de croissance, température, organe considéré…). Il y a de fortes chances pour que l'ABA interagisse avec d'autres signaux chimiques (pH ou autres hormones) ou avec le transport d'éléments nutritifs dans la sève brute (Bahrun *et al.*, 2002). En effet, il semble que la fermeture des stomates, la croissance et la sénescence des feuilles soient contrôlées en cas d'exposition à une contrainte hydrique, par un réseau complexe de signalisations hormonales incluant, entre autres, l'éthylène (Beaudoin *et al.*, 2000), l'acide jasmonique (Sembdner et Parthier, 1993), l'acide salicylique (Senaratna *et al*, 2000) ou les cytokinines (Tran *et al.*, 2007).

4.5.2.2 Les messagers secondaires

Les messagers secondaires sont des molécules permettant la transduction d'un signal provenant des récepteurs à la surface cellulaire jusqu'à des molécules cibles à l'intérieur de la cellule, dans le cytoplasme ou le noyau. La libération d'un messager secondaire se fait généralement par liaison d'un ligand (par exemple une hormone) sur un récepteur membranaire. Le messager secondaire peut alors induire une cascade de réponses qui permettent d'amplifier le signal, et ainsi de déclencher une réponse cellulaire. Il existe une multitude de messagers secondaires décrits pour répondre à la contrainte hydrique : l'adénosine3'-5' monophosphate ou AMP cyclique (Assman, 1995), l'acide phosphatidique (Munnik, 2001). Nous nous focaliserons dans le paragraphe suivant sur le signal calcique et les espèces activées d'oxygène (ROS).

Le signal calcique a été impliqué dans la réponse à divers stimuli extracellulaires associés à des réponses intracellulaires (Hong-Bo *et al.*, 2008). Jusqu'à présent, trois différentes classes de signaux calciques ont été décrits dans le cas de la contrainte hydrique : les canaux calciques perméables aux ions, les antiports Ca^{2+} / H^{+} et les pompes à calcium (Ca^{2+}-ATPases). Les effecteurs des voies de transduction des signaux calciques peuvent être classés en deux catégories : d'une part, les capteurs primaires sensibles au calcium comme les protéines kinases dépendantes du calcium (CDPK) et d'autre part, les substrats activés en aval qui représentent le dernier effecteur de la voie de signalisation (canaux ioniques, phosphorylation, concentration cytosolique en calcium).

La production de ROS résulte de l'excitation ou de la réduction incomplète d'oxygène moléculaire. Comme nous l'avons décrit auparavant, leur accumulation peut causer des dommages à l'environnement cellulaire en altérant les biomolécules par oxydation. Néanmoins, les variations spatiales et temporelles de concentration en ROS sont également interprétées comme des signaux nécessaires pour la croissance, le développement, la tolérance aux stress abiotiques, la réponse aux pathogènes et la mort cellulaire programmée. Les ROS activeraient les cascades MAPK (protéines kinases activées par les facteurs mitogènes), réguleraient les flux de calcium, et induiraient des facteurs de transcription spécifiques à chaque type de ROS (Gechev *et al.*, 2006). Par exemple, l'oxygène singulet induit des facteurs de transcription des familles ERF et MYB chez *Arabidopsis thaliana*, tandis que le peroxyde d'oxygène régule les protéines de choc thermique.

Avec les interactions possibles entre voies de signalisation, la complexité du réseau de transduction du signal via les messagers secondaires rend la caractérisation de ces messagers difficile, et de nombreuses questions restent encore en suspend.

4.5.2.3 La régulation transcriptionnelle

Comme son nom l'indique, la régulation de la transcription permet le contrôle de l'expression des gènes par action sur le niveau de transcription de l'ADN. Celle-ci est effectuée principalement grâce à des facteurs de transcription qui peuvent être classés en deux catégories : les éléments *cis*-régulateurs (reconnaissance d'élément de réponse directement sur la région promotrice du gène) ou les facteurs *trans*-régulateurs (reconnaissance d'un autre facteur de transcription lié ou de l'ARN polymérase). Les facteurs de transcription sont ainsi constitués de différents domaines : un domaine de fixation à l'ADN, un domaine d'activation de la transcription, et, pour certains facteurs *trans*-régulateurs, un 3ème domaine qui permet de fixer un élément annexe permettant de réaliser l'action d'un message extérieur à la cellule, comme un message hormonal. Les familles de facteurs de transcription ont été ainsi définies selon les domaines de liaison à l'ADN utilisés.

Comme nous l'avons vu précédemment, de nombreux facteurs de transcription sont impliqués dans la réponse à la contrainte hydrique, qui incluent les familles DREB (élément de réponse à la déshydratation), ERF (facteur de réponse à l'éthylène), protéines en doigt de zinc, WRKY, MYB, bHLH (Hélice-Boucle-Hélice), bZIP (facteur à glissière de leucine), NAC et des facteurs à homéodomaine (Shinozaki, 2003).

4.5.3 Les gènes codant pour des protéines de structure et des enzymes

Si l'activité de nombreux gènes peut être modifiée de façon passive par la contrainte hydrique, notamment dans le cas des gènes associés à la photosynthèse ou au métabolisme primaire qui sont sous exprimés, certains gènes ont été décrits comme étant potentiellement impliqués dans les mécanismes de tolérance actifs à la contrainte.

4.5.3.1 Osmorégulation

La diminution du contenu en eau de la plante peut entrainer des déséquilibres de concentration en solutés entre vacuole et cytoplasme. Différents mécanismes sont alors mis en place pour parer aux dégâts que cela pourrait engendrer :

i) l'évacuation des ions excédentaires du cytoplasme ou leur compartimentation dans des vacuoles. Les gènes impliqués codent pour des transporteurs ioniques, comme par exemple le gène *AtNHX1* codant pour un transporteur vacuolaire Na^+/H^+ (Shi et Zhu, 2002).

ii) L'ajustement de la pression osmotique par accumulation de solutés compatibles, tels que la proline, la glycine bétaine, des sucres et des polyols (revue par Seki *et al.*, 2007). Ces composés peuvent être accumulés à de fortes concentrations sans altérer l'activité enzymatique. De nombreux gènes intervenants dans la biosynthèse de ces composés ont été montrés comme impliqués dans la réponse à la contrainte hydrique, et certains ont même été utilisés en transformation génétique pour améliorer la tolérance à la contrainte hydrique chez certaines espèces (Umezawa *et al.*, 2006). Par exemple, l'introgression de gènes lié à la biosynthèse de la bétaine issus d'un micro-organisme améliore la photosynthèse et la croissance chez *Arabidopsis*, lorsque les plants sont soumis à déshydratation.

4.5.3.2 Protection des macromolécules

Lorsque l'environnement cellulaire devient inapproprié (changement de pH, stress osmotique…), la protection des macromolécules telles que les protéines peut être assurée par des protéines chaperonnes. Une protéine chaperonne est une protéine dont la fonction est d'assister d'autres protéines dans leur maturation, en leur assurant un repliement tridimensionnel adéquat. Le rôle des protéines chaperonnes est de prévenir les dommages potentiellement causés par une perte de fonction protéique due à un mauvais repliement tridimensionnel. Une des plus connues de ces protéines est une protéine de liaison se situant dans le réticulum endoplasmique appelée BiP. Isolée chez le soja, la surexpression de cette protéine chez le tabac lui a conféré une plus grande tolérance à la contrainte hydrique avec une amélioration de la croissance (Alvim *et al.*, 2001). Elle semble aussi être impliquée dans la sénescence foliaire (Valente *et al.*, 2009).

Les protéines LEA (Late Embryogenesis Abundant ; Wise, 2003) sont accumulées dans les dernières phases de la formation des graines et lorsque les plantes sont soumises à des déficits hydriques. Bien que leur rôle reste assez mal connu, il pourrait être similaire à celui des protéines chaperonnes (Umezawa *et al.*, 2006). Ces protéines sont associées à la protection des macromolécules, par liaison ou remplacement de l'eau pour les protéines les plus hydrophiles, ou en séquestrant les ions pour les protéines les plus hydrophobes (Bartels and Sunkar, 2005 ; Battaglia *et al.*, 2008).

4.5.3.3 Protéines de détoxification

Comme nous l'avons vu précédemment, la production de ROS peut permettre la transduction du signal de stress, tout en ayant des conséquences délétères sur le fonctionnement cellulaire. Des mécanismes de détoxification des ROS peuvent alors se développer afin de protéger les structures cellulaires. Ces réactions comprennent des mécanismes non enzymatiques, comme

l'oxydation de l'ascorbate et du glutathion. On sait également que des molécules telles que les tocophérols, flavonoïdes, alcaloïdes et caroténoïdes sont impliqués dans la détoxification non enzymatique des ROS, mais les mécanismes mis en jeu sont encore mal connus (Apel et Hirt, 2004). Les mécanismes enzymatiques comprennent les superoxydes dismutases, les peroxydases et les catalases (Bartels and Sunkar, 2005).

4.5.3.4 Aquaporines

Le contrôle des flux d'eau est un phénomène majeur pour permettre aux cellules et aux tissus de s'adapter à la contrainte hydrique. Parmi les protéines facilitant les mouvements d'eau, les aquaporines jouent un rôle majeur. Elles facilitent la diffusion de l'eau en formant des pores spécifiques dans la membrane. Les aquaporines sont membres d'une famille de protéines appelées protéines majeures intrinsèque (Major Intrinsic Protein, MIP), parmi lesquelles on retrouve celles localisées dans le tonoplaste (TIP), et dans la membrane plasmique (PIP). En général, les aquaporines sont très exprimées dans les tissus à croissance rapide et dans les cellules impliquées dans le passage de flux d'eau à haut débit. Si plusieurs études ont montré que les gènes codant pour des aquaporines sont induits par la déshydratation, permettant une plus grande perméabilité à l'eau et ainsi de faciliter le flux d'eau (Maurel et Chrispeels, 2001), il y a aussi des exemples d'aquaporines réprimées en cas de déshydratation, ce qui induirait une diminution de la perméabilité membranaire pour permettre la conservation de l'eau cellulaire (Bartels and Sunkar, 2005). Ainsi, la régulation de l'expression des aquaporines pourrait être fonction du délai de réponse, ainsi que de l'organe considéré (Maurel, 2007).

4.6 L'eucalyptus sous contrainte hydrique

L'eucalyptus a la réputation d'être un grand consommateur d'eau (Zahid *et al.*, 2010). De nombreuses études se sont attachées à décrire l'impact de l'alimentation hydrique sur la productivité, décortiquer le flux d'eau dans l'arbre, ainsi que mettre en évidence les mécanismes de tolérance à la contrainte hydrique. Ces études ont été conduites en serre (Merchant *et al.*, 2007 ; Martins *et al.*, 2008 ; Teixeira *et al.*, 2008), en plantation (Stape *et al.*, 2010 ; Warren *et al.*, 2007 ; Campion *et al.*, 2006) ou même en populations naturelles (MacFarlane *et al.*, 2010 ; Merchant *et al.*, 2007 ; Zeppel *et al.*, 2004).

En ce qui concerne la production, la majorité des études montrent que l'alimentation hydrique semble plus limitante que l'alimentation minérale. Par exemple au Brésil, Stape *et al.* (2010) ont montré que l'augmentation de fertilisation par rapport aux fertilisations traditionnellement effectuées n'augmente pas la productivité (autour de 46 m^3/ha/an), tandis que l'irrigation augmente la croissance de 30% (62 m^3/ha/an). La même tendance a été observée par Campion

et al. (2006) en Afrique du sud sur les biomasses aériennes et racinaires. Différentes études ont montré des capacités différentes des génotypes à maintenir leur croissance quand l'alimentation hydrique est limitée (Tatagiba *et al.*, 2009; Silva *et al.*, 2004), permettant ainsi de définir certains génotypes mieux adaptés pour certaines aires de répartitions.

Au niveau foliaire, les eucalyptus ont des feuilles relativement épaisses avec des faibles concentrations en azote et phosphore dues aux faibles teneurs en éléments minéraux des sols dans lesquels ils sont installés (Whitehead et Beadle, 2004). Leur index de surface foliaire (Leaf Area Index) augmente avec la disponibilité hydrique, ceci conduisant à de plus forts taux de transpiration (Hubbard *et al.*, 2010). Les feuilles ont un turn-over relativement long, pouvant aller jusqu'à 3 ans pour les espèces en régions tempérées, et leur croissance est maximale au cours des saisons les plus humides (Whitehead et Beadle, 2004). Lors de périodes de sécheresse prolongées, la sénescence des feuilles est très commune (Nouvelon *et al.*, 2010 ; Pook *et al.*, 1997). Sur les populations naturelles, les angles foliaires sont plus élevés dans les régions les plus sèches de l'Australie, tandis que les feuilles sont plus proches de l'horizontal dans les régions plus tempérées (King, 1997). Cette adaptation réduit la charge de chaleur subie en diminuant l'interception lumineuse.

Le bilan hydrique de plantations d'eucalyptus fait appel à différents paramètres qui ont été évalués afin d'améliorer les pratiques sylvicoles (Zhang *et al.*, 1996). Parmi ces paramètres, l'interception des précipitations par la canopée a été estimée sur divers peuplements (en Australie, Israël, Inde, Brésil) entre 10 et 34% des précipitations annuelles (Whitehead et Beadle, 2004). L'estimation des taux de transpiration est un des paramètres les plus analysés. Différentes méthodologies ont été développées (revues par Whitehead et Beadle, 2004) : les mesures des flux de sève sont les plus utilisées (MacFarlane *et al.*, 2010), mais il est également possible de mesurer les taux de transpiration par eddy covariance (mesure de la turbulence de l'atmosphère ; Cabral *et al.*, 2010), par mesure d'échange gazeux (Shem *et al.*, 2009) ou de les estimer grâce à l'équation de Penman-Monteith reliant l'évapo-transpiration potentielle à diverses données métorologiques (Carneiro *et al.*, 2008). Les différentes études montrent que la transpiration est réduite lorsque la température ou le déficit de pression de vapeur d'eau augmente ou lorsque l'humidité relative diminue (Cook *et al.*, 1999). Ce phénomène est particulièrement observable au cours de la saison sèche où la transpiration est rarement plus élevée que 2mm d'eau transpirée par jour, tandis qu'elle peut atteindre jusqu'à 8mm sur des plantations irriguées (Whitehead et Beadle, 2004). La régulation de la transpiration est une des questions clés de recherches en écophysiologie menées sur l'eucalyptus. Les différences saisonnières et entre espèces ou population pourraient être liées

à l'interaction entre surface foliaire, déficit de saturation de l'air, déficit hydrique du sol et potentiel hydrique foliaire (Whitehead et Beadle, 2004). La sensibilité de la conductance stomatique a été décrite comme un paramètre clé pour la régulation de la transpiration (Warren et al., 2007), bien que ce paramètre semble être plus sensible à un déficit hydrique du sol, qu'à un déficit de pression de vapeur d'eau dans l'air (Warren *et al.*, 2008). Pour conserver un taux de transpiration nécessaire à leur croissance, les eucalyptus semblent avoir une capacité importante à développer leur système racinaire à de très grandes profondeurs, et de façon opportuniste en profitant soit de gradients hydriques dans le sol soit des pluies (Drake *et al.*, 2009). Ainsi, Thongo-M'bou *et al.* (2008) ont montré au cours de la première année de croissance, que la formation de racines fines était fortement liée au contenu hydrique du sol, l'élongation des racines fines à la fin de la saison des pluies étant 11 fois supérieure à celle mesurée à la fin de la saison sèche au Congo.

Les eucalyptus, ayant un métabolisme en C3, semblent avoir une activité photosynthétique relativement élevée par rapport aux autres espèces pérennes, ceci étant principalement dû aux taux maximum de carboxylation et de transport d'électrons (revues par Whitehead et Beadle, 2004). La photosynthèse semble être particulièrement sensible aux changements de température et les taux de photosynthèse sont également limités par la contrainte hydrique (Cha-Um et Kirdmanee, 2010). Pour limiter les effets délétères de la contrainte hydrique, les eucalyptus mettent en place des mécanismes d'ajustement osmotique (Merchant *et al.*, 2007). Il a été montré que le quercétol, un flavonoïde de type flavonol, est un composé principal rentrant en jeu dans la tolérance de diverses espèces d'*Eucalyptus* à la contrainte hydrique (Merchant *et al.* 2010 ; Merchant *et al.* 2009 ; Arndt *et al.*, 2008). Concernant d'autres composés osmorégulateurs, des résultats contradictoires ont été publiés : chez *E. marginata*, Warren *et al.* (2007) mettent en évidence que les sucres solubles constituent une majeure partie des composés responsables de l'ajustement osmotique par rapport aux acides aminés, en particulier pour la proline, alors que Cha-Um et Kirdmanee (2010) montrent une forte augmentation de proline chez *E. camaldulensis* lors d'un stress hydrique ou salin.

Comparativement, très peu d'études ont été menées au niveau moléculaire. Comparant des séquences connues d'Arabidopsis avec des ESTs d'eucalyptus, Soares-Cavalcanti *et al.* (2009) ont mis en évidence 86 clusters représentant des facteurs de transcription de type MYB impliqués dans la réponse à la sécheresse et/ou au stress salin. De la même façon, dos Santos Barros *et al.* (2009) ont identifié 51 gènes orthologues correspondant à 9 familles multigéniques codant des osmoprotectants, les plus représentés appartenant à la voie de biosynthèse des tréhaloses et des cystéines. Par une approche de génomique fonctionnelle,

Tsuchihira *et al.* (2010) ont mis en évidence une corrélation positive entre tolérance à la sécheresse et le niveau de protéines codant les aquaporines des membranes plasmiques (PIP) chez *E. camaldulensis*. En transformant génétiquement un hybride *E. urophylla* x *E. grandis* par deux gènes de radis *RsPIP1* et *RsPIP2*, ils ont montré que la dérégulation de ces deux gènes avait des effets délétères, tandis que la surexpression de *PIP2* augmentait l'activité photosynthétique et la croissance chez l'eucalyptus. Enfin, certains gènes décrits pour d'autres espèces comme impliqués dans la réponse à la contrainte hydrique ont été mis en évidence chez l'eucalyptus, lors d'une réponse au froid, comme une tocopherol cyclase, *EgSXD1* (El Kayal *et al.*, 2006) et un facteur de transcription CBF (Gamboa *et al.*, 2007).

Etant donné le peu de données moléculaires obtenues pour la réponse du genre *Eucalyptus* à la contrainte hydrique, une vision globale de l'expression du transcriptome semble être indispensable à une meilleure compréhension des mécanismes moléculaires mis en jeu dans ce contexte. Tel était un des objectifs de cette thèse.

4.7 Comment étudier la réponse à la contrainte hydrique?

La réponse à la contrainte hydrique implique des processus complexes et affecte tous les organes de la plante. L'impact de la sécheresse a été particulièrement étudié sur un des paramètres majeurs ayant trait à l'économie : la production. Ainsi, Granier *et al.* (2007) ont montré que sur 11 sites forestiers dispersés dans toute l'Europe, la sécheresse de 2003 a conduit à une chute de la production primaire brute pour 8 sites par rapport à 2002, qui s'est répercutée les années suivantes, avec une intensité variable selon les sites et les espèces (figure 14).

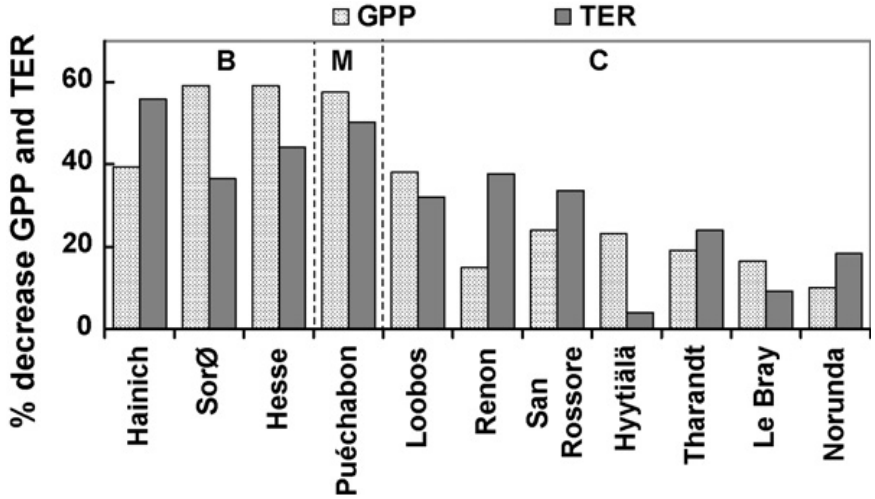


Figure 14 : Les effets de la sécheresse sur la production primaire nette (GPP) et la respiration des écosystèmes (TER) sont variables selon les espèces considérés. Ici sont représentées les diminutions de ces 2 paramètres estimés quotidiennement quand l'eau extractible diminue de 1 à 0,1. Les conifères (C) semblent enregistrer des pertes de productivité plus faibles que les espèces à feuilles caduques, qu'elles soient méditerranéennes (M) ou non (B). Issu de Granier *et al.*, 2007.

D'autre part, la compréhension des mécanismes sous-jacents de la réponse à la contrainte hydrique est cruciale pour pouvoir anticiper et/ou améliorer la réponse des arbres en fonction des variations environnementales. Ainsi, les scientifiques se sont attachés à décrire les mécanismes mis en jeu à différentes échelles. Par exemple, on peut décrire l'impact de l'alimentation hydrique au niveau morphologique par la description des allocations de biomasse entre partie aérienne et racinaire. Au niveau physiologique, les flux d'eau ou d'éléments minéraux dans le continuum sol / plante / atmosphère sont souvent approchés grâce à l'élaboration de modèles (figure 15) utilisant des variables environnementales (données climatiques et pédologiques) et physiologiques (flux de sève, potentiel hydrique foliaire, croissance radiale…).

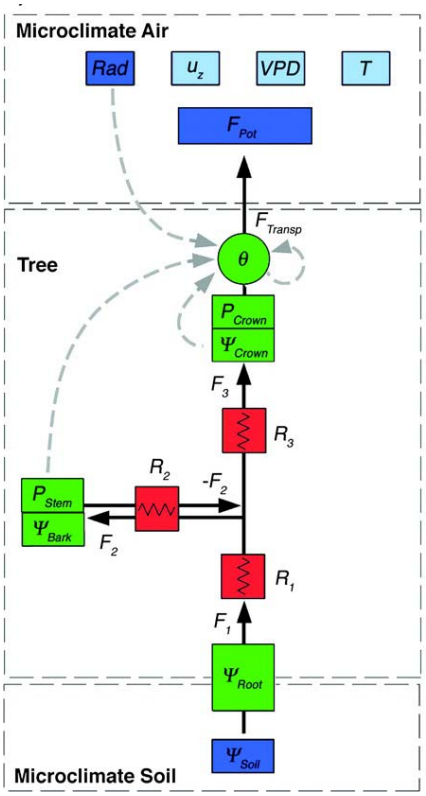


Figure 15 : Exemple de modélisation du continuum sol-plante-atmosphère prenant en compte des flux hydrauliques (F) et le stockage d'eau (P), et donc les potentiels hydrique (Ψ) et les résistance de flux (R) (Zweifel *et al.*, 2007).

Les études physiologiques peuvent être complétées par l'étude de coupes anatomiques de différents tissus. Par exemple, l'étude de l'anatomie des feuilles d'olivier a permis à Chartzoulakis *et al.* (1999) d'émettre l'hypothèse que les différences de rendement photosynthétique entre deux variétés soumises à une contrainte hydrique en l'absence de différences pour leurs paramètres hydrauliques pourraient être liées aux différences de conformation de leur mésophylle. En effet la variété qui présente la plus grande proportion

d'espace intercellulaire, possède également une plus forte conductance gazeuse à travers ces espaces, ce qui favoriserait le rendement photosynthétique.

De nombreuses études de dosage biochimiques ont également contribué à une meilleure compréhension des mécanismes de réponse à la contrainte hydrique. Le dosage biochimique de l'ABA (Julliard *et al.*, 1993), ainsi que de molécules osmorégulatrices telles que la proline (Troll et Lindsley, 1955) ou les sucres (Ashwel, 1957), ou des espèces activées d'oxygène ont permis de mieux comprendre les mécanismes oxydatifs (Bartoli *et al.*, 1998), ou l'osmorégulation ayant lieu au cours de la contrainte hydrique. Ces dosages font appel à différentes techniques, depuis les dosages immunologiques pour l'ABA, jusqu'à la résonnance magnétique nucléaire (RMN) pour les acides aminés et sucres, ou la chromatographie liquide haute pression (HPLC) pour les métabolites secondaires.

Enfin, au niveau moléculaire, les approches protéomiques se sont essentiellement faites grâce à des gels de séparation des protéines en deux dimensions (par la charge et par la taille) et identification par spéctrométrie de masse (Plomion *et al.*, 2006, Bedon *et al.*, 2010). Les analyses du transcriptome étant techniquement plus aisées, elles ont été largement utilisées ces dernières années pour tenter de déchiffrer les acteurs moléculaires de la réponse à la contrainte hydrique (Seki et al., 2007; Deyholos et al., 2009). Deux principales approches ont été utilisées : i/ sans *a priori* par le séquençage de transcrits et l'analyse *in silico* de leur distribution, ou grâce à des puces contenant des milliers de transcrits de fonction connues), ou pas et ii/ avec *a priori* grâce à la PCR quantitative sur des gènes candidats.

Cette thèse s'est focalisée sur la compréhension des mécanismes moléculaires mis en jeu dans la réponse à la contrainte hydrique de l'eucalyptus, notamment grâce à l'utilisation de méthodes de séquençage nouvelle génération (NGS), ce qui a permis d'obtenir une vision large du transcriptome exprimé au cours de la contrainte. La prochaine partie se propose d'introduire ces nouvelles technologies et de présenter leur intérêt dans le cadre d'études de génomique environnementale.

5 Les NGS pour répondre à des questions de recherche en génomique environnementale

5.1 Du séquençage « Sanger » aux NGS

Depuis 1977, les travaux de séquençage étaient largement dominés par l'utilisation d'une méthode de séquençage développée par Sanger et Coulson (Sanger *et al.*, 1977). Cette méthode est basée sur une technique de polymérisation de l'ADN incomplète grâce à l'utilisation de didésoxynucléotides (ddATP, ddCTP, ddGTP ou ddTTP) en mélange avec des désoxyribonucléotides utilisés classiquement en PCR. La terminaison se fait de manière statistique suivant que l'ADN polymérase utilise l'un ou l'autre de ces nucléotides. Il en résulte un mélange de fragments d'ADN de tailles croissantes, qui se terminent tous au niveau du didésoxyribonucléotide incorporé dans la séquence. Ces fragments sont ensuite séparés par électrophorèse, ce qui permet de repérer la position de chaque désoxyribonucléotides dans la séquence. Cette méthode de séquençage, relativement coûteuse et peu rapide a néanmoins permis de réaliser les 1ers séquençages complets de génome, comme celui de l'homme (Lander *et al.*, 2001 ; Venter *et al.*, 2001) et de la plante modèle *Arabidopsis thaliana* (AGI, 2000).

La révolution que représente les NGS, c'est de pouvoir séquencer des millions de fragments d'ADN en parallèle (Margulies *et al.*, 2005), sans étape de clonage nécessaire. Après être coupés mécaniquement et liés à des adaptateurs, les fragments à séquencer sont soit fixés sur des microbilles (454 et SOLiD), soit sur une lame de verre (Illumina) puis amplifiés séparément par PCR. Ensuite, le séquençage à proprement parler consiste à « lire » une base *in situ* sur chacun des fragments, à acquérir l'image correspondante puis à passer à la base suivante.

Les différentes technologies développées diffèrent principalement sur la préparation des échantillons à séquencer et sur la méthode de séquençage utilisée (Metzker, 2010).

La préparation des banques peut être basée sur une amplification clonale d'échantillons provenant de molécules d'ADN unique, ou à partir d'une seule molécule d'ADN.L'amplification clonale peut se faire grâce à une PCR en émulsion (les molécules d'ADN sont fixées sur des billes), ou une amplification en phase solide (fixation sur une lame de verre).

Ces méthodes sont relativement complexes, demandent beaucoup de matériel génomique et nécessitent des PCRs qui peuvent induire des erreurs telles que des mutations. Lorsque l'échantillon est réduit à une seule molécule d'ADN, il suffit d'immobiliser la molécule sur le support, soit grâce à une amorce, soit grâce à la molécule d'ADN elle-même, soit grâce à la polymérase.

Jusqu'alors, quatre méthodes de séquençage ont été développées : l'addition d'un seul nucléotide, la terminaison réversible cyclique, le séquençage en temps réel et le séquençage par ligation.

La méthode d'addition d'un seul nucléotide est utilisée dans la technologie Roche/454. Lors du pyroséquençage, les nucléotides sont ajoutés les uns après les autres. Si le nucléotide ajouté dans le milieu réactionnel correspond à celui attendu par la polymérase, il est incorporé dans le brin en cours d'élongation et libère un pyrophosphate, ce qui, par cascade enzymatique, produit un signal lumineux. Le signal lumineux est capté et reproduit sous forme d'un pic dont la hauteur est fonction du nombre de nucléotides incorporés en même temps.

La terminaison réversible cyclique (développée par Illumina/Solexa) utilise l'incorporation réversible de nucléotides fluorescents suivie de la lecture optique de la fluorescence. Il y a terminaison de synthèse grâce à l'utilisation d'un groupement de protection attaché au nucléotide qui stoppe la synthèse d'ADN. A la suite de l'incorporation, les nucléotides restants sont éliminés, et la prise d'image permet alors de déterminer quel nucléotide a été incorporé. L'élimination du groupement de protection permet la restauration du groupement fonctionnel du nucléotide incorporé, ce qui permet à l'ADN polymérase d'incorporer le prochain nucléotide, et ainsi de suite.

Contrairement à l'utilisation des terminateurs réversibles, le séquençage en temps réel ne fait pas appel à des arrêts du processus de synthèse de l'ADN (utilisé chez Pacific Biosciences). Cette méthode implique la prise d'image continue lors de la synthèse d'ADN grâce à des nucléotides marqués.

Enfin, le séquençage par ligation (technologie utilisé chez SOLiD, Life/APG) fait appel à une ligase pour identifier les nucléotides présents à une position donnée dans une séquence d'ADN. Contrairement aux autres méthodes, cette méthode n'utilise pas d'ADN polymérase pour synthétiser un deuxième brin d'ADN, mais la sensibilité d'inadéquation de la ligase pour déterminer la séquence de la molécule de l'ADN cible.

Au final, la différence entre ces principales technologies réside principalement en des différences du nombre de fragments que l'on peut séquencer et de leur longueur (tableau 1). Par exemple, la technologie 454 qui fait appel à des amplifications par PCR en émulsion de l'ADN cible et au pyroséquençage permet de séquencer de longs fragments (330pb en moyenne), mais pour relativement peu de séquences par run (environ 1M soit 0,45Gb), en comparaison avec la technologie Illumina qui utilise une amplification par PCR en phase solide et un séquençage par terminaison réversible cyclique et génère des lectures plus courtes (75-100pb) pour un grand nombre de séquences (6M de séquences par run soit 35Gb voire 200Gb annoncés avec la technique HiSeq2000).

Tableau 1 : Comparaison des caractéristiques de différentes méthodes de séquençage développées, d'après Metzker, 2010.

Platform	Library/ template preparation	NGS chemistry	Read length (bases)	Run time (days)	Gb per run	Machine cost (US$)	Pros	Cons	Biological applications	Refs
Roche/454's GS FLX Titanium	Frag, MP/ emPCR	PS	330*	0.35	0.45	500,000	Longer reads improve mapping in repetitive regions; fast run times	High reagent cost; high error rates in homo-polymer repeats	Bacterial and insect genome *de novo* assemblies; medium scale (<3 Mb) exome capture; 16S in metagenomics	D. Muzny, pers. comm.
Illumina/ Solexa's GA_{II}	Frag, MP/ solid-phase	RTs	75 or 100	4‡, 9§	18‡, 35§	540,000	Currently the most widely used platform in the field	Low multiplexing capability of samples	Variant discovery by whole-genome resequencing or whole-exome capture; gene discovery in metagenomics	D. Muzny, pers. comm.
Life/APG's SOLiD 3	Frag, MP/ emPCR	Cleavable probe SBL	50	7‡, 14§	30‡, 50§	595,000	Two-base encoding provides inherent error correction	Long run times	Variant discovery by whole-genome resequencing or whole-exome capture; gene discovery in metagenomics	D. Muzny, pers. comm.
Polonator G.007	MP only/ emPCR	Non-cleavable probe SBL	26	5§	12§	170,000	Least expensive platform; open source to adapt alternative NGS chemistries	Users are required to maintain and quality control reagents; shortest NGS read lengths	Bacterial genome resequencing for variant discovery	J. Edwards, pers. comm.
Helicos BioSciences HeliScope	Frag, MP/ single molecule	RTs	32*	8‡	37‡	999,000	Non-bias representation of templates for genome and seq-based applications	High error rates compared with other reversible terminator chemistries	Seq-based methods	91
Pacific Biosciences (target release: 2010)	Frag only/ single molecule	Real-time	964*	N/A	N/A	N/A	Has the greatest potential for reads exceeding 1 kb	Highest error rates compared with other NGS chemistries	Full-length transcriptome sequencing; complements other resequencing efforts in discovering large structural variants and haplotype blocks	S. Turner, pers. comm.

*Average read-lengths. ‡Fragment run. §Mate-pair run. Frag, fragment; GA, Genome Analyzer; GS, Genome Sequencer; MP, mate-pair; N/A, not available; NGS, next-generation sequencing; PS, pyrosequencing; RT, reversible terminator; SBL, sequencing by ligation; SOLiD, support oligonucleotide ligation detection.

5.2 De nombreuses possibilités d'utilisation

L'utilisation d'une technologie plutôt qu'une autre est souvent dictée par l'utilisation que l'on veut en faire. Les possibilités d'application des NGS sont diverses et peuvent intervenir dans de larges champs disciplinaires. Nous listerons ici les utilisations majeures qui en ont été faites.

5.2.1 Séquençage *de novo*

Les lectures obtenues par séquençage peuvent être assemblées les unes avec les autres afin de reconstruire des génomes. Ainsi, la première utilisation de NGS a permis de séquencer le génome bactérien de *Mycoplasma genitalium* (Margulies *et al.*, 2005). Au début, l'utilisation de ces technologies était réservée à des petits génomes tels que ceux des bactéries (Chaisson et Pevzner, 2008 ; Hernandez *et al.*, 2008). En effet, un des défis majeurs pour séquencer des génomes de grande taille est de trouver des zones de recouvrement spécifiques entre lectures relativement courtes, sachant qu'un grand nombre de séquences (voire des régions entières) sont répétées dans les génomes (Pop et Salzberg, 2008). Si un des éléments pour améliorer le recouvrement consiste en l'allongement de la longueur des séquences (Chaisson *et al.*, 2004), les séquences les plus longues obtenues avec les NGS sont encore trop courtes pour parer complètement à ce problème. L'obtention de paires de lectures (la distance entre deux lectures est connue) a permis également de pouvoir ordonner les séquences obtenues et ainsi améliorer l'assemblage (Fullwood *et al.*, 2009).

Deux stratégies de séquençage ont alors été développées :

- Le séquençage aléatoire global qui consiste à obtenir des séquences appariées à différentes échelles génomiques. On procède alors au séquençage des extrémités de grands fragments génomiques de plusieurs centaines de milliers de nucléotides, qui permettent d'assembler les contigs en vastes ossatures (scaffolds) et de détecter certaines erreurs à l'intérieur des contigs.

- Le séquençage hiérarchique qui consiste à assembler des lectures chevauchantes à l'échelle de grands fragments génomiques préalablement ordonnés en une carte physique. Ces grands fragments sont alors insérés dans des vecteurs (le plus souvent des chromosomes bactériens artificiels, BAC), puis séquencés aléatoirement. Cette stratégie demande un effort de cartographie physique supplémentaire, mais permet un assemblage plus aisé.

L'utilisation de ces différentes stratégies a ainsi conduit à la publication croissante de séquençages *de novo* de nombreux génomes (tableau 2) : chez les animaux, celui du panda

géant (Li *et al.*, 2010) et de la dinde domestique (Dalloul *et al.*, 2010) ; chez les plantes, celui de la pomme (Velasco *et al.*, 2010) du cacao (Argout et al., 2010) et du concombre (Huang *et al.*, 2009) par exemple, qui ont été partiellement séquencés grâce aux NGS. Actuellement le génome de l'eucalyptus est également en cours de séquencçage à l'aide des NGS.

Tableau 2: Projets d'assemblage de novo réalisés avec les NGS, revues par Shatz *et al.*, 2010

		Input sequence						Assembly								
								Contigs				Scaffolds				
Organism/genome size	Assembler/status[a]	Type	Pair size	Avgerage read (bp)	No. of reads	Read coverage[b]	Pair coverage[c]	No.	N50	Max	Total	No.	N50	Max	Total	Notes
Human (*H. sapiens*)/3.0 Gb	ABYSS published 2009	GA	210 bp	35–46	3.5 B	45×	120×	2.76 M	1.5 kb	18.8 kb	2.18 Gb	NR	NR	NR	NR	
Grapevine (*V. vinifera*)/500 Mb	Myriad published 2007	Sanger	2–10 kb	579	5.95 M	6.9×	21×	58,611	18.2 kb	238 kb	531 Mb[d]	2093	1.33 Mb	7.8 Mb	421 Mb[d]	
		Sanger	40 kb	460	144 k	0.13×	4.4×									
		Sanger	120 kb	369	68 k	0.02×	4.2×									
		454	None	169	12.5 M	4.2×	—									
Cucumber (*C. sativus*)/367 Mb	RePS2 published 2009	Sanger	2–6 kb	439	2.08 M	3.35×	9.9×	62,412	19,807	NR	226 Mb	47,837	1.15 Mb	NR	244 Mb	
		Sanger	40 kb	496	339 K	0.46×	16.7×									
		Sanger	140 kb	551	33.2 k	0.04×	5.6×									
		GA	200 bp	42	282 M	32.5×	76.8×	NR	2.6 kb	NR	204 Mb	NR	19 kb	NR	238 Mb	Assembly of only Sanger reads
		GA	400 bp	44	173 M	20.6×	94.4×									
		GA	2 kb	53	105 M	15.3×	286×	NR	12.5 kb	NR	190 Mb	NR	172 kb	NR	200 Mb	Assembly of only GA reads
Panda (*A. melanoleura*)/2.4 Gb	SOAPdenovo published 2010	GA	150	45	1.31 B	24.5×	43.3×	200,604	36,728	434,635	2.25 Gb	81,469[e]	1.22 Mb	6.05 Mb	2.30 Gb	
		GA	500	67	917 M	25.5×	90.2×									
		GA	2 kb	71	397 M	11.8×	192×									
		GA	5 kb	38	505 M	8.0×	533×									
		GA	10 kb	35	254 M	3.7×	571×									
Strawberry (*F. vesca*)/220 Mb	CABOG and Velvet announced	454	None	209	7.73 M	7.3×	—	16,487	28,072	215,349	202 Mb	3263	1.44 Mb	4.1 Mb	214 Mb	
		454	None	368	787 M	13.2×	—									
		454	2.5 kb	193	2.39 M	2.1×	6.9×									
		454	20 kb	236	1.58 M	1.7×	20×									
		GA	None	76	36 M	12.4×	—									
		SOLiD	2 kb	25	1.30 M	0.14×	6.4×									
Turkey (*M. gallopavo*)/1.1 Gb	CABOG announced	454	3 kb	180	6 M	1×	8×	128,271	12,594	90 kb	931 Mb	26,917	1.5 Mb	9 Mb	NR	
		454	20 kb	195	2 M	0.3×	18×									
		454	None	366	13 M	4×	—									
		GA	180 bp	74	200 M	13×	16×									
		GA	None	74	200 M	13×	—									

Results from de novo assembly of genomes by second-generation sequencing platforms. Summary of inputs and assembly results of recent genome assemblies using SGS reads.
[a]Status indicates when the assembly was published; "announced" assemblies have been described publicly but not yet published.
[b]The number of estimated genome size units contained in the sum of read lengths.
[c]The same value for the sum of lengths of fragments from which paired reads were sequenced.
[d]Contig total greater than scaffold total is largely attributable to "single haplotype contigs."
[e]Number of scaffolds includes single-contig scaffolds. There were 5201 multicontig scaffolds.
GA, Illumina Genome Analyzer; SOLiD, Applied Biosystems SOLiD System.

5.2.2 Reséquençage

Pour des génomes de faible taille, il est possible de reséquencer le génome entier pour différents génotypes afin d'analyser la variation génétique, et l'associer à la variatbilité phénotypique. Cette méthode est fréquemment utilisée chez les organismes procaryotes pour mettre en évidence des mutations entre des lignées pathogènes et non pathogènes, et ainsi identifier éventuellement des sites de résistances aux antibiotiques ou des sites liés à la pathogénicité. Par exemple, Andries *et al.* (2005) ont pu montrer, grâce à l'analyse comparative des séquences génomiques de lignées résistantes et sensibles à un antibiotique, que l'action de cet antibiotique sur *Mycobacterium tuberculosis* ciblait des ATP synthases.

Pour des génomes de plus grande taille, le reséquençage a été fréquemment utilisé pour rechercher des mutations dans des régions ciblées, sur de grandes populations. Par exemple, Dahl *et al.* (2007) ont mis au point une méthode pour pouvoir amplifier et séquencer 10 gènes liés au cancer colo-rectal sur 8 lignées cellulaires (6 lignées de cellules atteintes du cancer

colorectal, 1 lignées atteintes du cancer du sein, et 1 lignée témoin). Ils ont ainsi pu mettre en évidence 200 positions qui présentaient du polymorphisme entre les lignées étudiées, soit des substitutions d'une seule base, soit des insertions délétions. Lorsque l'on dispose déjà d'un génome de référence, comme c'est le cas pour les espèces modèles comme, entre autres, l'Homme (Lander *et al.*, 2001 ; Venter *et al.*, 2001), la drosophile (Adams *et al.* 2000), l'arabette (The Arabidopsis genome initiative, 2000), ou le peuplier (Tuskan *et al.*, 2006), la puissance des NGS permet de multiplier les échantillons séquencés. C'est ainsi que sont nés les projets Hap Map pour le génome humain (http://hapmap.ncbi.nlm.nih.gov/) et 1001 Genomes pour Arabidopsis (http://www.1001genomes.org/), qui se proposent de mettre en commun les efforts de reséquençage de différentes équipes pour identifier et cataloguer les mutations entre génomes de milliers de génotypes différents. Cela devrait permettre d'accélérer nettement la recherche de mutations et l'étude de leurs effets sur le phénotype.

5.2.3 Séquençage du transcriptome

5.2.3.1 Analyses d'abondance de transcrits

Le séquençage de transcrits peut être utilisé pour quantifier l'abondance des ARN messagers, en comptant le nombre de lectures trouvées dans les banques constituant tel ou tel contig. Cette technique est appelée RNA-Seq, et remplace peu à peu l'analyse d'expression de gènes en série (SAGE) développée par Velculescu *et al.* (1995), basée également sur le séquençage de fragments d'ADN, mais après restriction pour obtenir de courts fragments (13-17pb). La méthode RNA-Seq est non seulement plus aisée à mettre en place et moins chère, mais permet également d'atteindre les ARN faiblement exprimés, alors que la méthode SAGE se cantonnait aux ARN les plus exprimés (Wang, 2009). Le tableau 3 présente les différentes études publiées utilisant cette méthode.

Tableau 3 : Etudes récentes réalisées par RNA-Seq sur différentes espèces de plantes et de pathogènes.

Espèces étudiées	Contexte de l'étude	Plateforme	Référence
Concombre	processus de développement	454	*Guo et al., 2010*
Olivier	qualité du fruit	454	*Alagna et al., 2009*
Raisin	processus de développement	454	*Bellin et al., 2009*
Tomate	recherche de métabolites secondaires	454	*Schilmiller et al., 2010*
Ginseng	recherche de métabolites secondaires	454	*Sun et al., 2010*
Châtaigner américain et chinois	réponse aux stress abiotiques	454	*Barakat et al., 2009*
Burkholderia cenocepacia	réponse à l'environnement	Illumina	*Yoder-Himes et al., 2009*
Artémise	recherche de métabolites secondaires	454	*Wang et al., 2009*
Espèces de mangroves: *Rhizophora mangle* et *Heritiera littoralis*	adaptation à l'environnement	454	*Dassanayake et al., 2009*
Riz	atlas du transcriptome dans différents organes	Illumina	*Zhang et al., 2009*
Raisin	processus de développement	Illumina	*Zenoni et al., 2010*
Chlamydomonas	réponse à l'environnement	Illumina	*Gonzalez-Ballester et al., 2010*
Nicotiana attenuata	réponse aux stress abiotiques	454	*Guilardoni et al., 2010*
Aspergillus oryzae	réponse à l'environnement	Illumina	*Wang et al., 2010*
Maïs	impact de la sélection	454	*Vega-Arreguin, et al., 2009*

Différents protocoles de préparation des librairies ont été publiés, le plus simple et le plus utilisé consistant en la génération de banques d'EST non normalisées (Weber *et al.*, 2007 ; Shin *et al.* 2008 ; Barakat *et al.*, 2009). Nielsen *et al.* (2006) ont également publié une extension de la méthode SAGE avec séquençage des fragments par pyroséquençage 454, qui permet la détection d'ARN rare et d'améliorer l'analyse d'expression avec une plus grande profondeur de séquençage. Enfin Eveland *et al.* (2008) ont développé une méthode permettant de séquencer spécifiquement les régions 3' UTR des gènes, ce qui permet de mieux distinguer des transcrits très proches, tels que des paralogues ou pour étudier l'expression spécifique d'allèles.

Une des difficultés d'utilisation de cette méthode réside dans l'analyse statistique des résultats. En effet, pour pouvoir mettre en évidence les effets que l'on veut tester (par exemple entre deux conditions environnementales), il faut pouvoir s'affranchir des variations individuelles (liés aux répétitions), des variations liées à la construction des banques à séquencer, et liées au séquençage lui-même. Ainsi, les données de comptage ont des caractéristiques particulières qui rendent délicate l'élaboration d'un modèle statistique: ces données n'ont pas une distribution normale mais une distribution binomiale pouvant être approximée par une loi de Poisson (Jiang *et al.*, 2009), il est donc nécessaire d'appliquer des transformations de variables pour pouvoir utiliser des tests paramétriques. La capacité de détection de gènes différentiellement exprimés est directement liée à la longueur du transcrit (Oshlack et Wakefield, 2009), il faut donc normaliser les comptages par la longueur du transcrit. Les variations entre les répétitions techniques peuvent être expliquées par l'échantillonnage aléatoire, ou par la présence d'un bruit de fond tel que des concentrations différentes entre répétitions. Il faut donc tester si les variations entre répétitions sont bien liées à l'échantillonnage aléatoire (Marioni *et al.*, 2008). Enfin, les données digitales d'expression sont obtenues sur des milliers de gènes en parallèle, et des tests statistiques sont donc effectués sur des milliers de variables à la suite, l'hypothèse nulle risque alors d'être rejetée à tort de façon importante. Il faut donc prendre en compte le taux de fausses découvertes possibles (FDR). Ainsi, pour prendre en compte ces particularités, diverses méthodologies ont été développées, des plus simples tests statistiques comme le test exact de Fischer, le test du Chi-deux (Romualdi, 2001), des tests utilisés sur les banques d'ESTs obtenues à partir de séquençage Sanger (Audic et Claverie, 1997), ou des tests adaptés aux NGS depuis l'analyse de microarrays (package DEGseq ; Wang *et al.*, 2010).

La puissance de détection de transcrits différentiellement exprimés de la méthode RNA-Seq a engendré un grand engouement pour l'analyse d'expression sur différentes espèces et pour différentes questions de recherches (tableau 3), que ce soit pour étudier la réponse à des stress biotiques (Barakat *et al.*, 2009 ; Gilardoni *et al.*, 2010), abiotiques (Yoder-Himes, 2009) voire des processus de développement (Torres *et al.*, 2009).

5.2.3.2 Autres utilisations du séquençage de transcriptome

Le séquençage de transcrits est la base de la génomique fonctionnelle et peut avoir d'autres objectifs que l'analyse de l'abondance de transcrits: on peut y rechercher des mutations, observer d'éventuels réarrangements (épissage alternatif) ou étudier les ARN courts non-codants.

Séquencer les ADN complémentaires (ADNc) issus des ARN plutôt que l'ADN génomique (ADNg) permet de réduire considérablement la taille des séquences cibles. Ainsi, le séquençage de transcriptome est souvent préféré pour la recherche de mutations. Novaes *et al.* (2008) ont ainsi mis en évidence 23 742 polymorphismes d'un seul nucléotide (SNP) à partir de 21 génotypes d'*Eucalyptus grandis* grâce au séquençage du transcriptome exprimé dans un mélange de tissus. Les NGS semblent être particulièrement adaptées pour caractériser les transcriptomes d'espèces non modèles (Vera *et al.*, 2008), car elles permettent de couvrir une large partie de leur génome pour un coût très inférieur au séquençage Sanger.

La profondeur de séquençage obtenue avec les NGS permet également de détecter différentes formes d'épissage alternatif. L'épissage alternatif est un processus par lequel les exons des gènes sont différemment recombinés au niveau de l'ARN messager pour produire des transcrits différents, permettant à un seul gène de coder pour de multiples produits protéiques. Ce phénomène est très fréquent, par exemple, il est estimé que parmi les 32 000 gènes du génome humain, 35 à 59% des gènes présentent des formes d'épissage alternatif (Modrek et Lee, 2002). L'identification de ces différentes formes est fondamentale pour la compréhension de mécanismes fonctionnels.

Enfin, une autre application des technologies NGS pour l'analyse des transcriptomes est la découverte et la quantification des petits ARN non codants (ncRNA ; Morozova *et al.*, 2008 ; figure 16). Les ARNnc sont des molécules d'ARN qui ne sont pas traduites en protéines. Cette classe comprend les ARN de transfert (ARNt), les ARN ribosomaux (ARNr), les petits ARN nucléaires et nucléolaires, les microARN et les petits ARN interférents (ARNsi et ARNmi).

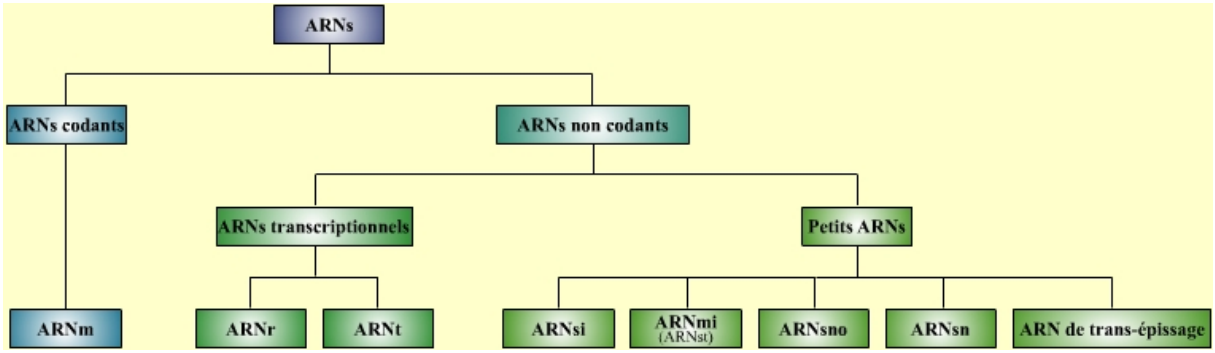


Figure 16 : Les différents types d'ARN trouvés dans les cellules végétales.

5.2.4 Métagénomique

Une des applications les plus prometteuses des technologies NGS est le séquençage d'échantillons issus d'environnement particulier, aussi connu sous le nom métagénomique. Dans ces projets, l'ADN est purifié à partir d'un environnement comme le sol, l'eau ou une partie du corps humain, et le mélange des espèces qui y sont présentes est séquencé de façon aléatoire ou ciblé, comme par exemple l'ARN 16S. Les lectures issues du séquençage peuvent provenir de plusieurs centaines, voire de milliers d'espèces différentes, présentant un défi beaucoup plus grand que l'assemblage d'un projet de séquençage d'un génome unique. Ces études permettent non seulement de faire un inventaire taxonomique mais aussi d'évaluer les processus métaboliques des micro-organismes qui vivent dans les environnements étudiés (Nowrousian, 2010).

5.2.5 Epigénétique

L'épigénétique est l'étude de la régulation de l'expression des gènes qui n'implique pas l'ADN génomique. Les deux principaux types de modifications épigénétiques régulant l'expression génique sont la méthylation de l'ADN par la modification covalente de la cytosine-5' et la modification post-traductionnelle des histones (Callinan et Feinberg, 2006). Ces modifications ont la particularité d'être héritables d'une génération de cellule à l'autre au cours de la mitose voire sur plusieurs générations d'organismes au cours de la méiose, même si leur cause a disparu.

Les modifications post-traductionnelles covalentes des queues d'histones peuvent être dues à des phénomènes de méthylation, d'acétylation, de phosphorylation, et d'ADP-ribosylation. Ces modifications pourraient contrôler l'expression des gènes par le contrôle de la force des interactions ADN-histones qui déterminent l'accessibilité de l'ADN aux facteurs de transcription (Schones et Zhao, 2008). Historiquement, les modifications d'histones étaient étudiées par précipitation de la chromatine (figure 17): après liaison des complexes ADN-

protéines, la protéine d'intérêt était alors précipitée grâce à un anticorps spécifique, puis l'ADN sur lequel elle était liée était caractérisé soit par PCR, qPCR, microrray ou séquençage Sanger couplé à la technique SAGE (Roh *et al.*, 2004). Le remplacement du séquençage Sanger par les NGS (méthode appelée ChIP-Seq) a permis de s'affranchir de l'étape de restriction, et a ainsi permis d'obtenir une meilleure résolution que les autres méthodes développées précédemment (Schones et Zhao, 2008).

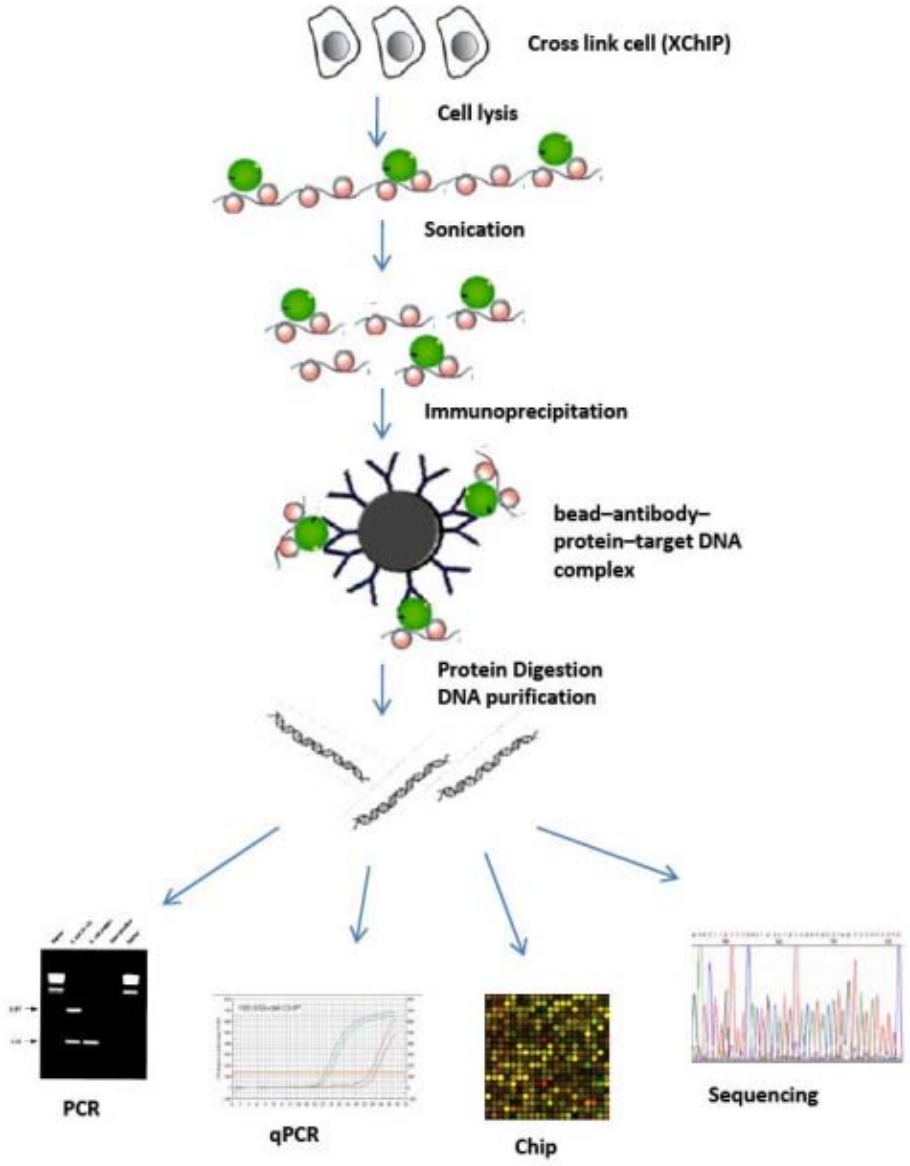


Figure 17 : L'immuno-précipitation de chromatine nécessite la liaison des protéines cellulaires à l'ADN par un traitement au formaldéhyde. L'ADN est ensuite clivé en fragments de 500 à 1000 pb par sonication. Ensuite, la protéine d'intérêt est immunoprécipitée, l'ADN sur lequel elle était liée récupéré et peut être amplifié par PCR, puis quantifiée par qPCR ou microarray, ou séquencée.

Pour étudier le niveau de méthylation des ADN par séquençage, on utilise le séquençage bisulfite qui est basé sur la propriété chimique de bisulfite d'induire la transformation des résidus cytosine en uracile tout en laissant les cytosines méthylées intactes (figure 18 ; Taylor *et al.*, 2007). Ainsi, le séquençage de l'ADN traité au bisulfite permettra de connaître la position des cytosines méthylées (ces positions seront restées cytosines après le traitement).

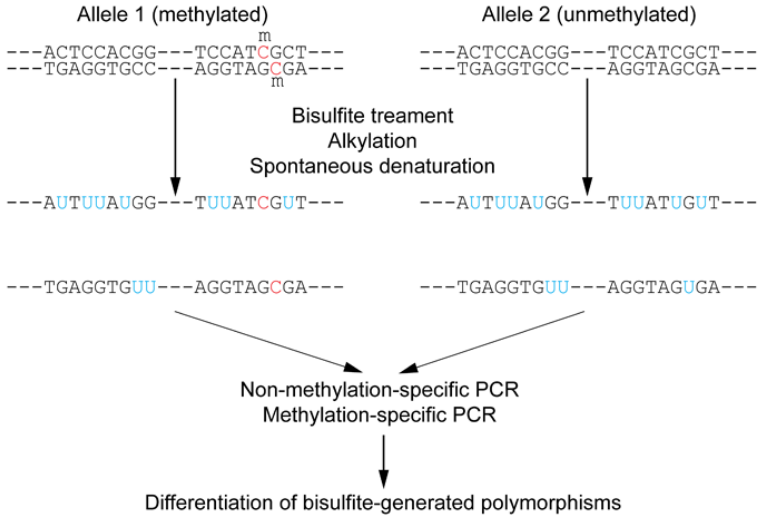


Figure 18 : Le séquençage bisulfite est basé sur la dénaturation en uracyl des cytosines qui ne sont pas méthylées par un traitement au bisulfite. Si l'ADN traité par bisulfite est amplifié puis séquencé, toutes les bases C non méthylées présentes dans la séquence originale (qui ont été transformée en U) seront remplacées par des T dans le produit d'amplification alors que les C méthylés resteront C. En comparant les séquences obtenus avec et sans traitement au bisulfite, on peut révéler le profil de méthylation des séquences étudiées.

5.3 L'émergence de la génomique environnementale favorisée par les NGS

La génomique environnementale est l'application de techniques moléculaires au service de l'écologie et des sciences environnementales (figure 19). Cette nouvelle discipline vise à comprendre les mécanismes de l'adaptation et de la variation phénotypique en utilisant des techniques de génomique (Ouborg et Vriezen, 2007). La réduction des coûts de séquençage et la large gamme d'utilisations possibles des NGS constituent une révolution majeure pour les études concernant des problématiques environnementales. D'une façon générale, la révolution majeure apportée par les NGS est de pouvoir accéder à des données génomiques pour des espèces non modèles, à un coût relativement faible.

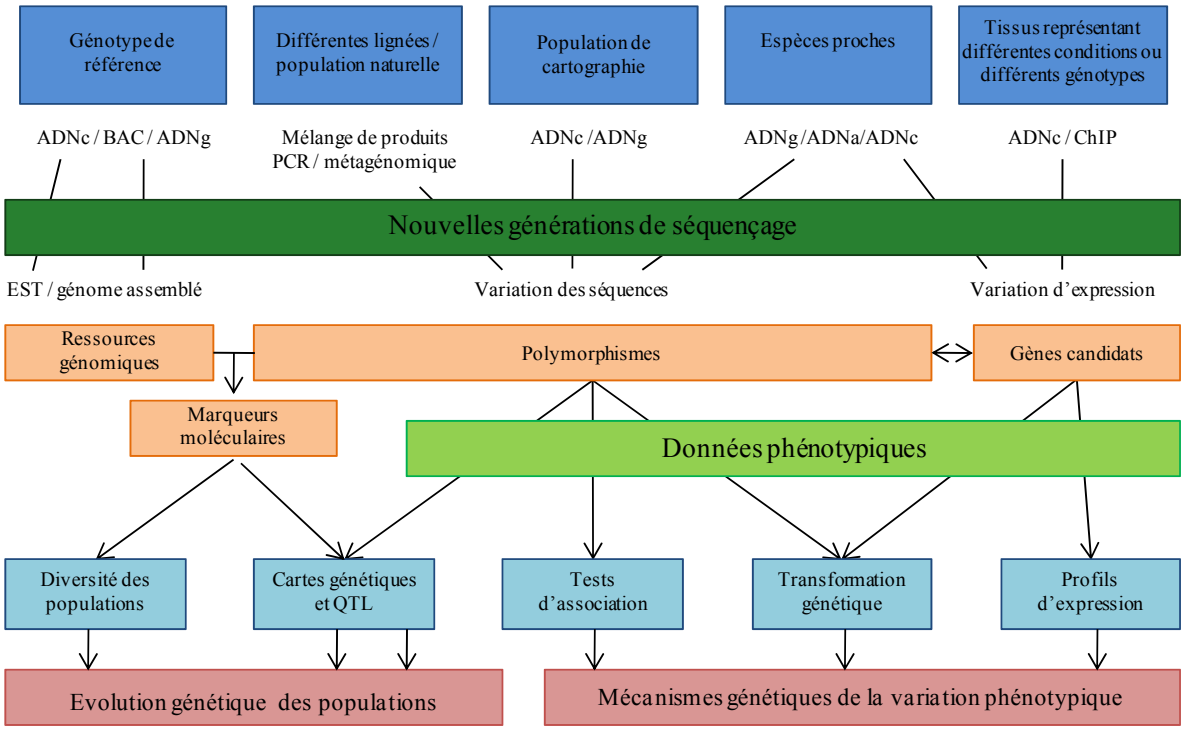


Figure 19 : Vue d'ensemble des applications possibles des NGS pour la génomique environnementale. Adapté de Varnshney *et al.*, 2009.

5.3.1 Les NGS pour l'étude de la biologie de l'évolution

Comme nous l'avons vu précédemment, les changements climatiques risquent d'influencer grandement la répartition des espèces. Si les modèles bioclimatiques peuvent être utilisés pour prédire les risques d'extinction et de migration des espèces, améliorer les connaissances sur l'histoire de la distribution des espèces et du climat est nécessaire pour pouvoir tester les modèles avant de les appliquer à des scénarios climatiques futurs (Guisan et Thuillier, 2005). L'histoire évolutive des populations est retracée essentiellement grâce à l'analyse de polymorphismes entre individus : elle peut être étudiée sur des espèces disparues en utilisant de l'ADN ancien (paléogénétique), ou sur des populations contemporaines. Dans ce cas, on

dispose généralement de grands effectifs pour pouvoir estimer les fréquences allèliques nécessaires à la caractérisation de la diversité nucléotidique des populations, et particulièrement des forces évolutives qui l'influencent.

5.3.1.1 La paléogénétique

L'analyse de l'ADN ancien permet de retracer l'histoire évolutive des populations sur de grandes périodes. Néanmoins, les ADN anciens ont des caractéristiques particulières qui rendent difficile la mise en place de ce type d'étude : on les trouve en faible quantité, ils sont souvent fragmentés, modifiés par des processus oxydatifs ou hydrolytique, et sont souvent présents en mélange de plusieurs espèces dans un échantillon donné. Les NGS sont particulièrement adaptées à l'étude de l'ADN ancien car elles permettent de séquencer des fragments relativement courts (de l'ordre de 250-500pb), une taille comparable à celles trouvée pour la plupart des génomes anciens dégradés (Millar *et al.*, 2008). Par leurs propriétés, les NGS devraient permettre un nouvel essor de ce type d'étude, que ce soit par analyse métagénomique, ou spécifique de certaines espèces.

Pour les études de métagénomique, l'identification et la quantification des différentes espèces présentes dans des échantillons représentant des environnements du passé sont possibles grâce à l'absence de clonage et à la grande profondeur de séquençage des NGS, ce qui permet d'étudier les changements subtils des dynamiques de population au cours du temps (Millar *et al.*, 2008). Ainsi, Sonstebo *et al.* (2010) ont utilisé le pyroséquençage 454 pour reconstruire la flore arctique de la dernière période glaciaire. Pour cela, ils ont comparé le mélange d'ADN contenu dans des échantillons de pergélisol avec une base de références issue de 850 espèces représentatives de la diversité de la flore locale. Parmi les 47 unités taxonomiques découvertes dans un échantillon de pergélisol, 20 ont pu être identifiées au niveau de l'espèce.

L'ADN ancien peut également être utilisé pour reconstruire la phylogéographie d'une unité taxonomique, même si à l'heure actuelle les études réalisées ne remontent qu'à quelques centaines d'années grâce aux échantillons détenus dans les herbiers, les études publiées à partir d'échantillons plus âgés étant soumises à controverse (Gugerli *et al.*, 2004).

A l'échelle de l'espèce, les analyses à partir d'ADN ancien sur les ancêtres de plantes cultivées permettent de retracer l'origine et l'étendue de l'agriculture, et l'impact de la domestication sur la diversité génétique et l'évolution des familles de gènes (Gugerli *et al.*, 2004). Dans ce cas, l'existence d'un génome de référence d'une espèce proche permet de

simplifier grandement l'assemblage et d'utiliser des technologies avec des lectures courtes (Millar *et al.*, 2008).

5.3.1.2 Génétique des populations-étude de marqueurs neutres

La génétique des populations permet non seulement de retracer l'évolution de l'espèce étudiée, mais également d'estimer son potentiel adaptatif aux changements environnementaux en cours, en fonction de leur niveau de diversité. Cette discipline se propose d'étudier la distribution et les changements de fréquence d'allèles dans les populations sous l'influence de pressions évolutives.

Ces études sont évidemment conditionnées par la mise en évidence de polymorphismes génétiques. Pour déterminer les fréquences allèliques à l'intérieur des populations, la profondeur de séquençage et la possibilité de mélanger des individus des NGS ont permis d'augmenter considérablement le nombre de données disponibles. S'il est encore difficile de comparer des génomes entiers, les études sont souvent restreintes à une fraction du génome moins large, tel que les EST (Novaes *et al.*, 2008), les ADN méthylés (Deschamps *et al.*, 2008), ou des gènes d'intérêt (Kulheim *et al.*, 2009). Néanmoins, la mise en place des projets de type Hap Map (http://hapmap.ncbi.nlm.nih.gov/) permet de penser que ce type de reséquençage va pouvoir s'étendre à l'ensemble du génome chez de nombreuses espèces.

Le séquençage peut également être couplé à d'autres techniques. Par exemple, van Orsouw *et al.* (2007) ont proposé une technique couplant AFLP et séquençage 454 pour réduire la complexité des séquences polymorphes, et permettant ainsi de comparer 2 ou plusieurs lignées de maïs. De plus, les polymorphismes mis en évidence grâce aux NGS peuvent être étudiés à l'aide de méthodes de génotypage à haut débit pour augmenter les effectifs génotypés (revue par Ragoussis, 2009).

La puissance de détection de polymorphismes des NGS a permit de grandes avancées dans l'étude de la conservation de la biodiversité (Ouborg *et al.*, 2010). En effet, la caractérisation des taux de mutation, de recombinaison, des flux géniques, de la dérive génétique permet une meilleure connaissance des niveaux de diversité des populations et donc de repérer les populations menacées, et de mettre en place des programmes de conservation du patrimoine génétique. L'analyse de variations neutres a largement été utilisée pour caractériser la diversité génétique d'espèces rares ou en danger. Néanmoins, la corrélation entre variations génétiques neutres et fitness des individus ou des populations menacées est soumis à controverse (Kohn *et al.*, 2006). Voilà pourquoi une des tâches majeures à laquelle s'attellent

les généticiens des populations est d'essayer d'identifier des locus dont la fréquence allèlique s'écarte des valeurs du modèle neutre et d'équilibre démographique, permettant ainsi de mettre en évidence des signatures de la sélection naturelle. Ces locus peuvent être identifiés de 3 façons : par une approche gène-candidat (étude du polymorphisme de gènes sur lesquels on a un *a priori*), par la détection de balayage sélectif (perte de polymorphisme pour des marqueurs voisins d'un site soumis à sélection directionnelle ayant subi un événement de fixation d'un allèle favorable), ou par la cartographie de déséquilibre de liaison (Kohn *et al.*, 2006). Si la détection de tels locus permet de mieux comprendre l'évolution de l'espèce ou de la population étudiée, cela permet également de mettre en évidence des régions du génome impliquées dans l'adaptation donc potentiellement dans la variation phénotypique de caractères adaptatifs.

5.3.2 Les mécanismes génétiques sous-jacent à la variation phénotypique

En plus de comprendre les mécanismes d'évolution des espèces pour anticiper les effets de leur environnement futur, les NGS peuvent également contribuer à caractériser les mécanismes génétiques mis en jeu dans la variation phénotypique intervenant dans la relation plante-environnement.

L'association de données phénotypiques aux données génotypiques en population peut permettre de mettre en évidence des régions génomiques associées à la variation de caractères d'intérêt. Dans le cas de pédigrées de cartographie génétique, il s'agit par exemple d'identifier de larges régions génomiques par la cartographie de QTL. Dans le cas de grandes populations d'individus faiblement apparentés, les études d'association permettent d'identifier jusqu'à des polymorphismes causaux de la variation des caractères. Comme nous l'avons vu précédemment, la grande capacité des NGS à détecter des polymorphismes en quantité importante bénéficie également à ce type d'approche (Stapley *et al.*, 2010). En effet, les NGS laissent entrevoir la possibilité de détecter ou génotyper des milliers de marqueurs SNP chez un grand nombre d'individus et surtout à l'échelle d'un génome entier. Ainsi, il devrait être bientôt possible de réaliser des tests d'association chez des espèces non modèles sans construire au préalable des cartes de liaison (Slate *et al.*, 2009). En 2010, grâce aux NGS, Huang *et al.* (2010) ont pu séquencer à grande échelle 517 variétés de riz. L'étude de la variabilité moléculaire en présence à permis aux auteurs de détecter un total de 3,6 millions de SNP qu'ils ont utilisés pour construire une carte haplotypique à haute densité. Cette analyse leur a permis d'identifier des régions génomiques associés à 14 caractères d'intérêt

agronomique, parmi lesquels 6 locus étroitement liés à des gènes précédemment identifiés par des méthodes lourdes de clonage positionnel.

Une autre façon de détecter des locus impliqués dans l'interaction avec l'environnement est de considérer non plus les variations nucléotidiques, mais la régulation de l'expression des gènes (Stapley *et al.*, 2010). Les NGS permettent non seulement d'établir une comparaison du catalogue des gènes exprimés entre différentes conditions, mais également d'identifier les différences génétiques de régulation de leur expression. La technologie RNA-Seq constitue un atout de taille pour l'analyse de la plasticité moléculaire, en permettant de dresser les normes de réaction de milliers de transcrits simultanément (Aubin-Horth et Renn, 2009) et de son contrôle génétique. De plus, l'analyse de séquences permet de déterminer les éléments régulateurs ou les sites d'épissage alternatif présents dans le génome, et les différentes méthodologies développées en relation avec les NGS permettent de caractériser l'action d'ARN non codants ou le contrôle épigénétique de l'expression. Avec l'avènement des NGS, la quantité de données ne va pas cesser de croître, et la compréhension de systèmes complexes sera conditionnée par la capacité d'intégration des différentes visions du génome disponibles (Hawkins *et al.*, 2010).

5.4 L'eucalyptus, un genre modèle pour la génomique environnementale

L'eucalyptus est un arbre tropical à croissance rapide et continue dont le bouturage est maîtrisé (Saya *et al.*, 2008), ce qui permet la mise en place aisée de dispositif expérimentaux en condition naturelle avec l'acquisition rapide de données pour une espèce pérenne.

La taille estimée du génome d'*Eucalyptus* est d'environ 600 Mpb (Grattapaglia et Bradshaw, 1994). C'est un genre diploïde (aucune preuve de polyploïdisation récente) avec un nombre de chromosomes égale à 11 (Eldridge *et al.*, 1993 ; Potts et Wiltshire, 1997). Grâce à son système de reproduction préférentiellement allogame et sa domestication récente, le genre *Eucalyptus* affiche un très haut niveau de diversité nucléotidique et des niveaux de déséquilibre de liaison très faibles (Grattapaglia et Kirst, 2008).

Plusieurs cartes génétiques avec des centaines de microsatellites publics et de nombreux gènes connus ont été publiés (Gion *et al.*, 2000 ; Thamarus *et al.* 2002 ; Marques *et al.*, 2002 ; Brondani *et al.*, 2006). Une forte synténie entre les génomes des différentes espèces a été mise en évidence (Myburg *et al.*, 2007), ce qui permet une exportation aisée de marqueurs génétiques d'une espèce à l'autre. De nombreux QTL sont connus pour des caractères liés à la qualité du bois ainsi que la propagation végétative, la tolérance au froid, et le contrôle de la

floraison (revue dans Poke *et al.*, 2005; Myburg *et al.*, 2007 ; Freeman *et al.*, 2009 ; Thumma *et al.*, 2010).

Des bases de données de séquences génomiques exprimées et les expériences d'expression microarray ont été produites par des groupes privés (Grattapaglia et Kirst, 2008 ; Vicentini *et al.*, 2005) et des institutions publiques (Kirst *et al.*, 2004; Paux *et al.*, 2004; Foucart *et al.*, 2006). A l'heure actuelle, 37853 EST d'*Eucalyptus* sont disponibles dans Genbank (figure 20) représentant majoritairement seulement quelques espèces du sous genre *Symphyomyrtus* : *E. camaldulensis*, *E. grandis*, *E. gunnii*, *E. globulus* et *E. tereticornis*. Elles ont principalement été développées sur les propriétés du bois (Paux *et al.* 2004) ou sur la résistance au froid pour les espèces tempérées.

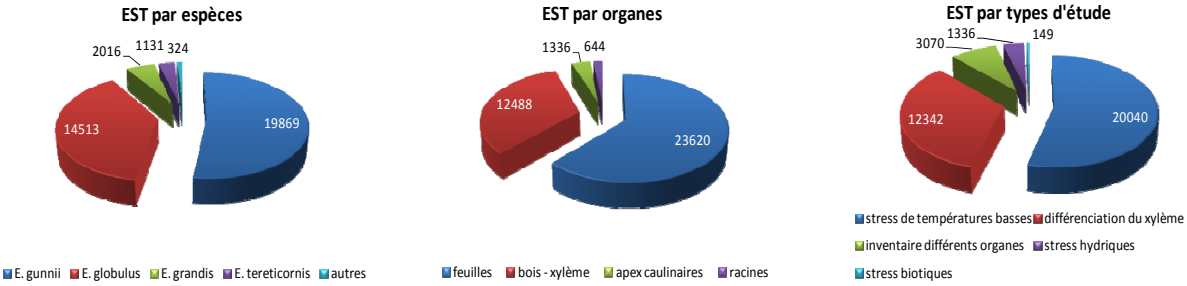


Figure 20: Nombre d'EST d'*Eucalyptus* publiés dans dbEST (NCBI) en Mars 2011.

De plus, 3 archives de séquences d'ADNc par NGS sont déposés sur la base de données NCBI : un concernant le séquençage par Illumina d'un hybride *E. grandis x E. urophylla*, un second de séquences 454 *d'E. grandis*, et le troisième de séquences 454 d'*E. globulus*.

Deux projets de séquençage du génome de deux espèces d'*Eucalyptus* sont en cours : il s'agit d'*E. grandis* séquencé par le JGI aux Etats-Unis et d'*E. camaldulensis* séquencé par le Kazusa DNA Research Institute du Japon. Par ailleurs, certaines caractéristiques de base du génome d'eucalyptus sont connues à partir du séquençage de banques BAC (Paiva *et al.*, 2011), particulièrement le fait que son génome contienne peu d'éléments répétés. En effet, il est estimé que seulement 4,2% du génome d'*E. camaldulensis* et 5,8% du génome d'*E.grandis* présente des éléments répétés, un taux plus bas que les 10% observés chez Arabidopsis, ce qui faciliterait son séquençage (Myburg *et al.*, 2006).

De par ses propriétés agronomiques, il a donc été l'objet de nombreux programmes de recherche dans différentes régions du monde (Brésil, Afrique du sud), tant au niveau génétique pour la sélection de génotypes améliorés, qu'au niveau éco-physiologique pour l'amélioration des méthodes de sylviculture.

REFERENCES BIBLIOGRAPHIQUES

Abe, H., K. YamaguchiShinozaki, et al. (1997). "Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression." Plant Cell **9**(10): 1859-1868.

Adams, M. D., S. E. Celniker, et al. (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185-2195.

Aitken, S. N., S. Yeaman, et al. (2008). "Adaptation, migration or extirpation: climate change outcomes for tree populations." Evolutionary Applications **1**(1): 95-111.

Alagna, F., N. D'Agostino, et al. (2009). "Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development." BMC Genomics **10**(399): (26 August 2009).

Allen, C. D., A. K. Macalady, et al. (2010). "A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests." Forest Ecology and Management **259**(4): 660-684.

Alpert, P. and E. L. Simms (2002). "The relative advantages of plasticity and fixity in different environments: when is it good for a plant to adjust?" Evolutionary Ecology **16**(3): 285-297.

Alvim, F. C., S. M. B. Carolino, et al. (2001). "Enhanced accumulation of BiP in transgenic plants confers tolerance to water stress." Plant Physiology **126**(3): 1042-1054.

Andries, K., P. Verhasselt, et al. (2005). "A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis." Science **307**(5707): 223-227.

Apel, K. and H. Hirt (2004). "Reactive oxygen species: Metabolism, oxidative stress, and signal transduction." Annual Review of Plant Biology **55**: 373-399.

Argout, X., J. Salse, et al. (2010). "The genome of Theobroma cacao." Nat Genet **43**(2): 101-108.

Arndt, S. K., S. J. Livesley, et al. (2008). "Quercitol and osmotic adaptation of field-grown Eucalyptus under seasonal drought stress." Plant Cell and Environment **31**(7): 915-924.

Ashwell, G. (1957). "Colorimetric analysis of sugar." Methods in Enzymology **4**: 73-105.

Assmann, S. M. (1995). "Cyclic AMP as a Second Messenger in Higher Plants (Status and Future Prospects)." Plant Physiology **108**(3): 885-889.

Aubin-Horth, N. and S. C. P. Renn (2009). "Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity." Molecular Ecology **18**(18): 3763-3780.

Audic, S. and J. M. Claverie (1997). "The significance of digital gene expression profiles." Genome Research **7**(10): 986-995.

Bahrun, A., C. R. Jensen, et al. (2002). "Drought-induced changes in xylem pH, ionic composition, and ABA concentration act as early signals in field-grown maize (Zea mays L.)." Journal of Experimental Botany **53**(367): 251-263.

Barakat, A., D. S. DiLoreto, et al. (2009). "Comparison of the transcriptomes of American chestnut (Castanea dentata) and Chinese chestnut (Castanea mollissima) in response to the chestnut blight infection." BMC Plant Biology **9**(51): (9 May 2009).

Barros, P. D., N. M. Soares-Cavalcanti, et al. (2009). "In Silico Evaluation of Osmoprotectants in Eucalyptus Transcriptome." Computational Intelligence Methods for Bioinformatics and Biostatistics **5488**: 66-77.

Bartels, D. and R. Sunkar (2005). "Drought and salt tolerance in plants." Critical Reviews in Plant Sciences **24**(1): 23-58.

Bartels, D. and R. Sunkar (2005). "Drought and salt tolerance in plants. Crit. Rev. in Plant Sci, 241-36."

Bartoli, C. G., M. Simontacchi, et al. (1999). "Drought and watering-dependent oxidative stress: effect on antioxidant content in Triticum aestivum L. leaves." Journal of Experimental Botany **50**(332): 375-383.

Battaglia, M., Y. Olvera-Carrillo, et al. (2008). "The enigmatic LEA proteins and other hydrophilins." Plant Physiology **148**(1): 6-24.

Beaudoin, N., C. Serizet, et al. (2000). "Interactions between abscisic acid and ethylene signaling cascades." Plant Cell **12**(7): 1103-1115.

Beck, E. H., S. Fettig, et al. (2007). "Specific and unspecific responses of plants to cold and drought stress." Journal of Biosciences **32**(3): 501-510.

Bedon, F., J. Majada, et al. (2010). "Interaction between environmental factors affects the accumulation of root proteins in hydroponically grown Eucalyptus globulus (Labill.)." Plant Physiology and Biochemistry **49**(1): 69-76.

Bellin, D., A. Ferrarini, et al. (2009). "Combining next-generation pyrosequencing with microarray for large scale expression analysis in non-model species." BMC Genomics **10**(555): (24 November 2009).

Boisvenue, C. and S. W. Running (2006). "Impacts of climate change on natural forest productivity - evidence since the middle of the 20th century." Global Change Biology **12**(5): 862-882.

Borel, C., A. Frey, et al. (2001). "Does engineering abscisic acid biosynthesis in Nicotiana plumbaginifolia modify stomatal response to drought?" Plant, Cell & Environment **24**(5): 477-489.

Bradshaw, A. D. (1965). The evolutionary significance of phenotypic plasticity in plants.

Bray, E. A. (1997). "Plant responses to water deficit." Trends in Plant Science **2**(2): 48-54.

Bray, E. A. (2002). "Abscisic acid regulation of gene expression during water-deficit stress in the era of the Arabidopsis genome." Plant, Cell & Environment **25**(2): 153-161.

Breda, N., R. Huc, et al. (2006). "Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences." Annals of Forest Science **63**(6): 625-644.

Brondani, R. P. V., E. R. Williams, et al. (2006). "A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus." Bmc Plant Biology **6**.

Buckman, H. O. and N. C. Brady (1965). Les Sols : Nature et propriétés Précis d'édaphologie. Intercontinental Editions, Inc.

Cabral, O. M. R., H. R. Rocha, et al. (2010). "The energy and water balance of a Eucalyptus plantation in southeast Brazil." Journal of Hydrology **388**(3-4): 208-216.

Callinan, P. A. and A. P. Feinberg (2006). "The emerging science of epigenomics." Human Molecular Genetics **15**: R95-R101.

Calvet, R. (2003). Sol, propriétés et fonctions, DUNOD.

Campion, J. M., M. Nkosana, et al. (2006). "Biomass and N and P pools in above- and below-ground components of an irrigated and fertilised Eucalyptus grandis stand in South Africa." Australian Forestry **69**(1): 48-57.

Carle, J. and P. Holmgren (2008). "Wood from Planted Forests." Forest Products Journal **58**(12): 6-18.

Carneiro, M., A. Fabiao, et al. (2008). "Effects of harrowing and fertilisation on understory vegetation and timber production of a Eucalyptus globulus Labill. plantation in Central Portugal." Forest Ecology and Management **255**(3/4): 591-597.

Chaisson, M., P. Pevzner, et al. (2004). "Fragment assembly with short reads." Bioinformatics **20**(13): 2067-2074.

Chaisson, M. J. and P. A. Pevzner (2008). "Short read fragment assembly of bacterial genomes." Genome Research **18**(2): 324-330.

Chambel, M. R., J. Climent, et al. (2005). "Phenotypic plasticity: a useful framework for understanding adaptation in forest species." Investigacion Agraria, Sistemas y Recursos Forestales **14**(3): 334-344.

Chartzoulakis, K., A. Patakas, et al. (1999). "Changes in water relations, photosynthesis and leaf anatomy induced by intermittent drought in two olive cultivars." Environmental and Experimental Botany **42**(2): 113-120.

Cha-um, S. and C. Kirdmanee (2010). "Effects of water stress induced by sodium chloride and mannitol on proline accumulation, photosynthetic abilities and growth characters of eucalyptus (<i>Eucalyptus camaldulensis</i> Dehnh.)." New Forests **40**(3): 349-360.

Chaves, M. M., J. P. Maroco, et al. (2003). "Understanding plant responses to drought - from genes to the whole plant." Functional Plant Biology **30**(3): 239-264.

Chevin, L. M., R. Lande, et al. (2010). "Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory." Plos Biology **8**(4).

Cook, P. G., T. J. Hatton, et al. (1999). "Water balance of a tropical woodland ecosystem, Northern Australia: a combination of micro-meteorological, soil physical and groundwater chemical approaches." Journal of Hydrology (Amsterdam) **210**(1/4): 161-177.

Dahl, F., J. Stenberg, et al. (2007). "Multigene amplification and massively parallel sequencing for cancer mutation discovery." Proceedings of the National Academy of Sciences of the United States of America **104**(22): 9387-9392.

Dalloul, R. A., J. A. Long, et al. (2010). "Multi-platform next-generation sequencing of the domestic turkey (Meleagris gallopavo): genome assembly and analysis." PLoS Biol **8**(9).

Dassanayake, M., J. S. Haas, et al. (2009). "Shedding light on an extremophile lifestyle through transcriptomics." New Phytologist **183**(3): 764-775.

Dat, J., S. Vandenabeele, et al. (2000). "Dual action of the active oxygen species during plant stress responses." Cellular and Molecular Life Sciences **57**(5): 779-795.

Davis, T. M., M. E. Shields, et al. (2009). "The strawberry genome is coming into view." Acta Horticulturae(842): 533-536.

Deschamps, S. and M. A. Campbell (2010). "Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery." Molecular Breeding **25**(4): 553-570.

Desprez-Loustau, M. L., B. Marcais, et al. (2006). "Interactive effects of drought and pathogens in forest trees." Annals of Forest Science **63**(6): 597-612.

deWitt, T. J., A. Sih, et al. (1998). "Costs and limits of phenotypic plasticity." Trends in Ecology and Evolution **13**: 77-81.

Deyholos, M. K. (2010). "Making the most of drought and salinity transcriptomics." Plant, Cell & Environment **33**(4): 648-654.

Drake, P. L., D. S. Mendham, et al. (2009). "A comparison of growth, photosynthetic capacity and water stress in Eucalyptus globulus coppice regrowth and seedlings during early development." Tree Physiology **29**(5): 663-674.

Ehleringer, J. R. and T. A. Cooper (1992). "ON THE ROLE OF ORIENTATION IN REDUCING PHOTOINHIBITORY DAMAGE IN PHOTOSYNTHETIC-TWIG DESERT SHRUBS." Plant Cell and Environment **15**(3): 301-306.

El Kayal, W., G. Keller, et al. (2006). "Regulation of tocopherol biosynthesis through transcriptional control of tocopherol cyclase during cold hardening in Eucalyptus gunnii." Physiologia Plantarum **126**(2): 212-223.

Eldridge, K., J. Davidson, et al. (1993). Eucalypt domestication and breeding. Oxford.

Eveland, A. L., D. R. McCarty, et al. (2008). "Transcript profiling by 3-Untranslated region sequencing resolves expression of gene families." Plant Physiology **146**(1): 32-44.

FAO, Ed. (2009). State of the world's forests 2009. Rome.

Foucart, C., E. Paux, et al. (2006). "Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in Eucalyptus." New Phytologist **170**(4): 739-752.

FRA (2010). "Global Forest Resources Assessment 2010: main report." FAO Forestry Paper(163): xxxi + 341 pp.

Freeman, J. S., S. P. Whittock, et al. (2009). "QTL influencing growth and wood properties in Eucalyptus globulus." Tree Genetics & Genomes **5**(4): 713-722.

Fullwood, M. J., C. L. Wei, et al. (2009). "Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses." Genome Research **19**(4): 521-532.

Gamboa, M. C., S. Rasmussen-Poblete, et al. (2007). "Isolation and characterization of a cDNA encoding a CBF transcription factor from E. globulus." Plant Physiology and Biochemistry **45**(1): 1-5.

Gechev, T. S., F. Van Breusegem, et al. (2006). "Reactive oxygen species as signals that modulate plant stress responses and programmed cell death." Bioessays **28**(11): 1091-1101.

Gilardoni, P. A., S. Schuck, et al. (2010). "SuperSAGE analysis of the Nicotiana attenuata transcriptome after fatty acid-amino acid elicitation (FAC): identification of early mediators of insect responses." BMC Plant Biology **10**(66): (14 April 2010).

Gion, J. M., P. Rech, et al. (2000). "Mapping candidate genes in Eucalyptus with emphasis on lignification genes." Molecular Breeding **6**(5): 441-449.

Gonzalez-Ballester, D., D. Casero, et al. (2010). "RNA-Seq Analysis of Sulfur-Deprived Chlamydomonas Cells Reveals Aspects of Acclimation Critical for Cell Survival." Plant Cell **22**(6): 2058-2084.

Gonzalez-Martinez, S. C., K. V. Krutovsky, et al. (2006). "Forest-tree population genomics and adaptive evolution." New Phytologist **170**(2): 227-238.

Granier, A., M. Reichstein, et al. (2007). "Evidence for soil water control on carbon and water dynamics in European forests during the extremely dry year: 2003." Agricultural and Forest Meteorology **143**(1-2): 123-145.

Grans, D., B. Hannrup, et al. (2009). "Genetic variation and relationships to growth traits for microfibril angle, wood density and modulus of elasticity in a Picea abies clonal trial in southern Sweden." Scandinavian Journal of Forest Research **24**(6): 494-503.

Grattapaglia, D. and M. Kirst (2008). "Eucalyptus applied genomics: from gene sequences to breeding tools." New Phytologist **179**(4): 911-929.

Gugerli, F., L. Parducci, et al. (2005). "Ancient plant DNA: review and prospects." New Phytologist **166**(2): 409-418.

Guisan, A. and W. Thuiller (2005). "Predicting species distribution: offering more than simple habitat models." Ecology Letters **8**(9): 993-1009.

Guo, S., Y. Zheng, et al. (2010). "Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types." BMC Genomics **11**(384): (17 June 2010).

Hall, N. (2007). "Advanced sequencing technologies and their wider impact in microbiology." Journal of Experimental Biology **210**(9): 1518-1525.

Halliwell, B. and J. M. C. Gutteridge (1999). Free radicals in biology and medicine. O. U. Press. Oxford. **3rd ed**.

Hamrick, J. L. (2004). "Response of forest trees to global environmental changes." Forest Ecology and Management **197**(1-3): 323-335.

Hamrick, J. L. (2004). "Response of forest trees to global environmental changes." Forest Ecology and Management **197**(1-3): 323-335.

Hernandez, D., P. Francois, et al. (2008). "De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer." Genome Research **18**(5): 802-809.

Hong-Bo, S., C. Li-Ye, et al. (2008). "Calcium as a versatile plant signal transducer under soil water stress." BioEssays **30**(7): 634-641.

Huang, S. W., R. Q. Li, et al. (2009). "The genome of the cucumber, Cucumis sativus L." Nature Genetics **41**(12): 1275-U29.

Huang, X., X. Wei, et al. (2010). "Genome-wide association studies of 14 agronomic traits in rice landraces." Nat Genet **42**(11): 961-967.

Hubbard, R. M., J. Stape, et al. (2010). "Effects of irrigation on water use and water use efficiency in two fast growing Eucalyptus plantations." Forest Ecology and Management **259**(9): 1714-1721.

IPCC (2007). Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland.

Jackson, M. (1997). "Hormones from roots as signals for the shoots of stressed plants." Trends in Plant Science **2**(1): 22-28.

Jenks, M. A., P. M. Hasegawa, et al. (2007). Transcriptome Analysis of Plant Drought and Salt Stress Response. Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops, Springer Netherlands**:** 261-283.

Jiang, H. and W. H. Wong (2009). "Statistical inferences for isoform expression in RNA-Seq." Bioinformatics **25**(8): 1026-1032.

Julliard, J., F. Pelese, et al. (1993). "TL-DNA TRANSFORMATION DECREASES ABA-LEVEL." Physiologia Plantarum **88**(4): 654-660.

Jump, A. S. and J. Penuelas (2005). "Running to stand still: adaptation and the response of plants to rapid climate change." Ecology Letters **8**(9): 1010-1020.

Kaul, S., H. L. Koo, et al. (2000). "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana." Nature **408**(6814): 796-815.

King, D. A. (1997). "The functional significance of leaf angle in Eucalyptus." Australian Journal of Botany **45**(4): 619-639.

Kirst, M., A. A. Myburg, et al. (2004). "Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus." Plant Physiology **135**(4): 2368-2378.

Kohn, M. H., W. J. Murphy, et al. (2006). "Genomics and conservation genetics." Trends in Ecology & Evolution **21**: 629-637.

Krutovskii, K. V. and D. B. Neale (2001). "Forest genomics for conserving adaptive genetic diversity." Forest Genetic Resources(29): 6-8.

Kube, P. D., C. A. Raymond, et al. (2001). "Genetic parameters for diameter, basic density, cellulose content and fibre properties for Eucalyptus nitens." Forest Genetics **8**(4): 285-294.

Kulheim, C., S. Yeoh, et al. (2009). "Comparative SNP diversity among four Eucalyptus species for genes from secondary metabolite biosynthetic pathways." BMC Genomics **10**(452): (24 September 2009).

Laclau, J. P., M. Arnaud, et al. (2001). "Spatial distribution of Eucalyptus roots in a deep sandy soil in the Congo: relationships with the ability of the stand to take up water and nutrients (vol 21, pg 129, 2001)." Tree Physiology **21**(9): 631-632.

Lam Son Phan, T., T. Urao, et al. (2007). "Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in Arabidopsis." Proceedings of the National Academy of Sciences of the United States of America **104**(51): 20623-20628.

Lam-Son Phan, T., K. Nakashima, et al. (2004). "Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter." Plant Cell **16**(9): 2481-2498.

Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Lawlor, D. W. (1995). The effects of water deficit on photosynthesis. Environment and plant metabolism. N. Smirnoff. Oxford, Bios Scientific Publishers**:** 129-160.

Levitt, J. (1972). Responses of Plants to Environmental Stresses. New York, London.

Li, R. Q., W. Fan, et al. (2010). "The sequence and de novo assembly of the giant panda genome." Nature **463**(7279): 311-317.

Macfarlane, C., C. Bond, et al. (2010). "Transpiration and hydraulic traits of old and regrowth eucalypt forest in southwestern Australia." Forest Ecology and Management **260**(1): 96-105.

Margulies, M., M. Egholm, et al. (2005). "Genome sequencing in microfabricated high-density picolitre reactors." Nature **437**(7057): 376-380.

Marioni, J. C., C. E. Mason, et al. (2008). "RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays." Genome Research **18**(9): 1509-1517.

Marques, C. M., R. P. V. Brondani, et al. (2002). "Conservation and synteny of SSR loci and QTLs for vegetative propagation in four Eucalyptus species." Theoretical and Applied Genetics **105**(2-3): 474-478.

Martinez-Meier, A., L. Sanchez, et al. (2009). "Ring density record of phenotypic plasticity and adaptation to drought in Douglas-fir." Forest Ecology and Management **258**(5): 860-867.

Martins, F. B., N. A. Streck, et al. (2008). "Soil water defcit affecting transpiration, growth and development of seedlings of two Eucalyptus species." Revista Brasileira De Ciencia Do Solo **32**(3): 1297-1306.

Maurel, C. and M. J. Chrispeels (2001). "Aquaporins. A molecular entry into plant water relations." Plant Physiology **125**(1): 135-138.

M'Bou, A. T., C. Jourdan, et al. (2008). "Root elongation in tropical Eucalyptus plantations: effect of soil water content." Annals of Forest Science **65**(6): 609p1-609p7.

Merchant, A., S. K. Arndt, et al. (2010). "Seasonal changes in carbohydrates, cyclitols, and water relations of 3 field grown Eucalyptus species from contrasting taxonomy on a common site." Annals of Forest Science **67**(1).

Merchant, A., A. Callister, et al. (2007). "Contrasting physiological responses of six Eucalyptus species to water deficit." Annals of Botany **100**: 1507-1515.

Merchant, A., A. D. Peuke, et al. (2010). "Phloem sap and leaf delta C-13, carbohydrates, and amino acid concentrations in Eucalyptus globulus change systematically according to flooding and water deficit treatment." Journal of Experimental Botany **61**(6): 1785-1793.

Metzker, M. L. (2010). "APPLICATIONS OF NEXT-GENERATION SEQUENCING Sequencing technologies - the next generation." Nature Reviews Genetics **11**(1): 31-46.

Millar, C. D., L. Huynen, et al. (2008). "New developments in ancient genomics." Trends in Ecology & Evolution **23**(7): 386-393.

Modrek, B. and C. Lee (2002). "A genomic view of alternative splicing." Nature Genetics **30**(1): 13-19.

Morozova, O. and M. A. Marra (2008). "Applications of next-generation sequencing technologies in functional genomics." Genomics **92**(5): 255-264.

Munnik, T. (2001). "Phosphatidic acid: an emerging plant lipid second messenger." Trends in Plant Science **6**(5): 227-233.

Myburg, A. A., B. M. Potts, et al. (2007). Eucalypts. Genome Mapping and Molecular Breeding in Plants, Forest Trees. C. Kole. Berlin Heidelberg, Deutschland, Springer-Verlag. **7:** 115-160.

Nielsen, K. L., A. L. Hogh, et al. (2006). "DeepSAGE - digital transcriptomics with high sensitivity, simple experimental protocol and multiplexing of samples." Nucleic Acids Research **34**(19).

Nouvellon, Y., J. P. Laclau, et al. (2010). "Within-stand and seasonal variations of specific leaf area in a clonal Eucalyptus plantation in the Republic of Congo." Forest Ecology and Management **259**(9): 1796-1807.

Novaes, E., D. R. Drost, et al. (2008). "High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome." BMC Genomics **9**(312): (30 June 2008).

Nowrousian, M. (2010). "Next-Generation Sequencing Techniques for Eukaryotic Microorganisms: Sequencing-Based Solutions to Biological Problems." Eukaryotic Cell **9**(9): 1300-1310.

Oshlack, A. and M. J. Wakefield (2009). "Transcript length bias in RNA-seq data confounds systems biology." Biology Direct **4**.

Ouborg, N. J., F. Angeloni, et al. (2010). "An essay on the necessity and feasibility of conservation genomics." Conservation Genetics **11**(2): 643-653.

Ouborg, N. J. and W. H. Vriezen (2007). "An ecologist's guide to ecogenomics." Journal of Ecology **95**(1): 8-16.

Paiva, J., E. Prat, et al. (2010). "Advancing Eucalyptus genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries." BMC Genomics **12**(1): 137.

Parry, M., C. Rosenzweig, et al. (2005). "Climate change, and risk global food supply of hunger." Philosophical Transactions of the Royal Society B-Biological Sciences **360**(1463): 2125-2138.

Paux, E., M. Tamasloukht, et al. (2004). "Identification of genes preferentially expressed during wood formation in Eucalyptus." Plant Molecular Biology **55**(2): 263-280.

Petit, R. J., C. Bodenes, et al. (2004). "Hybridization as a mechanism of invasion in oaks." New Phytologist **161**(1): 151-164.

Petit, R. J. and A. Hampe (2006). "Some evolutionary consequences of being a tree." Annual Review of Ecology Evolution and Systematics **37**: 187-214.

Pigliucci, M. (1996). "How organisms respond to environmental changes: From phenotypes to molecules (and vice versa)." Trends in Ecology & Evolution **11**(4): 168-173.

Pigliucci, M. (1996). "Modelling phenotypic plasticity .2. Do genetic correlations matter?" Heredity **77**: 453-460.

Pigliucci, M. and C. D. Schlichting (1996). "Reaction norms of Arabidopsis .4. Relationships between plasticity and fitness." Heredity **76**: 427-436.

Plomion, C., C. Lalanne, et al. (2006). "Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins." Proteomics **6**(24): 6509-6527.

Poke, F. S., R. E. Vaillancourt, et al. (2005). "Genomic research in Eucalyptus." Genetica **125**(1): 79-101.

Pook, E. W., A. M. Gill, et al. (1997). "Long-term variation of litter fall, canopy leaf area and flowering in a Eucalyptus maculata forest on the south coast of New South Wales." Australian Journal of Botany **45**(5): 737-755.

Poorter, H. and O. Nagel (2000). "The role of biomass allocation in the growth response of plants to different levels of light, CO2, nutrients and water: a quantitative review (vol 27, pg 595, 2000)." Australian Journal of Plant Physiology **27**(12): 1191-1191.

Pop, M. and S. L. Salzberg (2008). "Bioinformatics challenges of new sequencing technology." Trends in Genetics **24**(3): 142-149.

Potts, B. M. and R. J. E. Wiltshire (1997). Eucalypt genetics and genecology. Eucalypt Ecology: Individuals to Ecosystems. J. E. Williams and J. C. Z. Woinarski. Cambridge, Cambridge University Press**:** 56-91.

Ragoussis, J. (2009). "Genotyping Technologies for Genetic Research." Annual Review of Genomics and Human Genetics **10**: 117-133.

Reiser, V., D. C. Raitt, et al. (2003). "Yeast osmosensor Sln1 and plant cytokinin receptor Cre1 respond to changes in turgor pressure." Journal of Cell Biology **161**(6): 1035-1040.

Ricker, M., G. Gutierrez-Garcia, et al. (2007). "Modeling long-term tree growth curves in response to warming climate: test cases from a subtropical mountain forest and a tropical rainforest in Mexico." Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere **37**(5): 977-989.

Roh, T. Y., W. C. Ngau, et al. (2004). "High-resolution genome-wide mapping of histone modifications." Nature Biotechnology **22**(8): 1013-1016.

Romualdi, C., S. Bortoluzzi, et al. (2001). "Detecting differentially expressed genes in multiple tag sampling experiments: comparative evaluation of statistical tests." Human Molecular Genetics **10**(19): 2133-2141.

Sanger, F., S. Nicklen, et al. (1977). "DNA SEQUENCING WITH CHAIN-TERMINATING INHIBITORS." Proceedings of the National Academy of Sciences of the United States of America **74**(12): 5463-5467.

Saya, A. R, et al. (2008). Advances in mass clonal propagation of Eucalyptus urophylla x E. grandis in Congo. Montpellier, FRANCE, CIRAD - Forêt.

Scheiner, S. M. (1993). "GENETICS AND EVOLUTION OF PHENOTYPIC PLASTICITY." Annual Review of Ecology and Systematics **24**: 35-68.

Schilmiller, A. L., D. P. Miner, et al. (2010). "Studies of a Biochemical Factory: Tomato Trichome Deep Expressed Sequence Tag Sequencing and Proteomics." Plant Physiology **153**(3): 1212-1223.

Schlichting, C. D. (1986). "THE EVOLUTION OF PHENOTYPIC PLASTICITY IN PLANTS." Annual Review of Ecology and Systematics **17**: 667-693.

Schmitt, C. B., N. D. Burgess, et al. (2009). "Global analysis of the protection status of the world's forests." Biological Conservation **142**(10): 2122-2130.

Seki, M., M. Narusaka, et al. (2002). "Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray." The Plant Journal **31**(3): 279-292.

Seki, M., T. Umezawa, et al. (2007). "Regulatory metabolic networks in drought stress responses." Current Opinion in Plant Biology **10**: 296-302.

Sembdner, G. and B. Parthier (1993). "THE BIOCHEMISTRY AND THE PHYSIOLOGICAL AND MOLECULAR ACTIONS OF JASMONATES." Annual Review of Plant Physiology and Plant Molecular Biology **44**: 569-589.

Senaratna, T., D. Touchell, et al. (2000). "Acetyl salicylic acid (Aspirin) and salicylic acid induce multiple stress tolerance in bean and tomato plants." Plant Growth Regulation **30**(2): 157-161.

Shem, K., M. Catherine, et al. (2009). "Gas exchange responses of <i>Eucalyptus</i> , <i>C. africana</i> and <i>G. robusta</i> to varying soil moisture content in semi-arid (Thika) Kenya." Agroforestry Systems **75**(3): 239-249.

Shi, H. and J.-K. Zhu (2002). "Regulation of expression of the vacuolar Na⁺/H⁺ antiporter gene

<i>AtNHX1</i> by salt stress and abscisic acid." Plant Molecular Biology **50**(3): 543-550.

Shin, H., M. Hirst, et al. (2008). "Transcriptome analysis for Caenorhabditis elegans based on novel expressed sequence tags." BMC Biology **6**(30): (08 July 2008).

Shinozaki, K. and K. YamaguchiShinozaki (1997). "Gene expression and signal transduction in water-stress response." Plant Physiology **115**(2): 327-334.

Shinozaki, K. and K. Yamaguchi-Shinozaki (2007). "Gene networks involved in drought stress response and tolerance." Journal of Experimental Botany **58**(2): 221-227.

Shinozaki, K., K. Yamaguchi-Shinozaki, et al. (2003). "Regulatory network of gene expression in the drought and cold stress responses." Current Opinion in Plant Biology **6**(5): 410-417.

Silva, F. C. E., A. Shvaleva, et al. (2004). "Responses to water stress in two Eucalyptus globulus clones differing in drought tolerance." Tree Physiology **24**(10): 1165-1172.

Slate, J., J. Gratten, et al. (2009). "Gene mapping in the wild with SNPs: guidelines and future directions." Genetica **136**(1): 97-107.

Snsteb, J. H., L. Gielly, et al. (2010). "Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate." Molecular Ecology Resources **10**(6): 1009-1018.

Soares-Cavalcanti, N. M., A. C. Wanderley-Nogueira, et al. (2009). "Comparative In Silico Evaluation of MYB Transcription Factors in Eucalyptus, Sugarcane and Rice Transcriptomes." Computational Intelligence Methods for Bioinformatics and Biostatistics **5488**: 44-55.

Solomon, S., D. Qin, et al. (2007). "Climate change 2007: Synthesis Report. Contribution of Working Group I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers." Climate change 2007: Synthesis Report. Contribution of Working Group I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers: 22 pp.

Stackpole, D. J., R. E. Vaillancourt, et al. (2010). "Age trends in genetic parameters for growth and wood density in Eucalyptus globulus." Tree Genetics and Genomes **6**(2): 179-193.

Stape, J. L., D. Binkley, et al. (2010). "The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production." Forest Ecology and Management **259**(9): 1684-1694.

Stapley, J., J. Reger, et al. (2010). "Adaptation genomics: the next generation." Trends in Ecology & Evolution **25**(12): 705-712.

Sun, C., Y. Li, et al. (2010). "De novo sequencing and analysis of the American ginseng root transcriptome using a GS FLX Titanium platform to discover putative genes involved in ginsenoside biosynthesis." BMC Genomics **11**(262): (24 April 2010).

Tamura, T., K. Hara, et al. (2003). "Osmotic stress tolerance of transgenic tobacco expressing a gene encoding a membrane-located receptor-like protein from tobacco plants." Plant Physiology **131**(2): 454-462.

Tatagiba, S. D., J. E. M. Pezzopane, et al. (2009). "Performance of six clones of eucalyptus in response to substrate water availability." Engenharia na Agricultura **17**(3): 179-189.

Taylor, K. H., R. S. Kramer, et al. (2007). "Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing." Cancer Research **67**(18): 8511-8518.

Teixeira, P. C., J. L. M. Goncalves, et al. (2008). "Eucalyptus sp. seedling response to potassium fertilization and soil water." Ciencia Florestal **18**(1/2): 47-63.

Thamarus, K. A., K. Groom, et al. (2002). "A genetic linkage map for Eucalyptus globulas with candidate loci for wood, fibre, and floral traits." Theoretical and Applied Genetics **104**(2-3): 379-387.

Thomas, C. D., A. Cameron, et al. (2004). "Extinction risk from climate change." Nature **427**(6970): 145-148.

Thumma, B. R., S. G. Southerton, et al. (2010). "Quantitative trait locus (QTL) analysis of wood quality traits in Eucalyptus nitens." Tree Genetics & Genomes **6**(2): 305-317.

Torres, T. T., M. Dolezal, et al. (2009). "Expression profiling of Drosophila mitochondrial genes via deep mRNA sequencing." Nucleic Acids Research **37**(22): 7509-7518.

Tran, L. S. P., K. Nakashima, et al. (2004). "Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter." Plant Cell **16**(9): 2481-2498.

Tripiana, V., M. Bourgeois, et al. (2007). "Combining microsatellites, growth, and adaptive traits for managing in situ genetic resources of Eucalyptus urophylla." Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere **37**(4): 773-785.

Troll, W. and J. Lindsley (1955). "A PHOTOMETRIC METHOD FOR THE DETERMINATION OF PROLINE." Journal of Biological Chemistry **215**(2): 655-660.

Trommeter, M. and J. Weber (2005). Développement durable et changements globaux : le développement durable l'est-il encore pour longtemps ? Biodiversité et changements globaux. adpf. Paris, France, Robert Barbault & Bernard Chevassus-au-Louis**:** 242.

Tsuchihira, A., Y. T. Hanba, et al. (2010). "Effect of overexpression of radish plasma membrane aquaporins on water-use efficiency, photosynthesis and growth of Eucalyptus trees." Tree Physiology **30**(3): 417-430.

Tuskan, G. A., S. DiFazio, et al. (2006). "The genome of black cottonwood, Populus trichocarpa (Torr. & Gray)." Science **313**: 1596-1604.

Tyree, M. T. and F. W. Ewers (1991). "The Hydraulic Architecture of Trees and Other Woody-Plants." New Phytologist **119**(3): 345-360.

Umezawa, T., M. Fujita, et al. (2006). "Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future." Current Opinion in Biotechnology **17**(2): 113-122.

Uno, Y., T. Furihata, et al. (2000). "Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions." Proceedings of the National Academy of Sciences of the United States of America **97**(21): 11632-11637.

Urao, T., B. Yakubov, et al. (1999). "A transmembrane hybrid-type histidine kinase in arabidopsis functions as an osmosensor." Plant Cell **11**(9): 1743-1754.

Valente, M. A. S., J. A. Q. A. Faria, et al. (2009). "The ER luminal binding protein (BiP) mediates an increase in drought tolerance in soybean and delays drought-induced leaf senescence in soybean and tobacco." Journal of Experimental Botany **60**(2): 533-546.

Valladares, F., D. Sanchez-Gomez, et al. (2006). "Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications." Journal of Ecology **94**(6): 1103-1116.

van Kleunen, M. and M. Fischer (2005). "Constraints on the evolution of adaptive phenotypic plasticity in plants." New Phytologist **166**(1): 49-60.

van Orsouw, N. J., R. C. J. Hogers, et al. (2007). "Complexity Reduction of Polymorphic Sequences (CRoPS (TM)): A Novel Approach for Large-Scale Polymorphism Discovery in Complex Genomes." Plos One **2**(11).

Vega-Arreguin, J. C., E. Ibarra-Laclette, et al. (2009). "Deep sampling of the Palomero maize transcriptome by a high throughput strategy of pyrosequencing." BMC Genomics **10**(299): (6 July 2009).

Velasco, R., A. Zharkikh, et al. (2010). "The genome of the domesticated apple (Malus [times] domestica Borkh.)." Nat Genet **42**(10): 833-839.

Venter, J. C., M. D. Adams, et al. (2001). "The sequence of the human genome." Science **291**(5507): 1304-+.

Vera, J. C., C. W. Wheat, et al. (2008). "Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing." Molecular Ecology **17**(7): 1636-1647.

Vicentini, R., F. v. T. Sassaki, et al. (2005). "In silico evaluation of the Eucalyptus transcriptome." Genetics and Molecular Biology **28**: 487-495.

Violle, C., H. Castro, et al. (2009). "Intraspecific seed trait variations and competition: passive or adaptive response?" Functional Ecology **23**(3): 612-620.

Vitasse, Y. (2009). Déterminisme environnemental et génétique de la phénologie des arbres de climat tempéré. Thèse de doctorat.

Wang, B., G. W. Guo, et al. (2010). "Survey of the transcriptome of Aspergillus oryzae via massively parallel mRNA sequencing." Nucleic Acids Research **38**(15): 5075-5087.

Wang, L. K., Z. X. Feng, et al. (2010). "DEGseq: an R package for identifying differentially expressed genes from RNA-seq data." Bioinformatics **26**(1): 136-138.

Wang, W., Y. Wang, et al. (2009). "Global characterization of Artemisia annua glandular trichome transcriptome using 454 pyrosequencing." BMC Genomics **10**(465): (9 October 2009).

Wang, W. X., B. Vinocur, et al. (2003). "Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance." Planta **218**(1): 1-14.

Wang, Z., M. Gerstein, et al. (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nature Reviews Genetics **10**(1): 57-63.

Warren, C., T. Bleby, et al. (2007). "Changes in gas exchange versus leaf solutes as a means to cope with summer drought in <i>Eucalyptus marginata</i&gt." Oecologia **154**(1): 1-10.

Warren, C. R. (2008). "Does growth temperature affect the temperature responses of photosynthesis and internal conductance to CO2? A test with Eucalyptus regnans." Tree Physiology **28**(1): 11-19.

Weber, A. P. M., K. L. Weber, et al. (2007). "Sampling the arabidopsis transcriptome with massively parallel pyrosequencing." Plant Physiology **144**(1): 32-42.

Whitehead, D. and C. L. Beadle (2004). "Physiological regulation of productivity and water use in Eucalyptus: a review." Forest Ecology and Management **193**(1-2): 113-140.

Wise, M. J. (2003). "LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles." Bmc Bioinformatics **4**.

Wright, S. D. and K. D. M. McConnaughay (2002). "Interpreting phenotypic plasticity: The importance of ontogeny." Plant Species Biology **17**(2-3): 119-131.

Yamaguchi-Shinozaki, K. and K. Shinozaki (2005). "Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters." Trends in Plant Science **10**(2): 88-94.

Yang, J., J. Zhang, et al. (2004). "Activities of fructan- and sucrose-metabolizing enzymes in wheat stems subjected to water stress during grain filling." Planta **220**(2): 331-343.

Yoder-Himes, D. R., P. S. G. Chain, et al. (2009). "Mapping the Burkholderia cenocepacia niche response via high-throughput sequencing." Proceedings of the National Academy of Sciences of the United States of America **106**(10): 3976-3981.

Zahid, D. M., F. Shah, et al. (2010). "Planting Eucalyptus camaldulensis in arid environment - it is useful species under water deficit system?" Pakistanese Journal of Botany **42**(3): 1733-1744.

Zenoni, S., A. Ferrarini, et al. (2010). "Characterization of Transcriptional Complexity during Berry Development in Vitis vinifera Using RNA-Seq." Plant Physiology **152**(4): 1787-1795.

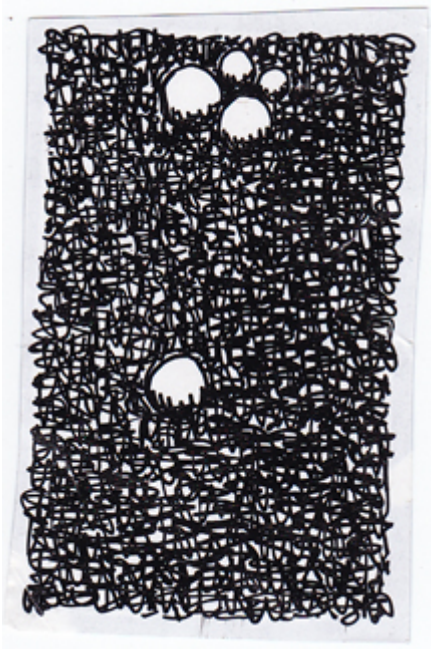
Zeppel, M. J. B., B. R. Murray, et al. (2004). "Seasonal responses of xylem sap velocity to VPD and solar radiation during drought in a stand of native trees in temperate Australia." Functional Plant Biology **31**(5): 461-470.

Zhang, G. J., G. W. Guo, et al. (2010). "Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome." Genome Research **20**(5): 646-654.

Zhang, L., W. R. Dawes, et al. (1996). "Modelling hydrologic processes using a biophysically based model--application of WAVES to FIFE and HAPEX-MOBILHY." Journal of Hydrology **185**(1-4): 147-169.

Zheng, J., J. Zhao, et al. (2004). "Isolation and analysis of water stress induced genes in maize seedlings by subtractive PCR and cDNA macroarray." Plant Molecular Biology **55**(6): 807-823.

Chapitre II - Contexte et objectifs



Ben

1 Contexte de l'étude

Au Congo, l'eucalyptus est exporté pour la fabrication de pâte à papier principalement, sa productivité en plantation industrielle avoisine 20m^3/ha/an. Les surfaces plantées sont d'environ 40 000 ha, avec un projet d'extension à 60 000 ha. En 2007, Eucalyptus Fibre Congo, l'industriel qui exploite les plantations monoclonales d'eucalyptus du Congo a expédié environ 200 000 tonnes de rondins pour des clients Européens.

La productivité des plantations d'eucalyptus au Congo est majoritairement limitée par les conditions écologiques peu favorables. Tout d'abord, le climat est caractérisé par une pluviométrie très variable au cours de l'année : les fortes pluies de la saison humide (souvent entre 50 et 100 mm par phénomène pluvieux) occasionnent des pertes importantes par drainage profond et engendrent une humidité propice au développement de pathogènes, essentiellement des champignons foliaires. La saison sèche (de Mai à Octobre) est exceptionnellement longue pour la latitude considérée, limitant ou empêchant la croissance sur des sols à faible capacité de réserve en eau. De plus, les sols sont assez pauvres chimiquement, et limitent l'approvisionnement minéral nécessaire à de fortes productivités (Laclau *et al.*, 2001). L'extrême saisonnalité du climat impose donc de disposer de génotypes capables de s'adapter à des contraintes continuellement changeantes et des ressources hydriques, minérales et organiques limitées. Ainsi, malgré un effort considérable d'introduction de matériel végétal, le nombre d'espèces potentiellement utilisable reste très réduit.

Les plantations industrielles ont été développées grâce à l'introduction de différences espèces par le Centre Technique Forestier Tropical (CTFT) à partir des années 1950 (Brézard, 1982) qui a conduit à la découverte d'hybrides parfaitement adaptés mais d'origine mal connue : *Eucalyptus* PF1 (*E. alba* x *E. spp*) et *E. 12ABL* x *E. saligna*. Par la suite, les recherches d'hybrides interspécifiques adaptés ont conduit à se focaliser sur le croisement *E. urophylla* x *E. grandis*. Un schéma de sélection récurrente réciproque a alors été mis en place dès 1989 pour améliorer la valeur des hybrides interspécifiques sur la base de leur croissance et de leur aptitude au bouturage (Vigneron, 1991). C'est ainsi qu'à l'heure actuelle une douzaine de génotypes d'*E. urophylla* x *E. grandis* sont utilisés en plantation industrielle.

On estime entre 20 et 30 ans la durée entre la création d'une variété et la fin de sa production. On peut alors se demander si la productivité des variétés clonales créées actuellement sera affectée par les changements climatiques d'ici une vingtaine d'années. Dans ce cas, on peut

imaginer intégrer dans les programmes de sélection des critères permettant de prendre en compte la capacité des génotypes à répondre aux contraintes environnementales futures. Ainsi, la caractérisation de la plasticité phénotypique potentielle des clones issus du programme d'amélioration ainsi que son contrôle génétique devrait permettre à la fois de créer des variétés susceptibles de supporter des évènements climatiques extrêmes (dont on prédit une fréquence plus élevée), et d'implanter l'eucalyptus sur des régions aux caractéristiques climatiques différentes.

2 Objectifs de la thèse

Etant donné le manque de connaissances préalables sur les caractères d'adaptation à la contrainte hydrique chez l'eucalyptus, l'objectif de cette thèse était de décrire un des mécanismes d'adaptation mis en place chez cette espèce lorsqu'elle est soumise à un déficit hydrique : La plasticité phénotypique et moléculaire, le volet diversité (autre moteur de l'adaptation) faisant l'objet d'un programme de recherche spécifique mené au CIRAD. La plasticité sera étudiée sur deux clones soumis à des régimes hydriques différents grâce à de nombreux caractères mesurés à différentes échelles, sur différents organes et sur une cinétique de 18 mois. Ces caractères sont liés à l'architecture aérienne et racinaire, à l'écophysiologie, à l'anatomie foliaire, et aux analyses du protéome et du transcriptome. La prise en compte de deux génotypes permettra d'aborder la plasticité en intégrant la composante "diversité" génétique.

Associant différents champs disciplinaires, ce projet devrait permettre d'avoir une vision globale de la réponse de la plante à la contrainte hydrique, de la plante entière jusqu'au niveau moléculaire. Cette vision "intégrative" d'un même dispositif expérimental est rarement abordée (Chaves *et al.*, 2003) et devrait permettre de mettre en évidence les caractères les plus pertinents à prendre en compte pour l'analyse de la réponse à la contrainte hydrique et les mécanismes moléculaires sous-jacent. L'expression génique constituant une des premières étapes de la cascade de réponses à la contrainte abiotique, le travail de thèse abordera principalement la plasticité moléculaire au niveau du transcriptome.

L'utilisation de deux clones issus de deux générations successives du programme d'amélioration a permis de caractériser leur plasticité à la contrainte hydrique au champ et de quantifier l'interaction génotype-environnement (GxE). Par ailleurs, ce travail a permis de mettre en évidence un groupe de gènes candidats pour cribler la diversité fonctionnelle

présente dans les populations naturelles en relation avec leur plasticité à la contrainte hydrique.

Ce travail s'articule autour de 3 grandes hypothèses et questions de recherche associées, abordées dans les différents chapitres:

H1 – *La gamme environnementale testée et les génotypes utilisés permettent de mettre en évidence des différences de plasticité entre génotypes*

La plasticité mis en évidence est toujours dépendante de la gamme environnementale testée. Dans cette étude, il a été choisi de comparer deux traitements hydriques, l'un correspondant aux conditions des plantations industrielles (non-irrigué pendant la saison sèche), l'autre consistant en une levée de contrainte par irrigation.

D'autre part, s'il est possible de mettre en évidence des mécanismes communs dans la réponse à la contrainte hydrique chez les plantes, et particulièrement chez les pérennes, à l'échelle de la culture, l'intensité de cette réponse dépend du génotype. Les deux génotypes ont été choisis parce qu'ils présentaient des différences en terme de production, d'efficience d'utilisation de l'eau, et de propriétés physiologiques. On peut donc penser qu'ils auront une réponse différente à la contrainte hydrique.

Deux questions de recherche permettront de tester cette hypothèse :
Comment varient différents caractères phénotypiques en réponse à la contrainte hydrique chez les deux clones d'eucalyptus étudiés? La plasticité observée au champ sous contrainte hydrique, présente-t-elle de l'interaction génotype x environnement?

H2 - *Les réponses moléculaires et phénotypiques à la contrainte hydrique varient au cours du développement de la plante et en fonction de son histoire.*

Selon la localisation et le stade de développement de l'organe dans la plante, le jeu des cascades hormonales et les conditions environnementales entraînent des processus physiologiques différents. L'étude étant menée au champ sur 2 saisons de croissance, différents stades de développement ontogéniques seront étudiés. De plus, la saisonnalité (changement de pluviométrie, d'intensité lumineuse) entrainera des variations de la réponse au cours de la cinétique.

Différentes questions de recherche permettront de tester cette hypothèse :
Quelle est la dynamique de la plasticité phénotypique et moléculaire au cours du développement de la plante : quel est le délai de réponse? Ces réponses sont elles réversibles?

Quels sont les effets d'une contrainte hydrique subie au jeune âge sur la *croissance* l'année suivante ?

H3 *– En réponse à la contrainte hydrique, la variation d'expression de certains gènes peut être mise en relation avec la variation observée pour certains caractères.*

Les processus mis en place par les plantes en réponse à la contrainte hydrique peuvent être étudiées à différentes échelles (moléculaire, anatomique, physiologique) et sur différents organes. Des questions importantes sur les mécanismes de passage de la réponse moléculaire à la réponse physiologique restent ouvertes. Dans cette étude, la caractérisation du phénotype à différentes échelles et l'étude du transcriptome permettra d'aborder le rôle possible de certains gènes exprimés lors de la réponse à la contrainte hydrique.

Deux questions de recherche permettront de tester cette hypothèse :
Peut-on associer les catégories fonctionnelles auxquelles appartiennent les gènes présentant des variations d'expression à des caractères phénotypiques mesurés dans le projet? Peut on observer des normes de réaction similaires entre données phénotypiques et moléculaires ?

3 Moyens mis en œuvre

3.1 Matériel Végétal

Les deux clones commerciaux choisis pour cette étude sont issus des programmes d'amélioration de l'eucalyptus au Congo menés par le CRDPI depuis les années 70. Le 1er des deux génotypes étudiés à avoir été utilisé en plantation est le clone PF1, 1-41. C'est un hybride naturel issu de pollinisation non contrôllée d'*E. alba* de provenance indonésienne. Le second clone, 18-50, est un hybride issu de pollinisation contrôlée entre *E. urophylla* et *E. grandis*. Ces deux clones ont été largement installés en plantation industrielle et ont donc été l'objet de nombreuses études pour caractériser leur croissance, leurs caractéristiques écophysiologiques et leurs besoins nutritifs afin d'améliorer les itinéraires sylvicoles associées à cette culture.

La production du clone 18-50 est relativement plus élevée que celle du clone 1-41 : leur accroissement annuel moyen est respectivement de 28,5 et 25,1 m^3/ha/an (Safou-Matondo *et al.*, 2005). Du point de vue physiologique, les paramètres photosynthétiques de ces 2 clones ont été précédemment caractérisés au stade mature par Yann Nouvellon (communication personnelle, tableau 4).

Tableau 4 : Caractéristiques de croissance et physiologiques au stade mature des deux clones utilisés dans cette étude. EUE_{GPP} : Efficience d'utilisation de l'eau, calculé par rapport à la production primaire nette.

	E. PF1 Genotype 1 -41	*E. urophylla* x *E. grandis* Genotype 18-50
Accroissement annuel (m^3 ha^{-1} a^{-1})	25.1	28.5
Surface spécifique moyenne (m^2 kg^{-1})	11.1	11.6
Surface foliaire moyenne (cm^2)	56.3	26.8
Disposition des stomates	amphistomatique	hypostomatique
EUE_{GPP} (gC kg^{-1} H_20)	2.2	4.3

Ainsi, ces deux génotypes ont été choisis dans cette étude pour leurs paramètres physiologiques et de croissance très contrastés. La comparaison à différentes échelles temporelles de ces 2 génotypes permettra de mieux appréhender les mécanismes biologiques à l'origine des variations de productivité au cours des saisons sèches et des pluies.

3.2 Expérimentation au champ

Dans la région du Kouilou, où sont installées les plantations d'eucalyptus au Congo, le climat est tropical humide, avec des saisons bien délimitées : au cours de la saison sèche de Juin à Septembre, les précipitations sont très faibles voire nulles, alors qu'au cours de la saison des pluies, les précipitations s'élèvent à 1200mm. Au cours de l'année, l'amplitude thermique est très faible (de l'ordre de 5°C), et la photopériode varie peu (figure 21). Les caractéristiques de cette région permettent donc d'avoir un terrain d'expérimentation dont peu de paramètres météorologiques varient au cours de l'année, mise à part la saison sèche très marquée avec arrêt des précipitations et diminution des radiations solaires. Cet arrêt des précipitations a pour conséquences une diminution de l'accroissement des arbres et représente une contrainte abiotique au champ qui réduit la productivité en plantation industrielle.

Figure 21: Photographie du dispositif expérimental en mars 2008, les parties irriguées et non irriguées étant indiquées par des flèches.

Le dispositif expérimental a été mis en place au champ, au bord du lac Yanika (4°20'S, 11°38'E ; figure 2), sur les terres du village de Lamba, situé à environ 60 km de Pointe Noire. Les sols de cette région sont des arenosols ferralitiques (classification FAO) caractérisés par une texture sableuse homogène avec une faible disponibilité en nutriments et en matière organique (Laclau, 2001).

Les 2 clones ont été soumis à 3 traitements hydriques différents : irrigués pendant 2 saisons sèches (IR) / Non irrigués pendant 2 saisons sèches (NI) / Non irrigués la 1ère saison sèche, irrigué la seconde (IR2). L'irrigation a été réalisée par aspersion tous les trois jours dont le volume correspond à l'évapotranspiration potentielle, estimée à 3mm/jour (Bouillet *et al.*, 1998). L'eau d'irrigation était puisée dans le lac voisin grâce à une pompe en immersion, et le volume d'irrigation a pu être contrôlé grâce à un compteur d'eau installé en sortie de pompe. Le débit de la pompe ne permettant pas d'arroser les placeaux simultanément, chaque placeau était irrigué l'un après l'autre pendant la nuit suivant une rotation de 3 jours (2 placeaux par nuit).

L'essai, situé sur une parcelle de 80m de large par 250m de long (soit 2ha) est constitué de 6 placeaux principaux correspondant à chaque point de prélèvement destructif (figure 22). Les deux génotypes ont été reproduits par bouurage en pépinière et 2176 plants ont été installés le 1er juin 2007 sur l'essai. Chacun de ces placeaux (P1 à P6) est constitué de 4 blocs de 24 arbres représentant les deux génotypes soumis aux 2 traitements IR ou NI. L'essai étant situé sur une légère pente, les différents blocs n'ont pas été installés de façon aléatoire. Ainsi, les blocs irrigués étaient situés sur la partie la plus basse de l'essai pour éviter l'écoulement des eaux d'irrigation dans la partie non irriguée. Ces blocs ont été séparés par des lignes de bordure de 8 arbres entre la partie irriguée et non irriguée. Quatre rangées d'arbres séparaient les blocs de l'extérieur de l'essai. Deux rangées d'arbres séparaient également les différents placeaux de prélèvement.

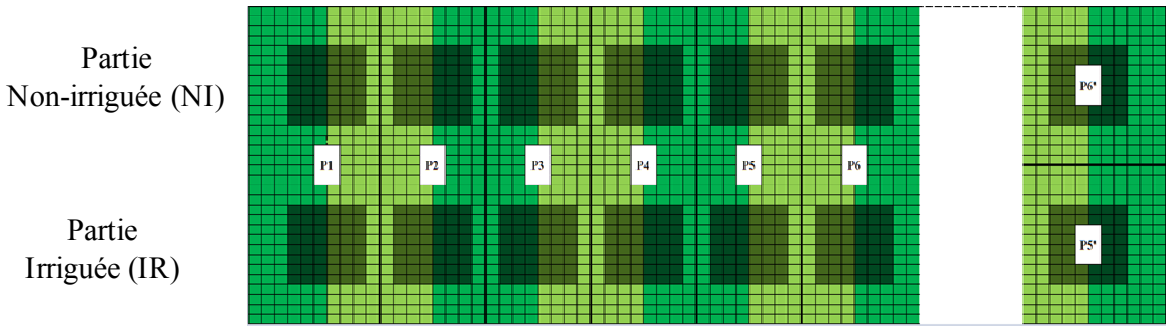


Figure 22 : Plan du dispositif expérimental : en vert clair sont indiqué les copies clonales du génotype 1-41, en vert sombre celles du génotype 18-50. Chaque placeau est délimité par une ligne plus sombre et les blocs de prélèvements sont d'une couleur foncée.

Par ailleurs, les placeaux 5' et 6' n'ont pas été irrigués au cours de la 1ère saison sèche mais ont été irrigués lors de la 2nde année afin d'étudier l'effet d'une contrainte hydrique sur la croissance des plants en 2nde année.

Ce dispositif a été étudié pendant 20 mois (2 saisons sèches) et a fait l'objet de mesures en continu (caractérisation du milieu : données météorologiques et humidité du sol, suivi de l'architecture aérienne) et de prélèvements destructifs ponctuels en fonction de la saison (mesures anatomiques, architecturales, physiologiques). Au total 6 points de prélèvements destructifs ont été réalisés de Juillet 2007 à Janvier 2009 correspondants aux placeaux de P1 à P6, les placeau P5' et P6' étant prélevés en même temps que les placeau 5 et 6, respectivement (Figure 23).

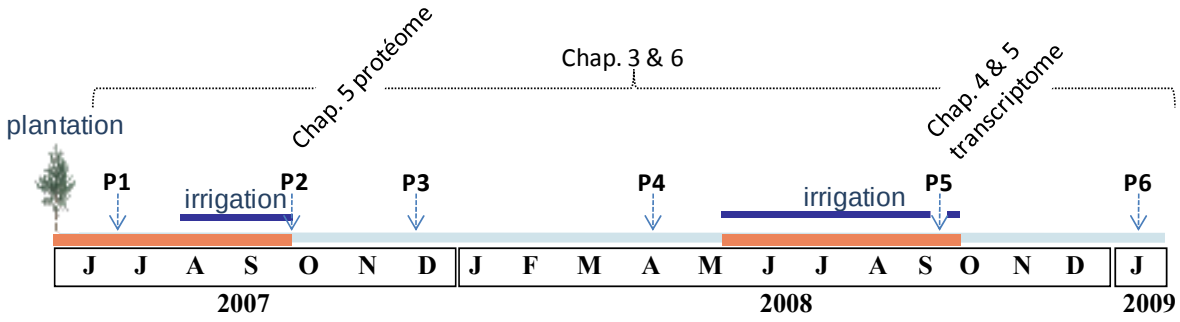


Figure 23 : Déroulement de l'expérimentation de terrain au cours des 20 mois de suivi. La 1ère année, le traitement irrigué (barres bleu foncé) n'a été appliqué qu'une partie de la saison sèche (barres rouge), tandis qu'elle a été apporté tout au long de la saison sèche en 2ème année. Les points de prélèvement (P) sont indiqués par des flèches.

3.3 Caractérisation des paramètres environnementaux

Diverses données météorologiques ont été enregistrées sur l'essai par une station météo (Davis Instruments Corp., Hayward, CA, USA). Cette station météo enregistrait les valeurs moyennes sur la demi-heure de la température, de l'humidité de l'air, de la vitesse et la direction du vent, ainsi que du rayonnement global. La pluviométrie tombée durant la demi-heure était également mesurée. L'absence de données de données dus aux problèmes ponctuels d'alimentation électrique de la station a été comblée grâce aux données météorologiques d'une station installée sur un autre essai du CRDPI, à Tchizalamou, à 10 km à vol d'oiseau de Yanika et à une distance comparable de la mer. La comparaison des données de Yanika et de Tchizalamou lors de période où les 2 stations fonctionnaient a permis de valider cette démarche.

Afin de faire le suivi hebdomadaire de l'humidité dans le sol, 72 sondes TDR (Time Domain Reflectometry) ont été réparties sur 8 fosses représentant les deux traitements et les deux clones sur le placeau 6 (soit deux fosses par condition). Sur les fosses de 5m de profondeur, trois profils différents ont été mesurés tandis qu'un seul profil a été mesuré sur les fosses d'1m de profondeur. Les sondes ont été installées à 0.15, 0.50, 1, 2, 3, 4 et 5 m de profondeur (figure 24).

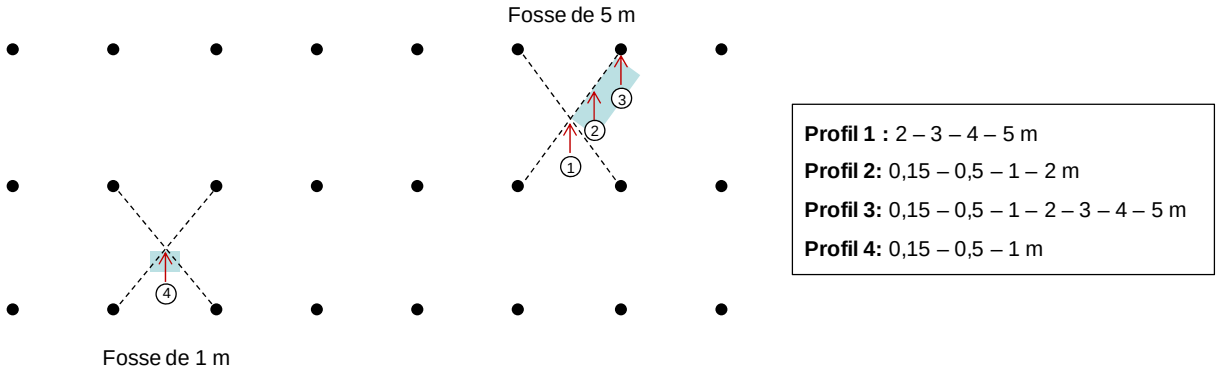


Figure 24: Disposition des sondes TDR sur chaque placeau de prélèvement. Chaque rond noir représente les arbres du placeaux, et les rectangles bleu, les fosses creusées. Pour chaque profil analysé est indiquée la profondeur à laquelle ont été installées les sondes TDR.

A la fin de l'expérimentation, une mesure des propriétés de drainage du sol a été effectuée grâce à une expérience de ressuyage contrôlé, afin de pouvoir dresser le bilan hydrique de l'essai. De même, des prélèvements de sol ont été effectués à 15 et 50 cm de profondeur, ces deux couches représentent l'horizon A contenant de la matière organique plus foncé et un horizon B plus clair. Les prélèvements ont été effectués à l'intérieur et à l'extérieur de la parcelle, sur les placeaux irrigués et non irrigués, sur toute la longueur de l'essai. Les échantillons de sol ont été séchés à l'air libre puis ramenés à l'UMR Eco&Sols de Montpellier pour des analyses en spectrométrie dans le proche infra-rouge (SPIR). L'analyse des spectres proche infra-rouge permet une analyse qualitative de la fertilité du sol. Elle a permis de savoir si : i/ la plantation d'eucalyptus avait modifié les propriétés du sol par rapport aux pourtours de l'essai ; ii/ si l'irrigation avait altéré la fertilité du sol (par lessivage ou captation plus importante par les arbres).

3.4 Diverses mesures phénotypiques associées au projet

Lors de ce projet, un grand nombre de mesures phénotypiques ont été réalisées en collaboration avec différentes équipes.

Tout d'abord, avant chaque point de prélèvement, la hauteur des arbres a été mesurée à la perche, et leur circonférence à l'aide d'un mètre ruban (supervisée par Andréas Ndeko, du

CRDPI). Ces données ont permis de sélectionner les arbres les plus représentatifs du placeau pour les prélèvements destructifs, qui, par leur coût et leur demande en main d'œuvre ne pouvaient pas être réalisés sur les 24 arbres de chaque placeau. Ainsi, à chaque point de prélèvement, 11 arbres dont la hauteur était située autour de la hauteur médiane du placeau ont été prélevés. Pour ces 11 arbres, la biomasse aérienne ainsi que les paramètres morphologiques des feuilles ont été mesurés. Parmi ceux-ci, 7 arbres ont été choisi pour évaluer l'architecture et la biomasse racinaire. Sur les 4 arbres restant, nous avons effectué des prélèvements de tissus pour les analyses anatomiques, chimiques, protéomiques et transcriptomiques.

Les biomasses aériennes ont été évaluées par compartimentation architecturale : le tronc est nommé axe 1, les branches s'insérant sur le tronc sont les axes 2, et ainsi de suite. Feuilles et tiges ont ainsi été traitées séparément, en suivant cette compartimentation. Ces mesures ont été supervisées par André Mabiala, au CRDPI.

Pour quantifier les biomasses racinaires, l'appareil racinaire a été excavé sur un cube d'un mètre de large autour de l'arbre. Dans ce cube, les racines ont été séparées en 3 catégories en fonction de leur taille : les racines grosses (diamètre >10mm), les racines moyennes (2mm<D<10mm), et les racines fines (D<2mm). Ces mesures ont été supervisées par Andréas Ndeko du CRDPI.

Pour mesurer la surface spécifique des feuilles, les aliquotes de 25 feuilles prélevées pour la mesure de biomasse foliaire ont été scannées, afin d'estimer leur surface. La surface spécifique est la surface mesurée rapportée à la masse sèche de l'échantillon. Elle renseignant sur la densité ou l'épaisseur de la feuille. Ces mesures ont été réalisées par André Mabiala du CRDPI.

La densité stomatique a été mesurée sur les faces supérieures et inférieures de 8 feuilles représentatives par arbre (4 arbres par condition), en comptant au microscope les stomates sur des empreintes faites au vernis. Ces mesures ont été effectuées à l'INRA de Pierroton par Pablo Hamlat et Anthony Reychler.

Pour les mesures de discrimination isotopique du carbone, 20 feuilles de lumière par arbre ont été sélectionnées et séchées à l'étuve. Après broyage, elles ont été analysées par spectrométrie de masse au laboratoire de Montpellier par Christophe Jourdan et Pascal Tillard de l'UMR IBIP de Montpellier.

Pour les coupes anatomiques, différents organes ont été prélevés et conservés dans un fixateur: feuilles juvéniles et matures, tiges, apex caulinaires et racinaires, racines âgées. Les coupes et mesures histologiques ont été réalisées au laboratoire PHIV de l'UMR AGAP de Montpellier, sous la supervision de Jean-Luc Verdeil.

L'extraction de polyphénols foliaires et leur mesure par HPLC ont été réalisées au sein de l'UMR Qualisud de Montpellier, et supervisées par Christian Mertz.

Enfin, l'analyse du protéome des feuilles et tiges a été réalisé par séparation sur gel 2D, puis identification des spots d'intérêt par spéctrométrie de masse en tandem. Ce travail a été réalisé par Frank Bedon au sein de l'UMR BioGeCo de Bordeaux.

3.5 Analyse du transcriptome

Si j'ai pu participer à la mise en place de l'essai et aux différents prélèvements et mesures présentée ci-dessus (en 2007 alors que j'étais en volontaire internationale, et par la suite lors de 3 missions de 2 à 3 semaines), l'essentiel du travail de laboratoire au cours de ma thèse s'est focalisé sur les analyses du transcriptome. Deux types d'approches complémentaires ont été suivies :

- Sans *a priori* : l'abondance des transcrits a été estimée grâce au pyroséquençage (454/Roche) de banque d'ADNc non normalisées construites à partir d'ARN d'apex caulinaires.

- Avec *a priori* : Les niveaux d'expression de différents gènes candidats ont été mesurés par PCR quantitative en temps réel (RT-qPCR) pour différents tissus (feuilles, tiges, xylème et apex caulinaires). Les gènes cibles ont été choisis de différentes façons : i) les protéines pour lesquelles ils codent ont été mis en évidence au cours de l'analyse du protéome, ii) ces gènes cibles étaient différentiellement exprimés dans les banques d'ADNc d'apex caulinaires, et d'après une revue bibliographique, ils pouvaient présenter un lien avec les caractères phénotypiques mesurés, iii) enfin, pour les gènes de la voie de biosynthèse des phényl-propanoïdes, nous disposions de séquences déjà publiées chez *Eucalyptus*.

REFERENCES BIBLIOGRAPHIQUES

Bouillet, J. P., J. D. Nzila, et al. (1998). Eucalyptus Plantations in the Equatorial Zone, on the Coastal Plains of the Congo. Site Management and productivity in Tropical Plantation Forests, Pietermaritzburg, South Africa, CIFOR.

Brezard, J. M. (1982). Les eucalyptus introduits au Congo de 1953 à 1981. Pointe-Noire, CTFT-Congo**:** 154.

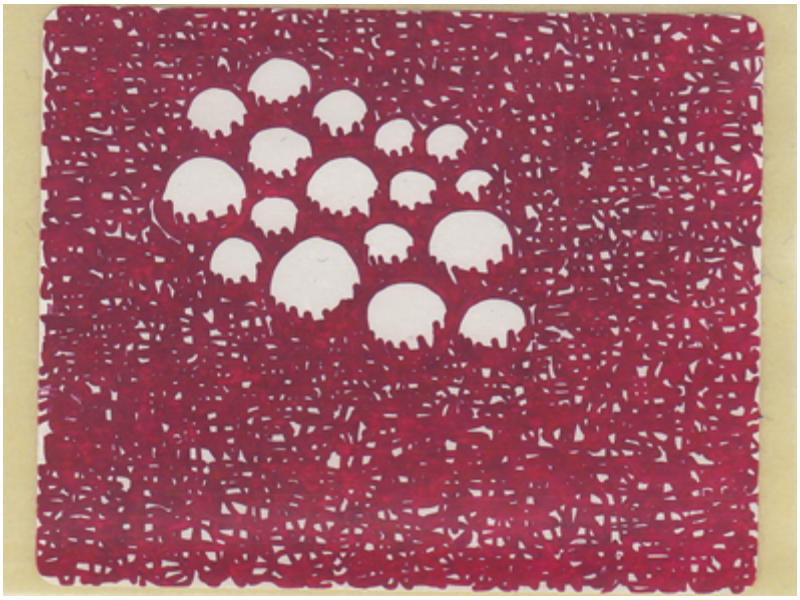
Chaves, M. M., J. P. Maroco, et al. (2003). "Understanding plant responses to drought - from genes to the whole plant." Functional Plant Biology **30**(3): 239-264.

Laclau, J. P., M. Arnaud, et al. (2001). "Spatial distribution of Eucalyptus roots in a deep sandy soil in the Congo: relationships with the ability of the stand to take up water and nutrients (vol 21, pg 129, 2001)." Tree Physiology **21**(9): 631-632.

Safou-Matondo, R., P. Deleporte, et al. (2005). "Hybrid and clonal variability of nutrient content and nutrient use efficiency in Eucalyptus stands in Congo." Forest Ecology and Management **210**(1-3): 193-204.

Vigneron, P. (1991). Création et amélioration des variétés hybrides d'Eucalyptus au Congo. UFRO «Intensive Forestry: the role of Eucalyptus», Durban.

Chapitre III - Plasticité phénotypique de l'*Eucalyptus* en réponse à un déficit hydrique saisonnier au champ



Ben

Phenotypic plasticity in response to seasonal water deficit on field of Eucalyptus

Emilie Villar[1,2,3,4], Philippe Deleporte[5], Agnès de Grandcourt[5], Laurent Saint-André[5], Andréas Ndeko[3], André Mabiala[3], Christophe Plomion[2,4], Christophe Jourdan[5], Jean-Marc Gion[1,4§]

[1]CIRAD, UPR39 Genetic Diversity and Breeding of Forest Tree Species, Campus de Baillarguet TA 10C, F–34398 Montpellier Cedex 5, France

[2]Université de Bordeaux, UMR1202 BIOGECO, F-33610 Cestas, France

[3]CRDPI, BP1291, Pointe Noire, République du Congo

[4]INRA, UMR1202 BIOGECO, F-33610 Cestas, France

[5]CIRAD, UPR80 s/c UMR Eco & Sols - INRA -Sup Agro - 2 Place Viala Bât 12 - 34060 Montpellier Cedex 02, France

[§] Corresponding author

E-mail addresses:

EV: villar@cirad.fr

DP: deleporte@cirad.fr

AdG: agnes.de.grandcourt@cirad.fr

LSA: standre@cirad.fr

AN: andreasndeko@yahoo.fr

AM: mabialangoma_andre@yahoo.fr

CJ: jourdan@cirad.fr

CP: plomion@pierroton.inra.fr

JMG: gion@cirad.fr

1 Introduction

With climate changes, efforts are needed to anticipate consequences on forest survival and productivity (Bonan, 2008; Allen *et al.*, 2010). Among predictable climate variations, drought constitutes one of the major threats on forest sustainability for several species, influencing mortality and productivity, both for natural forests and commercial plantations (Allen et al, 2010; Breda *et al.*, 2006). Numerous studies have already measured impact of drought on productivity in planted forests. As an example, in *Populus*, Monclus *et al.* (2006) have shown that the more productive genotypes seemed to display a larger reduction of biomass in response to mild drought stress, while the less productive genotypes assumed a larger variability in loss of biomass compared to irrigated treatments. In eucalyptus, several studies shown that irrigation increase growth: 30% of mean annual increment in brazilian plantations (Stape *et al.*, 2010) and 12.6% of stem utilizable volume at 4 years in south-african plantations (Campion *et al.* , 2006).

Several studies contributed to a better understanding of drought tolerance processes by studying the physiological responses of eucalyptus. Thus, it was shown that regulation of transpiration could be a key process in drought tolerance driven by interaction between leaf surface, air vapour deficit, soil water deficit and leaf water potential (Whitehead et Beadle, 2004), and it could be mainly controlled by stomatal conductance (Warren *et al.*, 2007). To preserve transpiration capacity, eucalyptus showed a high ability to develop deep root system following water gradient in soil (Drake, 2009; Thongo-Mbou *et al.*, 2008). Finally, the high photosynthetic capacity of eucalyptus has been shown to be greatly affected by water deficit (Cha-Um et Kirdmanee, 2010). Several processes to preserve leaf integrity and then photosynthetic activity have been described, through mechanisms of osmoregulation with polyols, sugars, or amino-acids (Merchant *et al.*, 2007; Warren *et al.* (2007) ; Cha-Um et Kirdmanee, 2010).

Such modifications of traits in response to environmental changes could be regarded as phenotypic plasticity. Phenotypic plasticity is considered as a major mean of adaptation for long-lived organisms such forest trees in the context of climate changes (Thuiller *et al.*, 2008; Valladares, 2007; Matesanz, 2010). Indeed, a large range of phenotypic plasticity may ensure trees to acclimate in many environmental conditions (West- Eberhard, 2003) or to colonize new environments (Sultan *et al.*, 2000). Trees have shown a wide ability to exhibit phenotypic plasticity for several traits (concerning growth, physiology, reproduction…), but it seems that

populations or genotypes facing common environmental variations could present large differences in their plasticity levels, as well as between species (e.g Rozendaal *et al.*, 2006), as at intra-specific level (Monclus *et al.*, 2005). Such differences may result from different selective pressure experienced (Pigliucci, 2005), or different constraints acting on plasticity (De Witt *et al.*, 1998).

Several barriers have to be considered when comparing phenotypic plasticity of different genotypes or species. First, size and age of trees must be taken into account to reveal true phenotypic plasticity. Indeed, phenotypic variation could be more related to growth variations than directly induced by environmental change (Wright and McConnaughay, 2002). As an example, Chambel *et al.* (2005) have shown that *Pinus canariensis* displayed no differences in biomass partitioning between wet and dry environments when size of the plant is taken into account, even if biomasses and shoot/root ratios were significantly different between the two environments.

Besides, growth or physiological traits are widely influenced by seasonality (Delagrange, 2010; Craven *et al.*, 2010). Craven (2010), studying eight tropical tree species, described two different strategies developed by different fast growing species in a site exhibiting large seasonal variations (dry and wet seasons). The first one is characterized by a high ability to adjust photosynthetic parameters when water became scarce ("drought avoider" strategy), while the other one is depicted by no change in photosynthetic characteristics, but maintenance of high growth rate, probably thanks to a deep rooting ("drought tolerant" strategy).

In Congo, Bouvet (1995) has shown that classification changes in breeding programs based on growth parameters occurs during the three first years of growth. This suggests that phenotypic plasticity part may be more important during these first years, when trees are more sensitive to environmental parameters. In this study, we studied growth and physiological parameters during the two first years of growth on an experimental design comparing two commercial clones which differs in their productivity level, subjected to two watering regimes (irrigated or not) during the dry seasons. We will examine variation of phenotypic plasticity and test if their differences in productivity could be explained by differences in phenotypic plasticity, taking into account, if possible, ontogenetic effects.

2 Materials and Methods

2.1 Study site

The study site is located in Yanika (4°20'S, 11°38'E, 100m a.s.l. 50m), on the Atlantic coastal plains of Republic of Congo. The plantation has been installed on native vegetation which was savannah dominated by the C4 Poaceae *Loudetia arundinacea* (Hochst.) Stend and *Hypparrhenia diplandra* (Hack.) Stapf (Laclau *et al.*, 2002). The deep sandy soils are classified as ferralic arenosols according to the FAO classification (Trouvé *et al.*, 1994). These soils contained mainly sands (80-90%) and few clay (8-10%) and silt (2-2.5%) (Mareschal *et al.*, 2011). Their cation exchange capacity (0.3 to 0.6 $cmol_c.kg^{-1}$ of soil) and C content (0.9% in the 0-10cm layer) are very low, indicating very poor soils (Nzila, 2002). Moreover these soils are characterized by a low water retention and low level of organic matter (Epron *et al.*, 2004; D'Annunzio *et al.*, 2008). Climate is characterized by high mean annual air humidity (85%) and temperature (25°C). Mean annual precipitations are around $1250mm.yr^{-1}$, with occurrence of a dry season from May to September marked by cloudiness lower temperatures and no rainfalls.

At the end of the field trial in January 2009, soil samplings were done to study effects of irrigation and plantation on soil fertility. Six samplings per plot were done with soil auger at 5 and 50 cm depth: 4 within the field trial and 2 outside the field trial, in the surrounding savannah. Two series of six sampling have been added in the surrounding savannah. Soil samples were air dried during 5 days. Near infrared reflectance spectroscopy (NIRS) was used to characterize the spectral signature of soil organic matter (SOM). Reflectance was measured between 1100 and 2500 nm at 2 nm intervals with a Foss NIRSystems 5000 spectrophotometer (Silver Spring, MD, USA), and analyzed following the protocol described by Brunet *et al.* (2007). ACP of spectrum were done with XLStat module of excel.

2.2 Plant material

We compared two eucalyptus hybrids used in clonal plantations, and obtained from the eucalypt breeding program conducted by CRDPI. The hybrid *E. PF1*, clone 1-41 (NCBI Taxonomy ID: 764271) was a genotype resulting from natural crossing of *E. alba* mother tree. At a mature stage, genotype 1-41 has a mean annual increment of 25.1 $m^3.ha^{-1}.y^{-1}$ (Safou-Matondo *et al.*, 2005), a water use efficiency of 2.2 $gC.kg^{-1}H2O$, a mean leaf area of

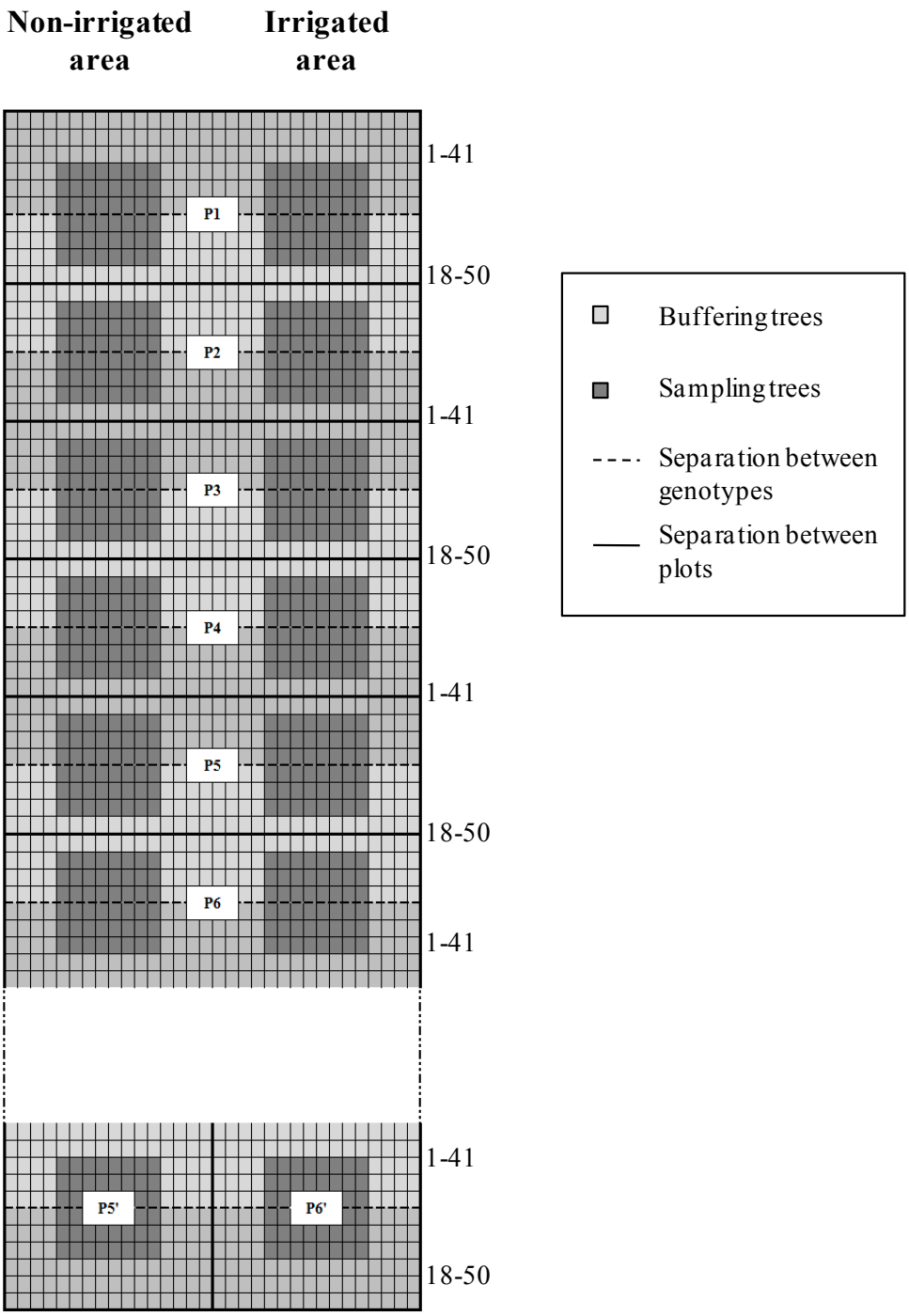


Figure 1 – The field trial was divided in 8 plots (P) : P1 to P6'. The plots were sampled at different sampling campaigns indicated by the number following "P". Each sampling plot consisted in 24 trees per genotype and per treatment and was separated by buffering zones. The irrigated part of the experiment was placed down slope. P5' and P6' were irrigated only during the second dry season. The two genotypes (1-41 and 18-50) were installed alternatively on the field.

56.3 cm^2, and is amphistomatic (personal communication, Dr. Y. Nouvellon, UMR80, CIRAD, France). The second one is issued from controlled pollination of *E. urophylla* and *E. grandis*, clone 18-50 (NCBI Taxonomy ID: 765255). At a mature stage, genotype 18-50 has a mean annual increment of 28.5 $m^3.ha^{-1}.y^{-1}$ (Safou-Matondo *et al.*, 2005), a water use efficiency of 4.3 $gC.kg^{-1}H2O$, a mean leaf area of 26.8 cm^2, and is hypostomatic (personal communication, Dr. Y. Nouvellon, UMR80, CIRAD, France). These two genotypes thus display contrasted growth and ecophysiological characteristics at the mature stage: one being more productive and more water-use efficient than the other. The two genotypes have been reproduced by cuttings, grown in nursery during three months before planting.

2.3 Field trial

The experimental trial has been installed in the field in June 2007 at the beginning of the dry season, and totalized 2176 plants with a spacing of 2.5 m x 3.2 m (1250 stems $ha^{-1)}$. The field trial was divided in 8 plots (figure 1). The main part of the experiment was constituted of 6 plots corresponding to 6 sampling campaigns. Each plot comprised 4 blocks of 24 trees corresponding to the two genotypes subjected to the different irrigation treatments. The irrigated treatments were located downslope to avoid water flow into the non-irrigated part. The different blocks were separated by buffering zones of 8 rows of trees between treatments, and 4 rows of trees between trial borders. Sampling parts in the different plots were separated by buffering zones of 2 rows of trees. Two supplementary plots (P5' and P6') were used to study recovery effects: the two genotypes were not irrigated during the first dry season but irrigated during the second dry season.

Before planting, an herbicide treatment (glyphosate, Round-Up®, Mosanto, USA) was applied, and after planting weeds were eliminated manually. During the first two weeks after plantation, each plant was watered (1 L day^{-1}). Starter-fertilization was applied two weeks after planting with 200 g per plant of NH_4NO_3. To prevent insect attacks, a pyrethroid insecticide has been applied.

2.4 Watering regimes

Trees were subjected to two main watering treatments: irrigation vs no-irrigation during the dry seasons. The irrigation was done by aspersion in order to recover potential evapotranspiration, estimated at 3 mm per day (Bouillet *et al.*, 2002). During the experimentation, irrigation was performed for the two dry seasons (2007 and 2008) during which the rainfalls ceased (table 1): from July 31st 2007 to October 24th 2007 (dry season 1,

DS1), and from June 05th 2008 to October 28th 2008 (dry season 2, DS2). Irrigation was also applied during a short period with less rainfalls during the wet season that occurred from February 19th 2008 to March 14th 2008.

Table 1– Irrigation was performed during three periods corresponding to the cessation of rainfalls. Plot 5 and 9 were sampled during the period 3, that is why irrigation had not been performed during the totality of this period on these plots.

				irrigation (mm.day-1)						
period	beginning	end	number of days	P2	P3	P4	P5	P6	P5'	P6'
period 1	31/07/2007	24/10/2007	85	3.6	4.5	3.5	3.3	3.6		
period 2	19/02/2008	14/03/2008	24			3.3	3.6	3.2	3.8	3.4
period 3 (plot 5)	05/06/2008	13/09/2008	100				4			
period 3 (plot 5')	05/06/2008	22/09/2008	109						3.9	
period 3	05/06/2008	28/10/2008	145					3.9		3.8

2.5 Environmental parameters

Meteorological parameters were recorded on the field with a weather station (Davis Instruments Corp., Hayward, CA, USA). Temperature, global radiation and precipitations were recorded every half-hour. Potential evapotranspiration was calculated using Hargreaves equation modified by Pai Wu (1997):

$$ET_0 = 0.0135(T - 17.78)R_s\left(\frac{10}{595.5 - 0.55T}\right)$$

where ET_0 is the potential evapotranspiration in mm per half-hour, T is the mean temperature per half-hour expressed in °C and R_s is the incident solar radiation per half-hour expressed in Langley.

Soil water content was measured weekly in the field with a time-domain reflectometer system (TDR; Trase system, Soil Moisture, Santa Barbara, CA) on one plot (P6). On each sampling part (2 genotypes x 2 treatments), 18 probes were installed at different distances from trees (4 profiles) and different depths (15 cm, 50 cm, 1 m, 2 m, 3 m, 4 m and 5 m). At each date, water reserves were estimated by differences between soil moisture measured by TDR probes and soil moisture at permanent wilting point (pF 4.2) for each soil layer, multiplied by the layer thickness. Water reserve was then computed as the addition of soil moisture reserve of all the layers up to 5.5 m deep. The permanent wilting point (pF4.2), as well as pF1, pF2 (field capacity) and pF3 were estimated through internal drainage experiment conducted at the end of the experiment. This internal drainage experiment enabled to adjust the relationship between humidity and soil water potential by the equation of Van Genuchten (1980).

2.6 Tree samplings measurements

Six sampling campaigns were done during the first two years of growth (table 2): in July 2007 (at the beginning of the dry season 1, no watering treatment was applied); in September 2007 (end of the dry season 1); in December 2007 (beginning of the wet season 1); in April 2008 (end of the wet season 1); in September 2008 (end of the dry season 2) and in January 2009 (middle of the wet season 2). The two supplementary plots (P5' and P6') were sampled in September 2008 and January 2009 respectively. At each sampling point 24 to 4 trees per condition were analyzed as described in the table 2.

Table 2 – Number of trees sampled during the experiment, at each sampling campaign.

				Number of trees measured per genotype and per treatment							
Sampling Plot	Sampling Campaign	Date	Trees age (months)	HT	C	VOB	above-ground biomass	below-ground biomass	delta C13	SLA / MLA	SD
P1	1	July-07	2	24			11	7		11	
P2	2	September-07	4	24			11	7	4	10	4
P3	3	December-07	7	24	24		11	5		9	4
P4	4	April-08	11	24	24	24	11	7	9	11	
P5	5	September-08	16	24	24	24	11	7	11	11	5
P5'	5	September-08	16	24	24	24	11	7	11	11	
P6	6	January-09	20	24	24	24	11	7	11	10	
P6'	6	January-09	20	24	24	24	11	7	11	8	

HT: height; C: circumference; VOB: volume over bark; SLA: specific leaf area; MLA: mean leaf area; SD: stomatal density.

2.7 Growth and biomass measurements

Before each sampling campaign, height (HT) and circumference at breast height (C) were measured on all the trees, on the plots remaining after previous destructive sampling campaigns. Volume over bark was computed using stem taper equation developed by Gomat *et al.* (2010). As destructive measurements were expensive and time-consuming, among the 24 trees available per condition on each sampling plot, 11 trees were selected for sampling according to the median height of the plot, and measurements were done on all or part of these 11 trees (table 2). For above-ground biomass, 11 trees per condition were harvested, divided in the following compartments: stem, dead branches, living branches, leaves. Living branches and leaves were sorted by their architectural compartments (trunk: Axe 1; branches inserted on stem : Axe 2, etc.). Below-ground biomass was estimated by root system excavation on 7 trees per conditions. Roots and stump were extracted from a one meter square centred on the tree.

Each part was weighted in the field. Each compartment was sub-sampled to evaluate the moisture content. Sub-samples were weighted before and after drying at 65°C to constant weight. Moisture content was computed for each sub-sample and applied to total measured

samples so as to retrieve total dry biomass for each compartments. Stem dry weight (SDW) correspond to the sum of trunk and branches biomasses, leaf dry weight (LDW) correspond to the sum of all leaves compartment biomasses, and root dry weight correspond to the sum of all roots compartment biomasses.

The different biomasses measured (SDW, LDW and RDW) were modeled in function of growth parameters (C and HT) as described in the supplemental file 1. Models were then applied to all the trees of the experimental design for which C and HT had been measured, to estimate SDWe, LDWe and RDWe. Biomass increment was then computed as differences in biomass between two measurement campaigns and reported in $T.ha^{-1}.y^{-1}$. Shoot / root ratio (S/R) has then been computed by dividing estimated dry above-ground (SDW and LDW) biomass with estimated root dry weight (RDW).

2.8 Physiological measurements

For each architectural compartment, 25 representative leaves were sub-sampled on the 11 trees sampled per condition. Their fresh mass and area were immediately measured. These leaves were then dried at 65°C to constant weight. The average of leaf areas were computed for each tree, giving the mean leaf area (MLA). The measured areas of these sub-samples were divided by their dry weights to calculate specific leaf area (SLA) for each compartment.

Twenty-five sun leaves from 6 to 11 trees sampled per condition were harvested, dried at 65°C to constant weight and grounded. An aliquot of 10g was weighted and analyzed for carbon isotopic composition ($\delta^{13}C$) using an elemental analyzer EA-2000 (EuroVector SpA, Milan, Italy) coupled to a mass spectrometer Elementar (Isoprime Ltd, Cheadle, United Kingdom). The isotopic ratios were expressed as delta values in parts per thousand and not as carbon isotope composition (Farquhar and Richards, 1984) because the isotopic composition of carbon dioxide at the site was not measured.

Stomatal density was determined with 6 leaves from 4 different trees per condition. A print of each side of leaves (adaxial and abaxial) was performed using transparent nail varnish and tape (Ceulemans *et al.*, 1995). These prints were observed using an Olympus BM2 RFCA light microscope (Olympus America Inc., Melville, New York, USA) and 5 microscopic fields per print of leaves (i.e. n=30 per leaves side and per condition) at a 20X magnification (representing 0.568 mm2) were analyzed.

2.9 Statistical analysis

For growth traits (HT, C and VOB), means and standard errors were computed with the maximum of data available, measured on different plots at each date. We evaluated plot (P), genotype (G), treatment (T) and genotype by treatment interaction (GxT) effects, by ANOVA, according to the following model:

$$X_{ijkl} = X_{mean} + aP_i + bG_{ij} + cT_{ik} + dGxT_{ijkl} + \varepsilon_{ijkl}$$

where X_{ijkl} is the trait measured in the plot I for the genotype j (1-14 or 18-50), subjected to the treatment k (NI or IR) and for the replicate l. a, b, c and d are the regressors, and ε_{ijkl} is the residual. Least-square means and standard errors were computed using lsmeans package from R (R Development Core Team).

For other measurements, means and standard errors were computed for each genotype and treatment at each sampling date. We evaluated genotype (G), treatment (T) and genotype by treatment interaction (GxT) effects, by two-way ANOVA, according to the following model:

$$X_{ijk} = X_{mean} + aG_i + bT_j + cGxT_{ij} + \varepsilon_{ijk}$$

where X_{ijk} is the trait measured in genotype i (1-14 or 18-50), treatment j (NI or IR) and for the replicate k. a, b and c are the regressors, and ε_{ijk} is the residual.

A plasticity index (PI) was computed as follow: $PI = (mean_{IR} - mean_{NI})/mean_{IR}$. PI was computed for each plot separately when data were available (HT, C, VOB, and estimated SDW, LDW, RDW).

3 Results

3.1 Soil homogeneity on the experimental trial

As expected, NIRS spectra of soil revealed large differences between the two sampled horizons. At 5 cm depth, soil was dark reflecting a high part of organic matter, while at 50 cm depth, soil was light highly composed of sands. A PCA was performed for each depth separately, with samples issued from the different plots inside the experimental site, and soils from surrounding savannah (figure 2). At 5 cm depth, axe F1 explained 98.47 % of the variation and F2 only 1.28 %, and at 50 cm depth, axe F1 explained 88.48 % of the variation and F2, 8.46 %. No specific structure was found both at 5 and 50 cm depth between the different plots, revealing a wide homogeneity of soils throughout the experimental site, whatever the treatment or the genotype. More surprisingly, savannah and eucalyptus soil samples seems to be confound, suggesting that soil organic matter had not been widely affected during the two first year of eucalyptus growth, soils under eucalyptus plantation remaining as poor as in savannah, as observed for C and N content by D'annunzio *et al.* (2008).

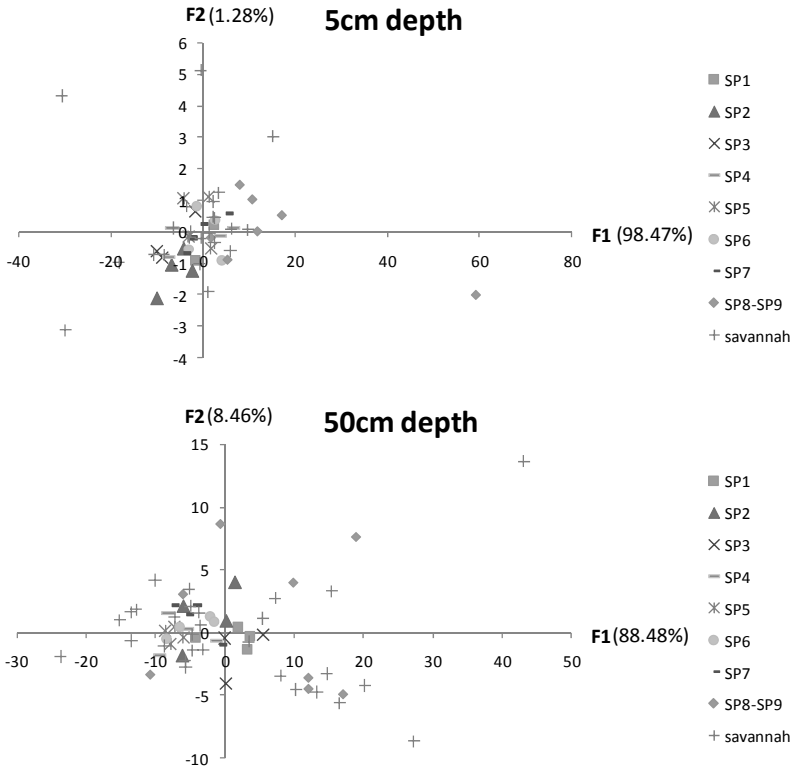


Figure 2 – PCA results for soil NIR spectra at 5cm and 50cm depth on the different sampling plots of the field trial and the surrounding savannah.

3.2 Climate parameters and water supply characterizations during the experiment

During the experiment, global radiations decreased from 16 $MJ.m^{-2}$ in average at the height of the wet season (February, March and April) to 9.6 $MJ.m^{-2}$ in average at the height of the dry season (August-September; figure 3). Such decreases in global radiations may be related to the increase of cloudiness occurring during dry seasons. Temperatures were relatively stables during the year, but were related to cloudiness and global radiations with minima around 22°C in June and July, and maxima around 26.5°C in February, March and April. At the end of the first dry season, precipitations began in October 16th 2007. During the first wet season studied, rainfalls cessation occurred between February 5th 2008 and March 15th 2008 (figure 3). This rainfall cessation is currently observed during the first quarter of the year, and is called "little dry season". During the summer 2008 (i.e. second dry season), no rainfall was observed from May 19th to October 27th 2008. Precipitations restarted with a significant lower level in 2008 (260 mm in October and November) compared to 2007 (882 mm for the same months). Apart from the first days of the dry seasons where irrigation was not yet performed, the amount of water supplied by irrigation recovered potential evapo-transpiration losses.

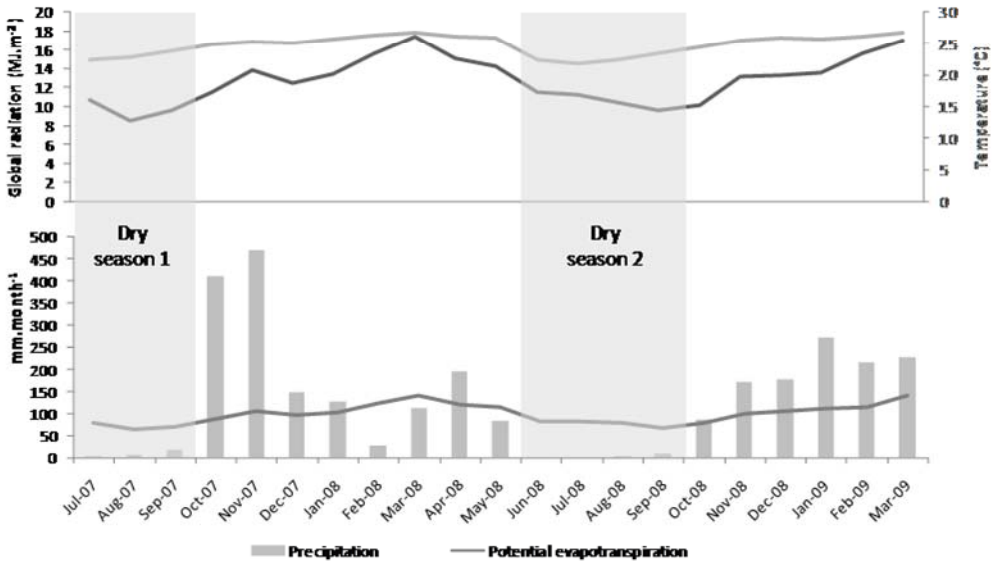


Figure 3 – Meteorological records done during the experiment: global radiations (dark grey), temperatures(light grey), precipitations and potential evapo-transpiration. Irrigations were performed during the two dry seasons, when the rainfalls ceased.

In 2007, soil water reserve (SWR; figure 4) was around 300 mm for all the genotypes and treatments before the start of irrigation, excepted for the plot containing 18-50IR which

contained 230 mm of water. Irrigation on this plot increased SWR until 350 mm at the end of the first dry season, and until 470 mm on the plot containing 1-41IR. For 18-50 and 1-41, the NI plots showed similar profiles for water reserve that slightly decrease during the first dry season to 270 mm. Thus, during the first dry season, water reserve in NI plots were higher than the field capacity (pF2, corresponding to a SWR of 220 mm), suggesting that water was available for the trees. Precipitations were very important in November 2007, and thus SWR reached more than 500 mm whatever the treatment or genotype.

Unfortunately, a technical problem with the acquisition station prevented soil moisture measurement on the field from December 2007 to February 2008.

In 2008, SWR in NI and IR plots presented similar profiles during the wet season (from February to June), and were up to the field capacity (pF2). During the second dry season (corresponding to the second application of irrigation), SWR in IR plots was up to pF1 with 430 mm in average. Differences between genotypes were observed in this treatment, SWR being higher for 18-50. Conversely, SWR decreased from 350 mm to 90 mm in average (below the pF3) in the NI treatment. SWR differences in NI treatment between October 2007 and 2008 could be due to a higher water consumption for 14 months-old trees (October 2008) compared to 2 months-old trees (October 2007).

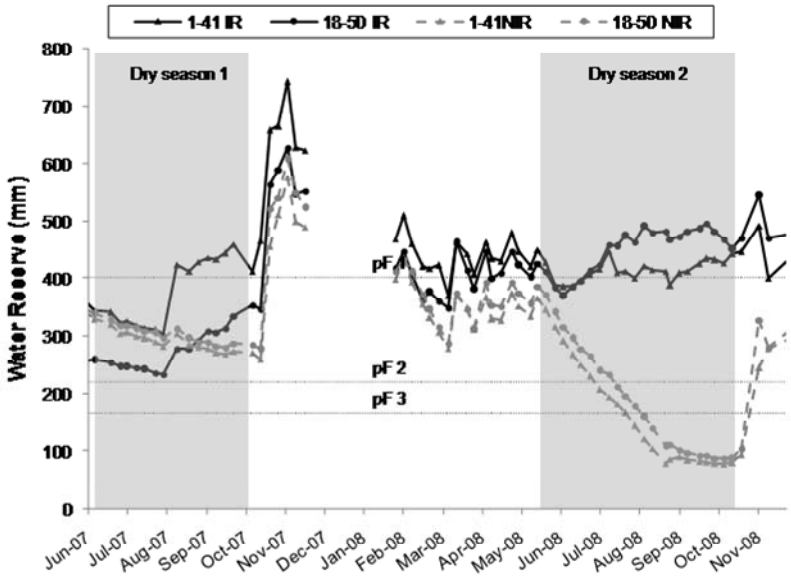


Figure 4 – Evolution of water reserve was modeled using soil volumetric water content measured with TDR trases installed on P6 for the two genotypes (1-41; 18-50) subjected to the two treatments (NIR; IR). Three period of irrigation were represented in grey. Values of the soil retention curves were indicated : pF1: maximal retention, pF2: field capacity, pF3: point of strong plant growth inhibition.

Table 3 – Summary statistics of growth traits measured (HT, C) and estimated according to stem taper equations (VOB) on the field trial, at different dates.

			Lsmeans ± SE						Explained variance (η2)			
date	**# SP**		**1-41IR**	**1-41NIR**	**18-50IR**	**18-50NIR**	**1-41IR2**	**18-50IR2**	**SP**	**G**	**T**	**GxT**
18/07/2007 (DS1)	8	HT	0.31 ^{a} ± 0.01	0.33 ^{a} ± 0.01	0.31 ^{a} ± 0.01	0.33 ^{a} ± 0.01			0.031**	0	0.001	0
14/09/2007 (DS1)	7	HT	0.63 ^{a} ± 0.01	0.5 ^{b} ± 0.01	0.59 ^{a} ± 0.01	0.48 ^{b} ± 0.01			0.013	0.014**	0.202***	0
		C	1.7 ^{b} ± 0.1	1 ^{c} ± 0	1.9 ^{a} ± 0.1	1.2 ^{c} ± 0			0.013	0.016**	0.274***	0.001
29/01/2008 (WS1)	5	HT	2.81 ^{c} ± 0.06	2.31 ^{d} ± 0.04	3.85 ^{a} ± 0.06	3.05 ^{b} ± 0.04			0.091***	0.325***	0.107***	0.009**
		C	7.6 ^{b} ± 0.2	5.1 ^{d} ± 0.2	9.3 ^{a} ± 0.2	6.4 ^{c} ± 0.2			0.071***	0.084***	0.175***	0.002
		VOB	1.5 ^{b} ± 0.1	0.7 ^{d} ± 0.1	2.2 ^{a} ± 0.1	1 ^{c} ± 0.1			0.054***	0.061***	0.182***	0.015**
07/04/2008 (WS1)	5	HT	4.19 ^{c} ± 0.08	3.65 ^{d} ± 0.06	5.92 ^{a} ± 0.08	4.84 ^{b} ± 0.06			0.081***	0.427***	0.087***	0.015***
		C	12.4 ^{b} ± 0.3	10 ^{d} ± 0.2	13.8 ^{a} ± 0.3	11.1 ^{c} ± 0.2			0.048***	0.047***	0.126***	0.001
		VOB	3.9 ^{b} ± 0.2	2.5 ^{d} ± 0.2	6 ^{a} ± 0.2	3.4 ^{c} ± 0.2			0.052***	0.123***	0.148***	0.017**
15/07/2008 (DS2)	2	HT	6.3 ^{d} ± 0.1	5.73 ^{e} ± 0.1	8.68 ^{a} ± 0.1	7.74 ^{b} ± 0.1	5.28 ^{f} ± 0.1	6.96 ^{c} ± 0.1	0.003	0.567***	0.052***	0.011**
		C	18.1 ^{a,b} ± 0.4	16.3 ^{c,d} ± 0.4	19.5 ^{a} ± 0.4	17.2 ^{b,c} ± 0.4	15.4 ^{d} ± 0.4	15.8 ^{c,d} ± 0.4	0.002	0.024**	0.083***	0.004
		VOB	9.8 ^{b,c} ± 0.4	7.7 ^{d,e} ± 0.4	14.5 ^{a} ± 0.4	10.6 ^{b} ± 0.4	6.5 ^{e} ± 0.4	8.5 ^{c,d} ± 0.4	0.002	0.158***	0.097***	0.021**
05/09/2008 (DS2)	2	HT	7.32 ^{c} ± 0.11	6.05 ^{d} ± 0.11	9.48 ^{a} ± 0.11	8.04 ^{b} ± 0.11	6.21 ^{d} ± 0.11	8.15 ^{b} ± 0.11	0	0.526***	0.157***	0.001
		C	19.8 ^{a} ± 0.4	17.6 ^{b} ± 0.4	21.1 ^{a} ± 0.4	17.8 ^{b} ± 0.4	17.3 ^{b} ± 0.4	17.7 ^{b} ± 0.4	0.002	0.012*	0.14***	0.006
		VOB	13.1 ^{b} ± 0.5	9.3 ^{c} ± 0.5	18.1 ^{a} ± 0.6	11.7 ^{b} ± 0.6	9.1 ^{c} ± 0.5	11.8 ^{b} ± 0.5	0.001	0.127***	0.188***	0.014*
07/01/2009 (WS2)	1	HT	9.63 ^{a} ± 1.01	7.68 ^{c} ± 0.91	10.09 ^{a} ± 0.69	8.52 ^{b} ± 0.62	8.43 ^{b} ± 0.77	9.88 ^{a} ± 0.79		0.149***	0.369***	0.03*
		C	25.2 ^{a} ± 4.1	19.2 ^{c} ± 3.9	24.6 ^{a} ± 3.5	19.9 ^{c} ± 2.4	21.2 ^{b,c} ± 2.9	22.9 ^{a,b} ± 3.4		0.005	0.296***	0.015
		VOB	25.4 ^{a} ± 8.1	13.3 ^{b} ± 5	25.1 ^{a} ± 6.9	15 ^{b} ± 3.7	16.7 ^{b} ± 4.9	22 ^{a} ± 7		0.021*	0.347***	0.023

Lsmeans ± SE : mean corrected from plot effects and standard error associated; letters indicate classification according to Bonferroni pairwise test. Explained variance of each modeled effect : SP (Sampling Plot), G (Genotype), T (Treatment), GxT (genotype by interaction effect). Stars indicate significance of the tested effects (* P<0.05; ** P<0.01; *** P<0.001).Circumferences were not measured until September 2007, because the trees were too small, and VOB was calculated from January 2008, because the stem taper equation was not calibrated for young trees.

3.3 Cumulated growth

Mortality was very low in this experiment (0.7 % of the trees) and happened during the first 4 months after plantation. No effect of genotype or treatment was observed for mortality.

First, results on cumulated growth indicated long-terms effects of environmental conditions. HT, C and VOB were analyzed by ANOVA (table 3), considering all existing plots at each sampling dates. The number of existing plot decreased along time due to the destructive measurements at each sampling campaign. Trees presented a higher growth on P5 and P6, compared to P4, P5', and P6' in January and April 2008 (during the first wet season, data not shown). At the beginning of the experiment (July 2007, DS1), no effects of G, T or GxT were detected on trees height. This result was expected because no different treatments were applied hitherto on the experiment, irrigation treatment beginning in August 2007. From the end of DS1, G effects were significant for all the traits measured (excepted for C in January 2009): genotype 18-50 presented a higher growth compared to genotype 1-41. Similarly, T effects were significant during all the experiment, from the end of DS1: growth in NI treatment was largely reduced.

On one side, GxT interaction was not significant for C whatever the sampling date. On the other side, HT and VOB during WS1 and even at the beginning of DS2 presented GxT effects. In this case, decrease of HT and VOB in NI treatment were higher for 18-50 than for 1-41. Nevertheless, at the end of the experimentation (January 2009, WS2), HT presented GxT effect, due to a higher decrease from IR to NI treatment in 1-41 compared to 18-50.

3.4 Growth increments

Second, analysis of growth increments indicated instantaneous effects of environmental conditions (figure 5). During dry seasons, the two genotypes did not present obvious differences in growth increment. Conversely, during WS1, genotype 18-50 had globally higher increments compared to genotype 1-41, whereas at the beginning of WS2, genotype 1-41 presented higher increments compared to genotype 18-50. Concerning T effects, growth increments were always higher in IR treatments, excepted between April and July 2008 (transition between WS1 and DS2). Finally, GxT interaction effects on HT increment were significant between January 2008 and April 2008 (end of WS1). During this period, HT increment was reduced in 18-50 subjected to NI treatment whereas no differences were observable between treatments in 1-41. This period correspond to the occurrence of the "little dry season" (between February 5th 2008 and March 15th 2008) that could affect 18-50 growth

in NI treatment. GxT effects were also significant for HT increment during DS2: HT increment were more reduced from IR to NI treatment in genotype 1-41 than in 18-50. The same trend was observed for C and VOB increment at the beginning of WS2.

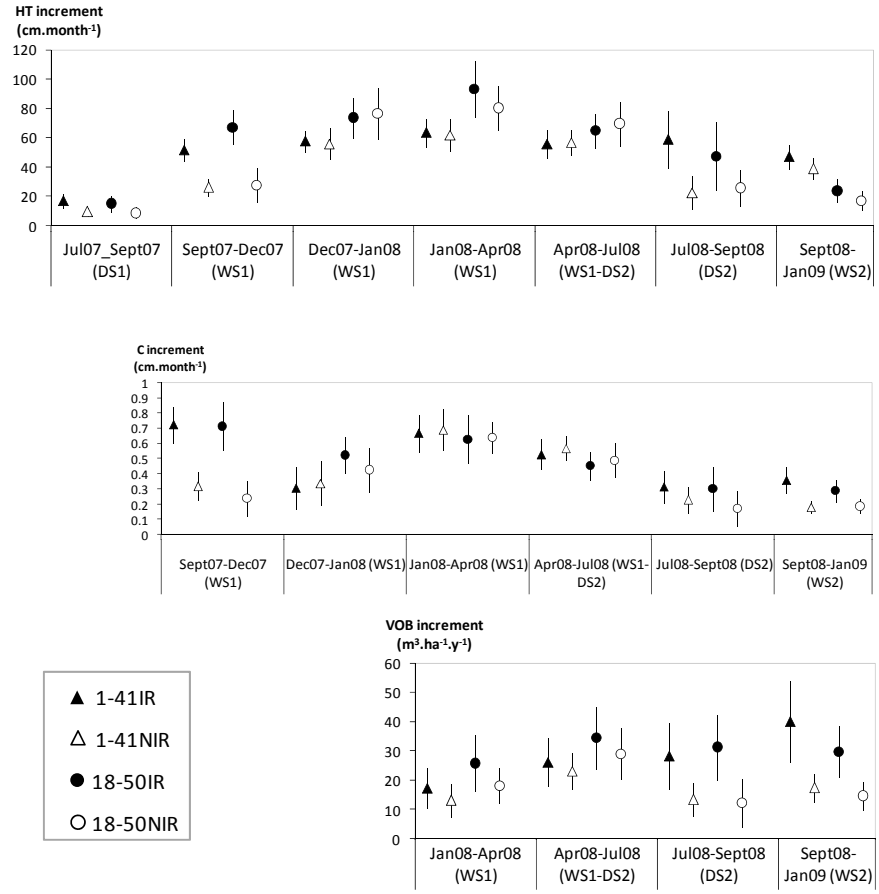


Figure 525 –Increment of height (HT), circumference (C) and estimated volume over-bark (VOB) between each sampling campaigns, corresponding to the first dry season (DS1), the first wet season (WS1), the second dry season (DS2), and the second wet season (WS2). Circumferences were not measured until September 2007, because the trees were too small, and VOB was calculated from January 2008, because the stem taper equation was not calibrated for young trees. Vertical bars indicate standard errors.

3.5 Biomass production

Dissection of biomass partitioning into stem (SDW), leaves (LDW) and root (RDW) compartments enabled to observe effects of environmental conditions on resources allocation during the experiment (table 4).

Observation of G effects confirmed that genotype 18-50 was the more productive clone: SDW were higher for 18-50 during all the experiment (e.g. in January 2009, mean SDW was of 15.0 kg *vs*. 11.5 kg for genotypes 18-50 and 1-41 respectively). LDW was similar between the 2 genotypes excepted in DS2 where LDW was higher for the genotype 18-50 (2.4 *vs.* 2.0 kg

Table 4 – Summary statistics of dry weights measured for stem (SDW), leaf (LDW) root (RDW) compartments and S/R ratio.

		Means ± SE						Explained variance (η^2)		
date	Trait	1-41IR	1-41NIR	18-50IR	18-50NIR	1-41IR2	18-50IR2	G	T	GxT
09/2007 (DS1)	SDW	9.41 [a] ± 3.04	4.3 [b] ± 1.47	11.02 [a] ± 2.79	5.64 [b] ± 0.82			0.046*	0.575***	0
	LDW	23.27 [a] ± 6.86	13.36 [b] ± 4.85	21.8 [a] ± 5.14	11.94 [b] ± 1.38			0.011	0.523***	0
	RDW	4.62 [a] ± 1.23	2.51 [b] ± 0.78	5.4 [a] ± 1.99	3.73 [a,b] ± 0.55			0.1	0.355**	0.005
	S/R ratio	7.54 [a] ± 1.11	6.32 [b] ± 1.05	6.67 [a] ± 1.69	4.6 [a,b] ± 0.5			0.177**	0.288**	0.019
12/2007 (WS1)	SDW	355.6 [b] ± 78.3	74.3 [c] ± 14	547.6 [a] ± 95	79 [c] ± 27.8			0.058***	0.844***	0.053***
	LDW	359.7 [a] ± 75.8	107 [b] ± 21.4	411.5 [a] ± 115.7	91.5 [b] ± 35.5			0.003	0.838***	0.012
	RDW	78.1 [b] ± 20.4	20.2 [b] ± 4.3	145.9 [a] ± 59.6	24.1 [b] ± 5.7			0.087*	0.544***	0.069*
	S/R ratio	7.54 [a] ± 1.11	6.32 [b] ± 1.05	6.67 [a] ± 1.69	4.6 [a,b] ± 0.5			0.574**	0.001	0.036
04/2008 (WS1)	SDW	3938 [a,b] ± 1144	2271 [c] ± 460	5036 [a] ± 1490	3413 [b,c] ± 1043			0.148**	0.32***	0
	LDW	1852 [a] ± 478	1533 [a] ± 258	1684 [a] ± 542	1682 [a] ± 446			0	0.034	0.033
	RDW	590 [a] ± 149	379 [a] ± 50	647 [a,b] ± 158	574 [b] ± 157			0.152*	0.194*	0.046
	S/R ratio	10.42 [a] ± 0.73	10.95 [a] ± 1	10.22 [a,b] ± 1.04	9.01 [b] ± 0.6			0.251**	0.025	0.165*
09/2008 (DS2)	SDW	7747 [b] ± 1054	5829 [c] ± 905	11112 [a] ± 1231	9601 [a] ± 1751	6491 [b,c] ± 950	7314 [b,c] ± 1765	0.355***	0.221***	0.087**
	LDW	2057 [a,b] ± 260	1667 [a] ± 240	2519 [a] ± 358	2311 [a] ± 423	2161 [a] ± 317	2344 [a] ± 447	0.249***	0.096*	0.049
	RDW	1106 [b] ± 73	1059 [b,c] ± 160	1469 [a] ± 175	1362 [a] ± 205	821 [c] ± 112	882 [b,c] ± 131	0.198***	0.487***	0.058*
	S/R ratio	8.88 [a,b] ± 0.69	7.33 [c] ± 0.55	9.88 [a] ± 0.88	8.34 [b,c] ± 0.67	10.45 [a] ± 1.18	10.26 [a,b] ± 0.93	0.049*	0.581***	0.042
01/2009 (WS2)	SDW	16090 [a,b] ± 3460	9061 [c] ± 1701	18971 [a] ± 6787	11873 [b,c] ± 2346	9867 [c] ± 1636	13932 [b,c] ± 3960	0.104**	0.354***	0.003
	LDW	2295 [a,b] ± 716	1532 [b] ± 250	2712 [a] ± 865	1949 [a,b] ± 529	1644 [b] ± 466	2297 [a,b] ± 953	0.104**	0.167**	0.005
	RDW	1593 [a,b] ± 530	1325 [a,b] ± 218	1811 [a] ± 441	1610 [a,b] ± 164	1069 [b] ± 388	1898 [a] ± 598	0.215**	0.05	0.081
	S/R ratio	11.57 [a] ± 2.11	7.74 [b] ± 0.64	8.44 [a,b] ± 3.83	9.19 [a,b] ± 1.06	11.82 [a] ± 1.63	9.26 [a,b] ± 1.39	0.082*	0.127*	0.174*

means ± SE : means and standard error associated; letters indicate classification according to Bonferroni pairwise test. Explained variance of each modeled effect : G (Genotype), T (Treatment), GxT (genotype by interaction effect). Stars indicate significance of the tested effects (* P<0.05; ** P<0.01; *** P<0.001). DS1 corresponds to the first dry season, WS1 to the first wet season, DS2 to the second dry season, and WS2 to the second wet season.

for genotypes 18-50 and 1-41 respectively). From WS1, RDW were also higher for genotype 18-50 (e.g., in January 2009, mean SDW was of 1.77 kg and 1.31 kg for genotypes 18-50 and 1-41 respectively).Treatment effect was mainly significant for all the biomass compartments at all sampling campaign: as expected, biomass production was higher in IR treatment compared to NIR treatment. Some exceptions were observed: no significant differences were displayed for LDW at the end of WS1, suggesting that trees from NIR plots were able to catch up with trees from IR plots for LDW during the wet season. Similarly, in January 2009 (middle of the WS2), RDW showed no significant effect of irrigation treatment.

Finally, few traits showed GxT effects: at the beginning of the WS1, SDW and RDW showed a higher decrease from IR to NI treatment in 18-50 (-86 % and –84 % for SDW and RDW, respectively) compared to 1-41 (-79 % and –74 % for SDW and RDW, respectively). These results suggested different behavior of the genotypes to cope with the water constraint imposed.

3.6 Biomass increment

The models used to predict biomasses in function of C and HT for all trees from P3, P4, P5 and P6 were detailed in supplementary file 1. Due to sampling plot heterogeneity, analysis of biomass increments must be considered carefully, but give nevertheless some trends on evolution of biomasses (figure 6).

First, SDW increment was larger in 18-50 during the wet seasons and larger in IR treatment during all the experiment. Concerning LDW, increments were higher in NI during WS1, confirming the first trend of catch-up observed on cumulated LDW. LDW increment considerably decrease in DS2 and WS2 compared to WS1 (7.4 $T.ha^{-1}.y^{-1}$ in WS1 and 0.9 $T.ha^{-1}.y^{-1}$ in DS2), until became negative for 1-41NI during the beginning of WS2. Finally, RDW increments were higher for genotype 18-50 and in IR treatments during the wet seasons, but during DS2, increments were higher in NI treatment, and particularly for 1-41.

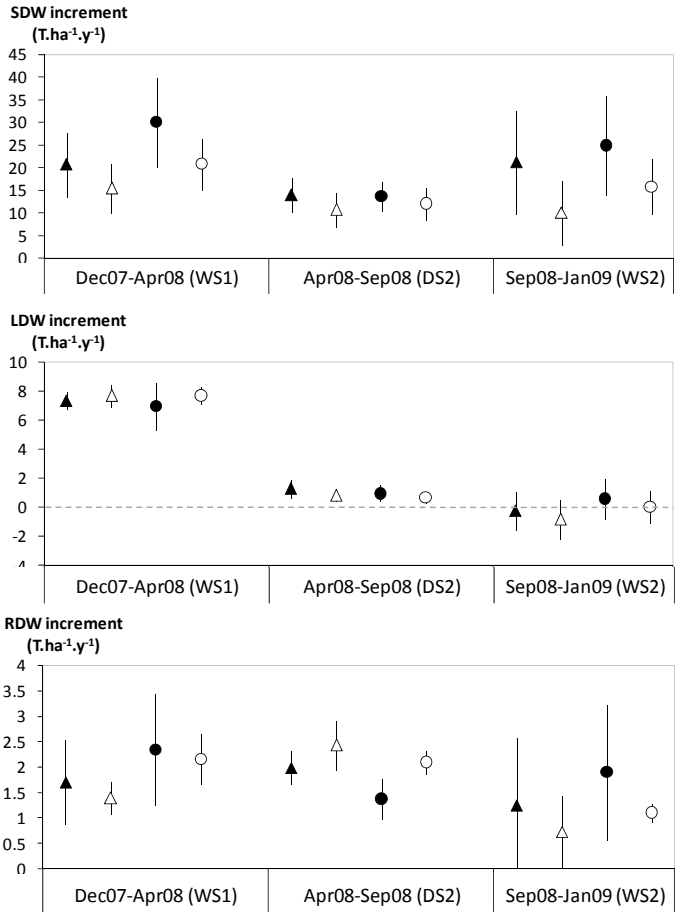


Figure 6 – Increments of stem (SDW), root (RDW) and leaf (LDW) estimated dry weights between each sampling campaign during the first wet season (WS1), the second dry season (DS2), and the second wet season (WS2). Increment estimation were done for the genotypes 1-41 (in triangles) and 18-50 (in circles), subjected to two watering treatments during the dry seasons: irrigated were in black, and non-irrigated in white. Vertical bars indicate standard errors.

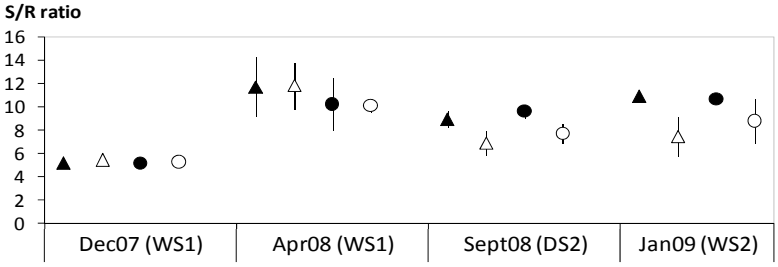


Figure 726 – S/R ratio repartition in the different genotypes and treatments calculated from estimated dry weights, during the first wet season (WS1), the second dry season (DS2), and the second wet season (WS2). S/R ratio are represented in triangles for the genotypes 1-41 and in circles for 18-50, irrigated treatment are in black, and non-irrigated in white. Vertical bars indicate standard errors.

3.7 Shoot/root ratios

At the beginning of WS1, S/R ratio were higher for 1-41 and in NI treatment (figure 7), but at the end of WS1, no effect of genotype or treatment was detected, suggesting that high rainfall during the wet season enabled trees to retrieve a common biomass partitioning independently from genotype or treatment.

Finally at the end of DS2 and beginning of WS2, S/R ratio were higher for genotype 18-50 suggesting that it could form more above-ground biomass at equal RDW compared to 1-41. As expected, trees from IR plots had a higher S/R ratio at DS2 and WS2, and thus seemed to allocate more biomass in above-ground compartment than in roots. Nevertheless, during these two periods increase in S/R ratio in IR seemed not due to the same processes: during DS2, it was mainly due to an high increase of root biomass in NI treatment, whereas during WS2, it was mainly due to an increase of stem biomass in IR treatments. GxT interaction effect was significant at WS2, indicating that 1-41 presented a higher decrease between IR and NIR (-31 %) compared to genotype 18-50 which was more stable between treatments (-17 % from IR to NI treatment). This result suggests that the two genotypes displayed differences in terms of response delay to recovery from seasonal water shortage.

3.8 Adjustment of physiological parameters

For the both genotypes, mean leaf areas (MLA) increased during the experiment until reach a maximum when trees were aged of 16 months (table 5). Nevertheless, MLA were highly different between genotypes whatever sampling campaign: 1-41 had larger leaves (67 dm^2 in mean at mature stage) compared to 18-50 (34 dm^2). MLA were larger in IR treatment at the beginning of the experimentation (September and December 2007), but not after, suggesting a high effect of ontogeny on leaf area during the first year of growth. Specific leaf area (SLA) were generally higher for genotype 18-50 (12.2 $m^2.kg^{-1}$ in mean on the whole experiment) compared to 1-41 (11.2 $m^2.kg^{-1}$; table 5, figure 8).

During the both dry seasons, SLA were higher for IR treatment compared to NIR treatment. Interestingly, during WS2 (January 2009), SLA were higher in NIR treatments, whatever genotypes (10.72 in NIR vs. 9.6 $m^2.kg^{-1}$ in IR treatment). GxT interaction effect on SLA was significant in April 2008 (WS1) and September 2008 (DS2): SLA decrease from IR to NI treatment was higher in 18-50 compared to 1-41.

Table 5 – Summary statistics of physiologic traits: Specific Leaf Area (SLA), Mean Leaf Area (MLA), carbon isotopic composition ($d^{13}C$), carbon composition (C%), stomatal density on lower face of the leaves (SD inf) and on upper face of the leaves (SDsup).

		Means ± SE						Explained variance (η2)		
date	Trait	1-41IR	1-41NIR	18-50IR	18-50NIR	1-41IR2	18-50IR2	G	T	GxT
09/2007 (DS1)	SLA	14.7 ^{a} ± 3	11.2 ^{a} ± 0.8	14.2 ^{a} ± 2.4	12.1 ^{a} ± 2			0.001	0.263 **	0.019
	MLA	29.2 ^{a} ± 3.8	20.5 ^{b} ± 2.9	16 ^{c} ± 2	10.7 ^{d} ± 1.5			0.632 ***	0.23 ***	0.014 *
	d13C	-29.7 ^{a} ± 0.4	-29.6 ^{a} ± 0.8	-29.5 ^{a} ± 0.4	-29 ^{a} ± 1.3			0.061	0.044	0.015
	C%	44 ^{a} ± 0.6	43.8 ^{a} ± 1.6	44 ^{a} ± 0.3	45.1 ^{a} ± 1.5			0.078	0.033	0.071
	SD inf	114 ^{a} ± 32	129 ^{a} ± 23	110 ^{a} ± 6	136 ^{a} ± 9			0.002	0.243	0.017
	SD sup	36 ^{a} ± 6	40 ^{a} ± 11	2 ^{b} ± 0	3 ^{b} ± 0			0.901***	0.006	0.003
12/2007 (WS1)	SLA	13.4 ^{b} ± 1.5	14.5 ^{b} ± 1.3	17.8 ^{a} ± 1.6	17.9 ^{a} ± 2.3			0.586 ***	0.013	0.009
	MLA	70.4 ^{a} ± 6.8	52 ^{b} ± 5.9	29.1 ^{c} ± 2.5	20.1 ^{c} ± 3			0.695 ***	0.097 ***	0.012 *
	SD inf	123 ^{a} ± 9	113 ^{a} ± 16	113 ^{a} ± 9	109 ^{a} ± 4			0.122	0.106	0.017
	SD sup	46 ^{a} ± 4	41 ^{a} ± 2	3 ^{b} ± 1	3 ^{b} ± 1			0.983***	0.002	0.004
04/2008 (WS1)	SLA	9.9 ^{b} ± 0.9	9.7 ^{b} ± 0.3	12.4 ^{a} ± 1.1	10.6 ^{b} ± 1			0.33 ***	0.131 **	0.078 *
	MLA	64.9 ^{a} ± 2.9	65.6 ^{a} ± 6.7	25.1 ^{b} ± 2.8	26.6 ^{b} ± 4.2			0.916 ***	0.001	0
	d13C	-30.8 ^{b} ± 0.3	-30.8 ^{b} ± 0.3	-29.4 ^{a} ± 0.5	-30.3 ^{b} ± 0.5			0.452***	0.112**	0.109**
	C%	46.7 ^{b} ± 4.3	47.2 ^{a,b} ± 1	47.6 ^{a,b} ± 2.1	50.1 ^{a} ± 1.3			0.13*	0.078	0.036
09/2008 (DS2)	SLA	10.5 ^{b,c} ± 0.8	9.5 ^{d} ± 0.5	11.8 ^{a} ± 0.7	9.8 ^{c,d} ± 0.9	10.4 ^{b,c,d} ± 0.5	11.3 ^{a,b} ± 0.9	0.133 **	0.422 ***	0.042 *
	MLA	69.4 ^{a} ± 6.9	67 ^{a} ± 5.7	35.2 ^{b} ± 2.5	32.5 ^{b} ± 2.5	66.6 ^{a} ± 4.7	35.1 ^{b} ± 2	0.931 ***	0.005	0
	d13C	-31.6 ^{c} ± 0.4	-30.6 ^{b} ± 0.5	-30.6 ^{b} ± 0.6	-29.5 ^{a} ± 0.6	-31.7 ^{c} ± 0.2	-30.8 ^{b} ± 0.6	0.33***	0.375***	0.001
	C%	51.3 ^{a} ± 3.2	49.6 ^{a,b} ± 1.5	47.5 ^{b,c} ± 1	47.7 ^{b,c} ± 1	48.6 ^{b,c} ± 1.2	46.3 ^{c} ± 2.3	0.308***	0.112**	0.03
	SD inf	122 ^{a} ± 4	115 ^{a} ± 3	130 ^{a} ± 20	113 ^{a} ± 3			0.015	0.271*	0.05
01/2009 (WS2)	SLA	9.7 ^{a} ± 0.5	10.5 ^{a} ± 0.7	9.6 ^{a} ± 1.1	10.9 ^{a} ± 1.8	9.7 ^{a} ± 0.7	10.6 ^{a} ± 2.4	0.007	0.181 **	0.008
	MLA	66.8 ^{a} ± 4.4	70.2 ^{a} ± 6.2	34 ^{b} ± 4.9	30.2 ^{b} ± 2.6	67 ^{a} ± 6.4	35.4 ^{b} ± 2.8	0.924 ***	0	0.009 *
	d13C	-30.4 ^{b} ± 0.7	-30.2 ^{b} ± 0.4	-29.2 ^{a} ± 0.6	-29.1 ^{a} ± 0.4	-30.3 ^{b} ± 0.3	-29.3 ^{a} ± 0.7	0.525***	0.008	0.002
	C%	48.6 ^{a} ± 1.2	48.3 ^{a,b} ± 0.6	47 ^{a,b} ± 1.8	47.4 ^{a,b} ± 1.1	47.3 ^{a,b} ± 1.8	46.7 ^{b} ± 1.1	0.127**	0.077	0.021

means ± SE : means and standard error associated; letters indicate classification according to Bonferroni pairwise test. Explained variance of each modeled effect : G (Genotype), T (Treatment), GxT (genotype by interaction effect). Stars indicate significance of the tested effects (* P<0.05; ** P<0.01; *** P<0.001). DS1 corresponds to the first dry season, WS1 to the first wet season, DS2 to the second dry season, and WS2 to the second wet season.

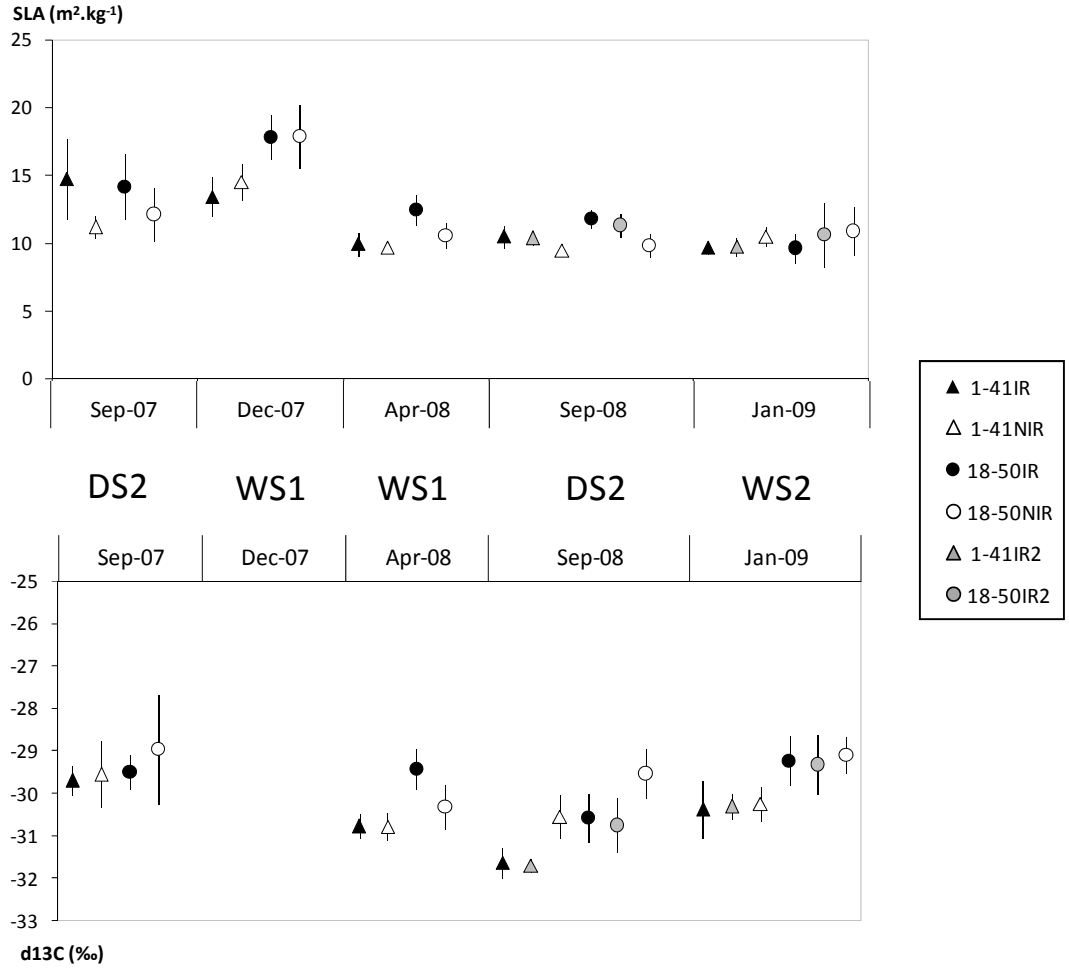


Figure 8 – SLA and d13C measured at the different sampling campaigns and on the recovery plots. DS1 corresponds to the first dry season, WS1 to the first wet season, DS2 to the second dry season, and WS2 to the second wet season. Vertical bars indicate standard errors.

Carbon isotope composition analysis revealed also differences between genotypes (table 5, figure 8) : from April 2008 (WS1), genotype 18-50 had higher $\delta^{13}C$ value compared to 1-41 (−29.8 ‰ vs. -30.8 ‰). Nevertheless, $\delta^{13}C$ was also influenced by environment: in September 2008 (end of DS2), $\delta^{13}C$ was higher in NIR treatments compared to IR treatments. Interestingly, in April 2008 (WS1), $\delta^{13}C$ was higher in IR treatment compared to NI treatment for genotype 18-50, while no change was observed for 1-41, revealing GxT interaction effect.

Finally, concerning stomatal density (SD), we confirmed first observations done on mature trees leaves: genotype 1-41 had stomata on the both adaxial and abaxial surfaces of leaves (in average, 41 stomata.mm^{-2} on adaxial surfaces and 119 stomata.mm^{-2} on abaxial surfaces), while 18-50 was hypostomatic (in average, 3 stomata.mm^{-2} on adaxial surfaces and 119 stomata.mm^{-2} on abaxial surfaces). On abaxial surfaces of the leaves, no differences in SD were found between genotypes, neither between samplings. Nevertheless, treatment effect

was significant in September 2008 (end of DS2): leaves from IR treatments presented higher SD compared to NI treatments (125 vs 114 stomata.mm^{-2}, respectively).

3.9 Recovery plots

Recovery plots (P5' and P6') were only irrigated during the second dry season. In July 2008, IR2 plots had globally a lower growth compared to IR and NI plots (e.g. VOB of 7.5, 9.1 and 12.1 dm^3, for IR2, NI and IR plots respectively; table 3). In September 2008 (end of DS2), trees from recovery plots presented growth traits similar to trees from NI treatment whatever the genotype (VOB of 11.8 and 11.7 dm^3 for 18-50 subjected respectively to IR2 and NI treatment, and VOB of 9.1 and 9.3dm^3, similarly for genotype 1-41). In January 2009 (middle of WS2), no differences were found between IR and IR2 treatments for genotype 18-50 (25.1 and 22.0 dm^3 respectively), while genotype 1-41 showed lower growth on IR2 treatment compared to IR treatment (16.7 and 25.4 dm^3 respectively).

In September 2008, SDW in IR2 treatment were not significantly different from the other treatments for 1-41, whereas SDW were lower than in NI and therefore in IR treatments (table 4). Concerning biomass partitioning, S/R ratio of 18-50 in IR2 treatment was similar to 18-50IR and higher compared to 18-50NI, while for genotype 1-41, S/R ratio was significantly higher compared to IR and NI treatments. In January 2009 (WS2), SDW of recovery plots were similar to NIR treatments whatever the genotype, and lower than IR treatments (table 4). S/R ratio was similar between treatments for 18-50, while S/R ratio for 1-41 in IR2 was similar to IR but higher than NIR treatment.

Finally, leaves parameters measured in September 2008 showed that trees from recovery plots had physiologic parameters similar to trees from IR plots: $\delta^{13}C$ and SLA were higher in these plots compared to NIR plots (figure 8). This suggests that this two parameters were highly responsive to watering, and not time integrated such growth parameters.

4 Discussion

4.1 Experimental trial capacity

Establishment of trials to study trees phenotypic plasticity has encountered several difficulties to fulfill. Theoretically, the definition of phenotypic plasticity induced the use of genetically similar individuals such clones (Valladares *et al.*, 2006). Nevertheless, for several forest species, propagation by cuttings is not plenty controlled, and the genetic unit studied is

dependant of the characteristics of the species of interest, often constituted of related individuals such full-sibs or populations (Chambel *et al.*, 2005). In those cases, comparisons of phenotypic plasticity across species could suffer from problems due to the extent of uncontrolled genetic variability (Valladares *et al.*, 2006). In our study, we tried to cope with these difficulties by installing more than 1000 coppices of the same 2 genotypes, known to have different growth performance and physiological properties. Samplings were done to be as representative as possible of the mean phenotype, by selecting individual trees close to the median height of the considered plot (buffer trees excluded).

Moreover, in phenotypic plasticity studies, the ecological gradient is often simulated under greenhouses with controlled conditions, by manipulation of one or few parameters. These experiments were conducted at short-term, and seemed maladapted for tall and long-lived organisms such forest trees. An alternative consisted in the establishment of multi-locality common gardens, where environmental parameters were not controlled, but measured. This type of field trials was more representative of natural conditions, but analysis is complicated by the fact that environmental variation were multiples (e.g. soil properties, climate parameters). Our study was installed in a field experiment of 1.8 ha, where soils were homogenous, and where the only environmental parameter variation was soil moisture. Despite this, it still seemed that some micro-environmental variations had influenced growth during the first wet season (P5 and P6 had higher growth compared to P4, P8 and P9), inducing difficulties to conclude about biomass evolution. Despite the fact that eucalyptus are known to develop quickly deep-root (Christina *et al.*, 2011), and that it was possible that roots could reach water table (located at 8m depth) before the second dry season experienced, physiological traits measurements such Specific Leaf area (SLA) and carbon isotopic composition ($\delta^{13}C$) were significantly different between treatments at the end of DS2, ensuring that trees experienced water scarcity in NIR treatments. Consequently, we could expect that this trial reflect as well as possible industrial conditions in NIR treatments, and then is fully adapted to the study of phenotypic plasticity of eucalyptus in response to seasonal drought.

4.2 Effect of seasons on growth and physiological parameters

In this study, when no irrigation was performed during the dry season, growth was characterized by contrasted seasonal patterns. As DS1 corresponded to the first months of growth, seasonal decrease of growth was not obvious, nevertheless seasonal effects could be observed on DS2. Effects of dry season on tree morphology were immediate: daily increment

in height, circumference and volume over-bark reduced directly from DS2 (between July 2008 and January 2009). It seemed that effects of water shortage were observed latter on daily increment in SDW which were reduced during WS2 (between September 2008 and January 2009), especially for genotype 18-50. When irrigation was applied, tree growth was more constant, as suggested by constantly rising daily increments of SDW and VOB.

On the other hand, leaves component seemed to present more sensitivity to the seasonal variations. First, reduction of daily increment in LDW affected trees independently from treatment during DS2. This decrease continued at the beginning of WS2, becoming a loss of leaf biomass in NIR and IR2 treatments. As leaf biomass increment resulted from leaf production and senescence, the loss of leaf biomass must be induced by a higher leaf-shedding rate compare to leaf production rate. Leaf senescence or shedding has been widely described as an adaptive process to drought avoidance, by decreasing surface area for transpiration, and preserving moisture in perennial tissues (Kozlowski, 1976). Nevertheless, here, it seems that leaf shedding was independent from soil moisture, but could be related to developmental processes. Litterfall measurement revealed that proportion of leaf litter related to LDW were higher in 1-41 subjected to NIR treatments in June and July 2008, suggesting leaf senescence due to water shortage, but were higher in IR treatments in September and October 2008. In this case, it is possible that trees from IR plots began pruning, in a more intense way compared to trees from NIR plots. Indeed, in Congolese eucalyptus plantations, Laclau (2001) has shown that leaves begin to fall around the 10th month of growth, and litterfall are more important at the beginning of the wet season (October-November) and the dry season (May to July).

Nouvellon *et al.* (2009) have already described seasonal variation of SLA for *Eucalyptus* species in Rep. of Congo. They showed that SLA was higher at the height of the wet season, in January, and SLA was lower at the end of the dry season, in October. Regarding at the same growing conditions (i.e. NIR treatment), we found also the same trends, with a contrasted pattern during the first growing season (comparing SLA in September and December 2007), and at a lesser extent during the second growing season (September 2008 and January 2009). SLA in December 2007 was particularly high (14 and 17.8 $m^2.kg^{-1}$ for 1-41 and 18-50) compared to other published data, that could be related to intense precipitations that occurred at the beginning of WS1. In this study we also highlight seasonal variation of $\delta^{13}C$ in sun leaves: carbon isotopic composition was very low at the end of DS1, and high during the dry seasons. Interestingly, $\delta^{13}C$ in January 2009 was high for all the trees measured

suggesting that water supply could be a limiting factor during the beginning of the WS2. Finally, the fact that SLA remained low and $\delta^{13}C$ was high in January 2009, in addition to the observation of low rainfall could confirm that the beginning of WS2 didn't provide enough water for ensure a high leaf production, that could partly explain the low increment of LDW during this period.

4.3 Plasticity of the two genotypes

Plasticity of growth parameters seemed to be relatively similar between genotypes (figure 9), but plasticity index (PI) displayed seasonal trends: excepted for HT, PI decreased at the end of WS1 and during DS2. Growth (HT, VOB) and biomass (SDW) increments were higher during WS1 for 18-50 whatever the treatment, and moreover, during WS2, growth on recovery plots were higher for genotype 18-50 compared to genotype 1-41: 18-50IR2 trees had similar growth parameters to 18-50IR trees, while 1-41IR2 trees had similar growth parameters to 1-41NI trees. This result suggests that if the both genotypes decreased growth similarly during drought seasons, the differences between genotypes could reside primarily in ability to recover from seasonal water deficit and to benefit from favorable conditions.

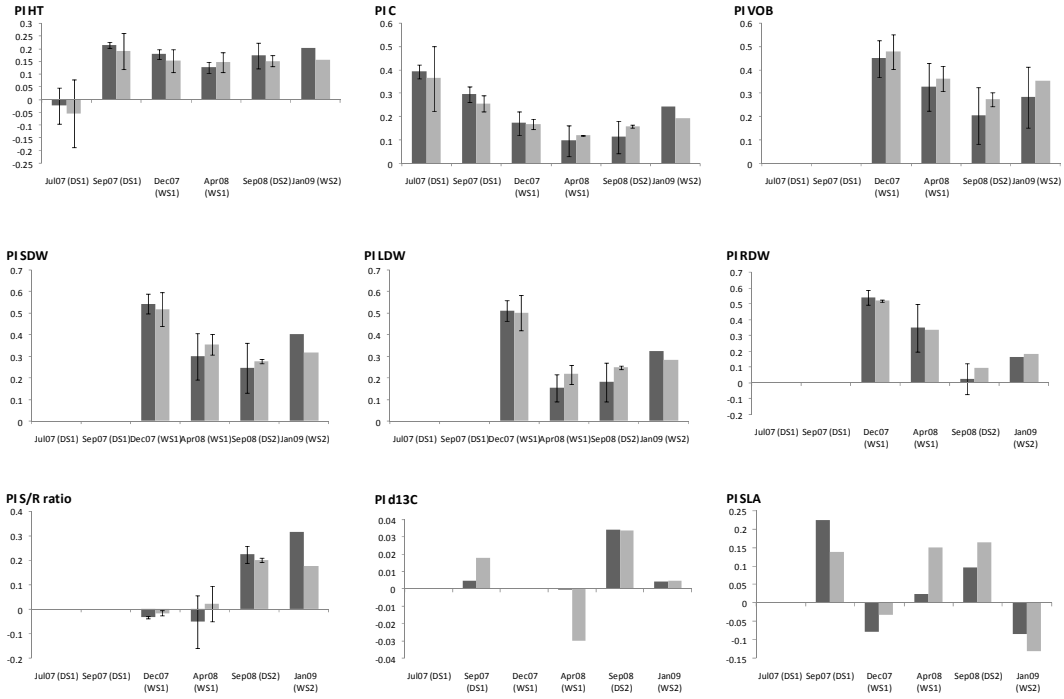


Figure 9 – Plasticity index computed for growth, biomass and physiological traits analyzed. PI for HT, C, VOB, SDW, LDW, RDW and S/R ratio were computed by plot (with simulated datas for DW). Means are indicated with standard errors (vertical bars) when possible.

Comparing two *Eucalyptus* hybrids, Drew (2007) defined two strategies to describe their growth properties. One hybrid had a regular growth but accumulates late in ontogenic trajectories (tortoise strategy), while the other benefit from favorable conditions with a discontinuous growth (hare strategy). Here, the both genotypes showed a discontinuous growth, but genotype 1-41 seems to accumulate late in ontogenesis compared to 18-50.

These differences between the two hybrids could be explained by several hypotheses. First, on *Eucalyptus globulus*, Drew *et al.* (2008) suggested that immediate benefits from irrigation on stem diameter increment could be due to the high responsiveness of cambium and xylem developmental zones to water status. We could hypothesize that the two hybrids doesn't display the same properties concerning stem tissues sensitivity to water availability. Moreover, as suggested by seasonal patterns, physiological traits seemed to be more plastic than growth traits. The both genotypes showed high differences in carbon isotope composition, which suggests that the genotypes differed in water use efficiency (Farquhar and Richards, 1984; Olbrich *et al.*, 1993). Concerning leaf morphology, genotype 18-50 presented higher PI in MLA by highly reducing leaf area in NIR treatment, during DS1, beginning of WS1 and WS2. SLA also showed higher plasticity for genotype 18-50. High SLA has been shown to be related to high photosynthetic capacity (Lusk *et al.*, 2003). These results suggest that genotypes present differences in ability to retrieve high photosynthetic capacity when environmental condition became favorable. Finally, the both genotypes seemed to differ in strategy of biomass partitioning: first, PI of S/R ratio varied along the experiment (figure 9): genotype 18-50 was highly plastic for this trait during WS1, while genotype 1-41 was more plastic at the beginning of WS2. Moreover, from DS2, S/R ratio were higher for genotype 18-50 suggesting that it could form more above-ground biomass at equal RDW compared to 1-41.

5 Conclusions

This study has shown that differences of productivity of the two genotypes studied reside primarily in efficiency differences between the two genotypes: the more productive genotype (*E. urophylla* x *E. grandis*) produce generally more above-ground biomass for a same below-ground biomass. Carbon isotope composition analysis confirms this trend, suggesting that this genotype use less water for the same amount of fixed carbon. If differences in productivity between genotypes could hardly be explained by differences in growth plasticity, it seems that it could be more related to the large ability of the genotype 18-50 to recover from water

shortage and to use favorable conditions to increase above-ground biomass. Physiologically, responses to drought seem to be more instantaneous, and differences in plasticity were larger, suggesting that trees developed a “drought avoidance” strategy. More investigations are needed to have a better understanding of which mechanisms are involved by the genotype 18-50 to recover quickly from dry season. It seemed that resources allocation processes and leaf biomass management may reflect key mechanisms involved in this "drought avoidance' strategy.

SUPPLEMENTAL FILE 1 : BIOMASS PREDICTIONS

1 Methods

For each sampling campaign, we compared 4 different linear models establishing prediction of dry weight (DW) of each compartment to analyze (stem, leaf or roots), in function of circumference at 1.3m (C) and total tree height (HT), as described by Saint-André *et al.* (2005). The studied models are the following:

Global model: $DW = a + b \times C^2HT + \varepsilon$

a, b, are the model parameters to be estimated and ε the residual error.

Local G model: $DW = a_{1\text{-}41} + b_{1\text{-}41} \times C^2HT + a_{18\text{-}50} + b_{18\text{-}50} \times C^2HT + \varepsilon$

where $a_{1\text{-}41}$, $b_{1\text{-}41}$ are model parameters specific to genotype 1-41, $a_{18\text{-}50}$, $b_{18\text{-}50}$ are model parameters specific to genotype 18-50, and ε the residual error.

Local T model: $DW = a_{IR} + b_{IR} \times C^2HT + a_{NIR} + b_{NIR} \times C^2HT + \varepsilon$

where a_{IR}, b_{IR} are model parameters specific to IR treatment, a_{NIR}, b_{NIR} are model parameters specific to NIR treatment, and ε the residual error.

Local GxT model: $DW = a_{1\text{-}41IR} + b_{1\text{-}41IR} \times C^2HT + a_{1\text{-}41NI} + b_{1\text{-}41NI} \times C^2HT + a_{18\text{-}50IR} + b_{18\text{-}50IR} \times C^2HT + a_{18\text{-}50NI} + b_{18\text{-}50NI} \times C^2HT + \varepsilon$

where $a_{1\text{-}41IR}$, $b_{1\text{-}41IR}$ are model parameters specific to genotype 1-41 subjected to IR treatment, $a_{1\text{-}41NI}$, $b_{1\text{-}41NI}$ are model parameters specific to genotype 1-41 subjected to NI treatment, $a_{18\text{-}50IR}$, $b_{18\text{-}50IR}$ are model parameters specific to genotype 18-50 subjected to IR treatment, $a_{18\text{-}50NI}$, $b_{18\text{-}50NI}$ are model parameters specific to genotype 18-50 subjected to NI treatment, and ε the residual error.

For all the equations, we fit simultaneously the model for the mean (as described above) and the model for the variance (variance = $\sigma^2 \times C^2HT^{2k}$, where σ^2 and k were estimated parameters), using the proc NLP of SAS software (package OR). Parameters non significant from zero were removed, and the equation was refitted. Maximum likelihood (ML) was estimated for each model. Comparisons between models were performed using the Akaike Information Criterion (AIC = 2ML + 2p) and the Bayesian Information Criterion (BIC = 2ML

+ p log(n)) where p was the number of parameters of the model and n the number of observations (Akaike, 1973; Schwarz, 1978). When the two criteria led to different results, we chose the model with the fewest parameters.

2 Results

The nature of best models to predict biomasses in function of C and HT could inform on the effects acting on biomass partitioning (table S1, figure S1).

Thus, for SDW prediction, global models were the best for samplings done in December 2007 and April 2008 (beginning and end of WS1), indicating that whatever the genotype and the treatment partition of biomass into stem was similar. In September 2008 (DS2), local model with parameters specific to treatment was the best, indicating that for the same volume, trees from NIR treatments displayed larger SDW. In January 2009 (WS2), the best model was a local model with specific parameters for GxT, indicating that trees from NIR treatment displayed larger SDW for the same volume, and differences in biomass equations between treatments were more important for genotype 18-50 than for 1-41. Similarly, for LDW, global models were selected for December 2007, April and September 2008, indicating that no differences in biomass partitioning could be detected between genotypes or treatment. In January 2009 (WS2), the same trends were observed than for SDW, local GxT model with higher partitioning for NIR, especially 18-50 NIR. Finally, partitioning in RDW show different patterns between dry season and wet seasons: for a same volume, trees from IR plots had a larger RDW during WS1 and WS2, but during DS2 trees from NIR plots had a larger RDW. These results suggest that trees could switch resources allocation to root system when water supply is limited.

Table S 1 – Final weighted equations for dry weight of each compartment (stem, leaf and root), sampling campaign by sampling campaign.

			Mean estimation						
Variable	Sampling Campaign	Model	1-41IR	1-41NI	18-50IR	18-50NI	Error term	R^2	RMSE
SDW	SC3	Global	$\mu = 0.04 + (33.48\ C^2HT)$	$\mu = 0.04 + (33.48\ C^2HT)$	$\mu = 0.04 + (33.48\ C^2HT)$	$\mu = 0.04 + (33.48\ C^2HT)$	$e = 1.72(C^2HT)0.61$	0.904	0.040
	SC4	Global	$\mu = 0.92 + (37.82\ C^2HT)$	$\mu = 0.92 + (37.82\ C^2HT)$	$\mu = 0.92 + (37.82\ C^2HT)$	$\mu = 0.92 + (37.82\ C^2HT)$	$e = 87.51(C^2HT)1.12$	0.863	0.387
	SC5	Local T	$\mu = 2.39 + (19.99\ C^2HT)$	$\mu = 0.97 + (26.83\ C^2HT)$	$\mu = 2.39 + (19.99\ C^2HT)$	$\mu = 0.97 + (26.83\ C^2HT)$	$e = 4.84(C^2HT)0.94$	0.891	0.530
	SC6	Local GxT	$\mu = 0 + (23.18\ C^2HT)$	$\mu = 0 + (29.51\ C^2HT)$	$\mu = 0 + (27.49\ C^2HT)$	$\mu = 0 + (34.24\ C^2HT)$	$e = 5.65(C^2HT)1.19$	0.934	0.895
LDW	SC3	Global	$\mu = 0.07 + (26.26\ C^2HT)$	$\mu = 0.07 + (26.26\ C^2HT)$	$\mu = 0.07 + (26.26\ C^2HT)$	$\mu = 0.07 + (26.26\ C^2HT)$	$e = 0.44(C^2HT)0.46$	0.839	0.039
	SC4	Global	$\mu = 1.07 + (8.54\ C^2HT)$	$\mu = 1.07 + (8.54\ C^2HT)$	$\mu = 1.07 + (8.54\ C^2HT)$	$\mu = 1.07 + (8.54\ C^2HT)$	$e = 0.64(C^2HT)0.31$	0.353	0.253
	SC5	Global	$\mu = 0.99 + (3.82\ C^2HT)$	$\mu = 0.99 + (3.82\ C^2HT)$	$\mu = 0.99 + (3.82\ C^2HT)$	$\mu = 0.99 + (3.82\ C^2HT)$	$e = 0.44(C^2HT)0.74$	0.606	0.195
	SC6	Local GxT	$\mu = 0 + (3.45\ C^2HT)$	$\mu = 0 + (4.97\ C^2HT)$	$\mu = 0 + (4.29\ C^2HT)$	$\mu = 0 + (5.59\ C^2HT)$	$e = 0.22(C^2HT)0.32$	0.726	0.290
RDW	SC3	Global	$\mu = 0.02 + (11.89\ C^2HT)$	$\mu = 0.02 + (11.89\ C^2HT)$	$\mu = 0.02 + (11.89\ C^2HT)$	$\mu = 0.02 + (11.89\ C^2HT)$	$e = 0.8(C^2HT)0.69$	0.834	0.018
	SC4	Local GxT	$\mu = 0 + (6.73\ C^2HT)$	$\mu = 0.27 + (1.25\ C^2HT)$	$\mu = 0 + (6.47\ C^2HT)$	$\mu = 0.16 + (5.34\ C^2HT)$	$e = 421.85(C^2HT)2.17$	0.701	0.050
	SC5	Local T	$\mu = 0.61 + (1.87\ C^2HT)$	$\mu = 0.67 + (2.34\ C^2HT)$	$\mu = 0.61 + (1.87\ C^2HT)$	$\mu = 0.67 + (2.34\ C^2HT)$	$e = 0.05(C^2HT)0.42$	0.625	0.073
	SC6	Local GxT	$\mu = 0 + (2.45\ C^2HT)$	$\mu = 0.48 + (2.79\ C^2HT)$	$\mu = 0 + (2.99\ C^2HT)$	$\mu = 1.12 + (1.23\ C^2HT)$	$e = 0.05(C^2HT)0.32$	0.813	0.084

Units : C and H are in meter, RMSE in kg/tree, R^2 is dimensionless.

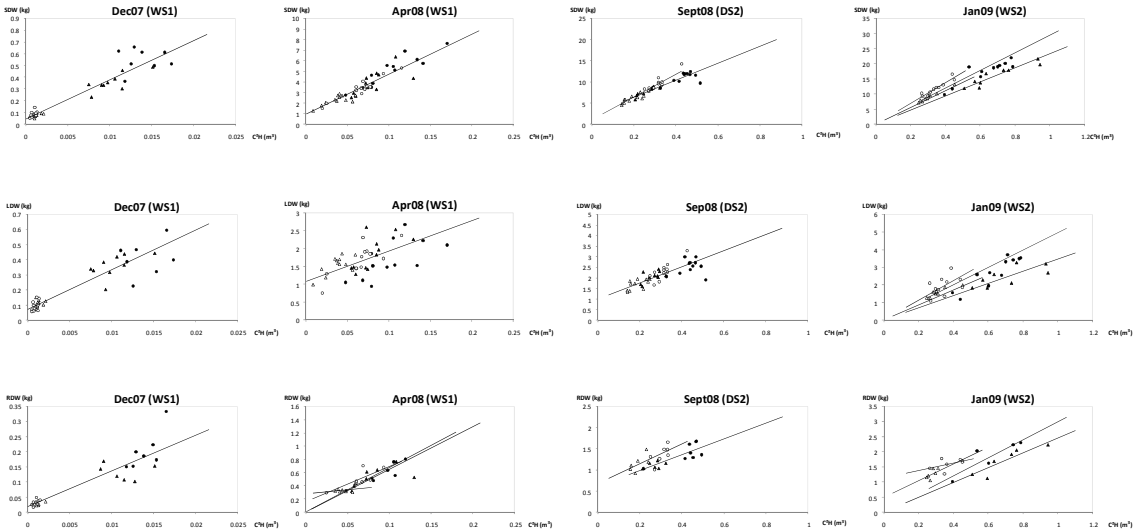


Figure S 1 – Biomass equations and measured sampled used to construct the equations for stem (SDW), leaf (LDW) and root (RDW) at 4 sampling campaigns.

REFERENCES

Akaike, H. (1973). "Information theory and an extension of the maximum likelihood principle." 2nd International Symposium on Information Theory.

Allen, C. D., A. K. Macalady, et al. (2010). "A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests." Forest Ecology and Management **259**(4): 660-684.

Bonan, G. B. (2008). "Forests and climate change: forcings, feedbacks, and the climate benefits of forests." Science (Washington) **320**(5882): 1444-1449.

Bouillet, J.-P., J.-P. Laclau, et al. (2002). "Changes with age in the spatial distribution of roots of Eucalyptus clone in Congo: Impact on water and nutrient uptake." Forest Ecology and Management **171**(1-2): 43-57.

Bouvet, J. M. and P. Vigneron (1995). "Age trends in variances and heritabilities in Eucalyptus factorial mating designs." Silvae Genetica **44**(4): 206-216.

Breda, N., R. Huc, et al. (2006). "Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences." Annals of Forest Science **63**(6): 625-644.

Brunet, D., B. G. Barthès, et al. (2007). "Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity." Geoderma **139**(1-2): 106-117.

Campion, J. M., M. Nkosana, et al. (2006). "Biomass and N and P pools in above- and below-ground components of an irrigated and fertilised Eucalyptus grandis stand in South Africa." Australian Forestry **69**(1): 48-57.

Ceulemans, R., L. Vanpraet, et al. (1995). "EFFECTS OF CO2 ENRICHMENT, LEAF POSITION AND CLONE ON STOMATAL INDEX AND EPIDERMAL-CELL DENSITY IN POPLAR (POPULUS)." New Phytologist **131**(1): 99-107.

Chambel, M. R., J. Climent, et al. (2005). "Phenotypic plasticity: a useful framework for understanding adaptation in forest species." Investigacion Agraria, Sistemas y Recursos Forestales **14**(3): 334-344.

Cha-um, S. and C. Kirdmanee (2010). "Effects of water stress induced by sodium chloride and mannitol on proline accumulation, photosynthetic abilities and growth characters of eucalyptus (<i>Eucalyptus camaldulensis</i> Dehnh.)." New Forests **40**(3): 349-360.

Christina, M., J. P. Laclau, et al. (2011). "Almost symmetrical vertical growth rates above and below ground in one of the world's most productive forests." Ecosphere **2**(3): art27.

Craven, D., D. Dent, et al. (2010). "Seasonal variability of photosynthetic characteristics influences growth of eight tropical tree species at two sites with contrasting precipitation in Panama." Forest Ecology and Management **In Press, Corrected Proof**.

d'Annunzio, R., S. Conche, et al. (2008). "Pairwise comparison of soil organic particle-size distributions in native savannas and Eucalyptus plantations in Congo." Forest Ecology and Management **255**(3-4): 1050-1056.

de Dieu Nzila, J., J.-P. Bouillet, et al. (2002). "The effects of slash management on nutrient cycling and tree growth in Eucalyptus plantations in the Congo." Forest Ecology and Management **171**(1-2): 209-221.

Delagrange, S. (2010). "Light- and seasonal-induced plasticity in leaf morphology, N partitioning and photosynthetic capacity of two temperate deciduous species." Environmental and Experimental Botany **70**(1): 1-10.

DeWitt, T. J. (1998). "Costs and limits of phenotypic plasticity: Tests with predator-induced morphology and life history in a freshwater snail." Journal of Evolutionary Biology **11**(4): 465-480.

Drake, P. L., D. S. Mendham, et al. (2009). "A comparison of growth, photosynthetic capacity and water stress in Eucalyptus globulus coppice regrowth and seedlings during early development." Tree Physiology **29**(5): 663-674.

Drew, D. M., A. P. O'Grady, et al. (2008). "Daily patterns of stem size variation in irrigated and unirrigated Eucalyptus globulus." Tree Physiology **28**(10): 1573-1581.

Drew, D. M. and N. W. Pammenter (2007). "Developmental rates and morphological properties of fibres in two eucalypt clones at sites differing in water availability." Southern Hemisphere Forestry Journal **69**(2): 71-79.

Epron, D., Y. Nouvellon, et al. (2004). "Spatial and temporal variations of soil respiration in a Eucalyptus plantation in Congo." Forest Ecology and Management **202**(1-3): 149-160.

Farquhar, G. D. and R. A. Richards (1984). "ISOTOPIC COMPOSITION OF PLANT CARBON CORRELATES WITH WATER-USE EFFICIENCY OF WHEAT GENOTYPES." Australian Journal of Plant Physiology **11**(6): 539-552.

Genuchten, M. T. v. (1980). "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils." Soil Science Society of America Journal **44**(5): 892-898.

Gideon, S. (1978). "Estimating the Dimension of a Model." The Annals of Statistics **6**(2): 461-464.

Gomat, H., P. Deleporte, et al. (2010). "What factors influence the stem taper of <i>Eucalyptus</i>: growth, environmental conditions, or genetics?" Annals of Forest Science: 1-12.

Laclau, J. P., J. P. Bouillet, et al. (2001). "Dynamics of nutrient translocation in stemwood across an age series of a eucalyptus hybrid." Annals of Botany **88**(6): 1079-1092.

Laclau, J. P., W. Sama-Poumba, et al. (2002). "Biomass and nutrient dynamics in a littoral savanna subjected to annual fires in Congo." Acta Oecologica-International Journal of Ecology **23**(1): 41-50.

Lusk, C. H., I. Wright, et al. (2003). "Photosynthetic differences contribute to competitive advantage of evergreen angiosperm trees over evergreen conifers in productive habitats." New Phytologist **160**(2): 329-336.

Mareschal, L., J. D. D. Nzila, et al. (2010). "Mineralogical and physico-chemical properties of Ferralic Arenosols derived from unconsolidated Plio-Pleistocenic deposits in the coastal plains of Congo." Geoderma **162**(1-2): 159-170.

Matesanz, S., E. Gianoli, et al. (2010). "Global change and the evolution of phenotypic plasticity in plants." Year in Evolutionary Biology **1206**: 35-55.

M'Bou, A. T., C. Jourdan, et al. (2008). "Root elongation in tropical Eucalyptus plantations: effect of soil water content." Annals of Forest Science **65**(6): 609p1-609p7.

Merchant, A., A. Callister, et al. (2007). "Contrasting physiological responses of six Eucalyptus species to water deficit." Annals of Botany **100**: 1507-1515.

Monclus, R., E. Dreyer, et al. (2005). "Productivity, leaf traits and carbon isotope discrimination in 29 Populus deltoides x P-nigra clones." New Phytologist **167**(1): 53-62.

Monclus, R., E. Dreyer, et al. (2006). "Impact of drought on productivity and water use efficiency in 29 genotypes of Populus deltoides x Populus nigra." New Phytologist **169**(4): 765-777.

Nouvellon, Y., J.-P. Laclau, et al. (2010). "Within-stand and seasonal variations of specific leaf area in a clonal Eucalyptus plantation in the Republic of Congo." Forest Ecology and Management **259**(9): 1796-1807.

Olbrich, B. W., D. Leroux, et al. (1993). "VARIATION IN WATER-USE EFFICIENCY AND DELTA-C-13 LEVELS IN EUCALYPTUS-GRANDIS CLONES." Journal of Hydrology **150**(2-4): 615-633.

Pereira, J. S. and T. T. Kozlowski (1976). "LEAF ANATOMY AND WATER RELATIONS OF EUCALYPTUS-CAMALDULENSIS AND E-GLOBULUS SEEDLINGS." Canadian Journal of Botany-Revue Canadienne De Botanique **54**(24): 2868-2880.

Pigliucci, M. (2005). "Evolution of phenotypic plasticity: where are we going now?" Trends in Ecology & Evolution **20**(9): 481-486.

Rozendaal, D. M. A., V. H. Hurtado, et al. (2006). "Plasticity in leaf traits of 38 tropical tree species in response to light; relationships with light demand and adult stature." Functional Ecology **20**(2): 207-216.

Safou-Matondo, R., P. Deleporte, et al. (2005). "Hybrid and clonal variability of nutrient content and nutrient use efficiency in Eucalyptus stands in Congo." Forest Ecology and Management **210**(1-3): 193-204.

Saint-André, L., A. T. M'Bou, et al. (2005). "Age-related equations for above- and below-ground biomass of a Eucalyptus hybrid in Congo." Forest Ecology and Management **205**(1-3): 199-214.

Stape, J. L., D. Binkley, et al. (2010). "The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production." Forest Ecology and Management **259**(9): 1684-1694.

Sultan, S. E. (2000). "Phenotypic plasticity for plant development, function and life history." Trends in Plant Science **5**(12): 537-542.

Thuiller, W., C. Albert, et al. (2008). "Predicting global change impacts on plant species' distributions: Future challenges." Perspectives in Plant Ecology, Evolution and Systematics **9**(3-4): 137-152.

Trouve, C., A. Mariotti, et al. (1994). "Soil organic carbon dynamics under Eucalyptus and Pinus planted on savannas in the Congo." Soil Biology and Biochemistry **26**(2): 287-295.

Valladares, F., E. Gianoli, et al. (2007). "Ecological limits to plant phenotypic plasticity." New Phytologist **176**: 749-763.

Valladares, F., D. Sanchez-Gomez, et al. (2006). "Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications." Journal of Ecology **94**(6): 1103-1116.

Warren, C., T. Bleby, et al. (2007). "Changes in gas exchange versus leaf solutes as a means to cope with summer drought in <i>Eucalyptus marginata</i&gt." Oecologia **154**(1): 1-10.

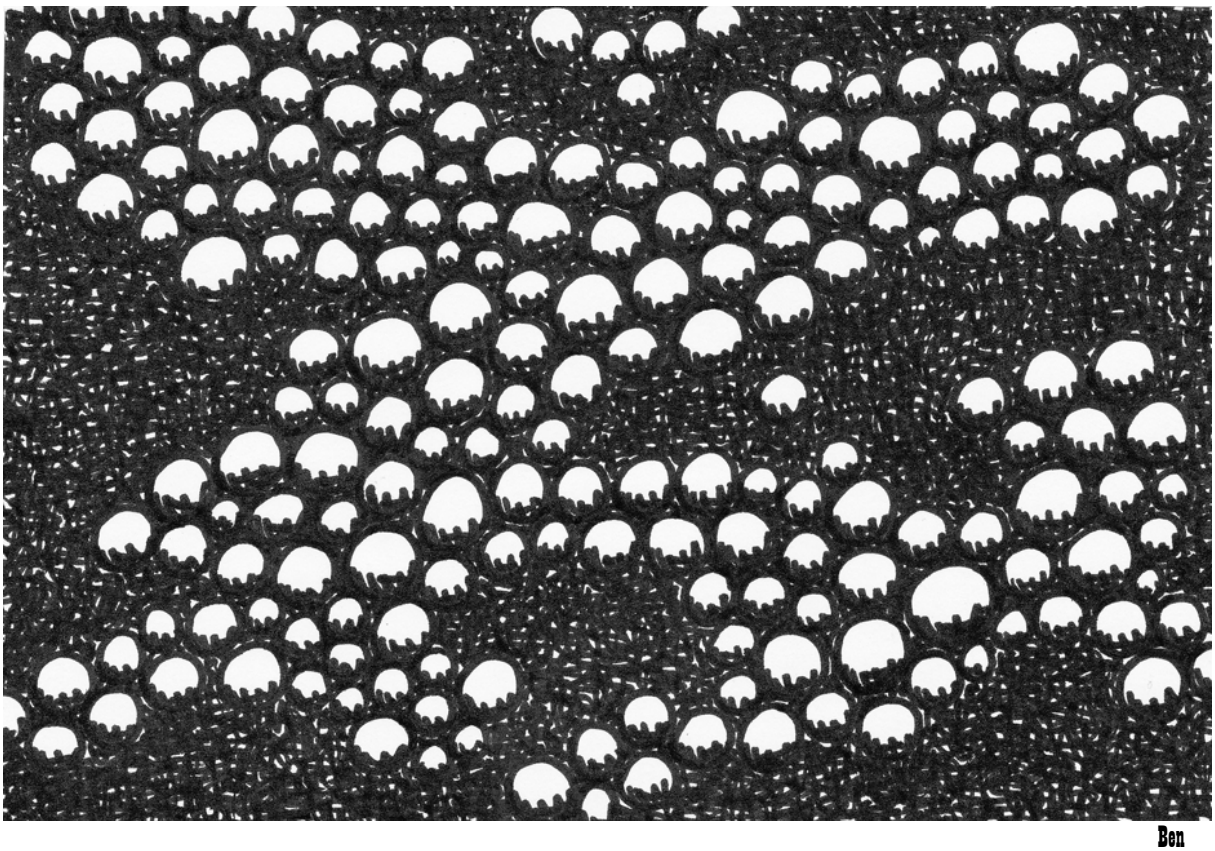
West-Eberhard, M. J. (2003). Developmental Plasticity and Evolution. New York.

Whitehead, D. and C. L. Beadle (2004). "Physiological regulation of productivity and water use in Eucalyptus: a review." Forest Ecology and Management **193**(1-2): 113-140.

Wright, S. D. and K. D. M. McConnaughay (2002). "Interpreting phenotypic plasticity: The importance of ontogeny." Plant Species Biology **17**(2-3): 119-131.

Wu, I.-P. (1997). "A Simple Evapotranspiration Model for Hawaii: The Hargreaves Model." Engineer's Notebook; EN-106.

Chapitre IV - Analyse de la plasticité foliaire par une approche intégrative



The multi-scale plasticity of leaves seems to influence the growth of two *Eucalyptus* genotypes subjected to water deficit

Emilie Villar[1,2,3], Andréas Ndeko[3], André Mabiala[3], Jean-Luc Verdeil[4], Christian Mertz[5], Christophe Plomion[2,6], Christophe Jourdan[7], Jean-Marc Gion[1§]

[1]CIRAD, UPR39 Genetic Diversity and Breeding of Forest Tree Species, Campus de Baillarguet TA 10C, F–34398 Montpellier Cedex 5, France

[2]Université de Bordeaux, UMR1202 BIOGECO, F-33610 Cestas, France

[3]CRDPI, BP1291, Pointe Noire, République du Congo

[4]CIRAD, UMR DAP, TA A-96 / 02 - Avenue Agropolis - 34398 Montpellier Cedex 5, France

[5]CIRAD, UMR QUALISUD, Bâtiment 16, 73 rue Jean-François Breton - 34398 Montpellier Cedex 5, France

[6]INRA, UMR1202 BIOGECO, F-33610 Cestas, France

[7]CIRAD, UPR80 s/c UMR Eco & Sols - INRA -Sup Agro - 2 Place Viala Bât 12 - 34060 Montpellier Cedex 02, France

[§] Corresponding author

E-mail addresses:

EV: villar@cirad.fr

AN: andreasndeko@yahoo.fr

AM: mabialangoma_andre@yahoo.fr

JLV: verdeil@cirad.fr

CM: christian.mertz@cirad.fr

CJ: jourdan@cirad.fr

CP: plomion@pierroton.inra.fr

JMG: gion@cirad.fr

1 Background

Phenotypic plasticity, is defined as the capacity of a genotype to adjust its phenotype in response to an environmental change (Bradshaw, 1965; Schlichting and Pigliucci, 1998). This component of adaptation is particularly important for sessile organisms and long-lived organisms, such as trees (Rehfeldt *et al.*, 2001; Grulke, 2010) because it may enable them to cope with the predicted climate changes with a single generation. There has been increasing interest in characterization of the phenotypic plasticity displayed by trees, and various degrees of plasticity have been reported, depending on the traits and environmental changes studied. Phenotypic plasticity has been detected for numerous traits in response to light constraints (Valladares *et al.,* 2008; Delagrange, 2011), temperature gradients (Vitasse *et al.*, 2010), nutrient scarcity (Sultana *et al.*, 2010; Sampedro *et al.*, 2010) and water stress (Ramirez-Valiente *et al.,* 2010; Trentacoste *et al.*, 2011). These analyses revealed that several genetic units (genotypes or populations) subjected to a common set of environmental conditions displayed plastic responses (environmental effect) and/or other response patterns (genotype-by-environment effect, GxE). In forest tree breeding, the GxE component of phenotypic plasticity is of particular importance for the improvement of growth performance because it is linked to the potential of the improved varieties and/or breeding populations to respond in an adaptive manner to large environmental ranges, whilst keeping productivity levels in an acceptable range. However, such plasticity may entail high costs. Indeed, the establishment of specific mechanisms for coping with environmental constraints may have costs in terms of information acquisition, adapted phenotype production and maintenance. Moreover, plasticity may lead to developmental instability and may affect other traits of interest, because of genetic linkage, pleiotropy or epistasis (de Witt *et al.*, 1998). High plasticity costs might lead to specialist strategies, with genotypes displaying lower levels of phenotypic plasticity and being restricted to a certain environment, whereas lower costs might lead to more generalist strategies, with genotypes adapting to varying environments through the adjustment of phenotypic parameters (van Tienderen, 1991; Sultan and Spencer, 2002).

Traits displaying plasticity have frequently been studied in resource-acquiring organs, such as roots (Kleczewski *et al.*, 2010) and leaves (Delagrange, 2011; Sultana *et al.*, 2010). The plasticity of leaf traits is crucial, given the key role of the leaves in photosynthesis and carbon sequestration. Leaf acclimation to environmental change may be studied at several scales (e. g. morphology, anatomy and physiology of leaves), reflecting the complexity of the leaf (Delagrange, 2011). The costs and benefits of foliar plasticity have often been investigated in

terms of structural and symplastic components, corresponding to morphological and physiological plasticity, respectively (Lusk *et al.*, 2008; Maherali *et al.*, 2010). Investments in the structural components of the leaf are assumed to be expensive and essentially irreversible for the plant (Poorter *et al.*, 2006), due to organogenesis and maintenance. However, such investments may be seen as long-term investments, increasing both the lifespan of the leaf and the amount of carbon acquired (Lusk *et al.*, 2008). Conversely, investments in symplastic components (i.e. plasticity in physiological traits) may entail fewer costs (Maherali *et al.*, 2010; Caruso *et al.*, 2006) and provide a more short-term response.

In recent years, much attention has focused on leaf plasticity in response to water deficit in eucalyptus (Warren *et al.,* 2006; Carter and White, 2009; Franks *et al.,* 2009; Turner *et al.,* 2010). These studies have shown that trait plasticity differs considerably between species. For example, a comparison of 14 *Eucalyptus* species in two common gardens under two rainfall regimes revealed differences in the degree of phenotypic plasticity for carbon discrimination, specific leaf area (SLA) and leaf nitrogen content (Turner *et al.*, 2010). In another study, Warren *et al.* (2006) compared 29 *Eucalyptus* species originating from areas with different rainfall levels and concluded that the physiological traits of leaves played a less important role in adaptation to water stress than morphological traits. Given this high degree of diversity among *Eucalyptus* species, it is important to characterise the response of a particular trait to site conditions, if we are to predict survival and growth, particularly in cases of climate change (Carter and White, 2009).

In the Republic of Congo, the dry season extends from May to October and is characterized by an absence of rainfall for four months, leading to a decrease in the productivity of *Eucalyptus* plantations (Nouvellon *et al.*, 2010). In this study, we characterized the foliar plasticity of two commercial genotypes (1-41 and 18-50) known to differ in terms of their growth and ecophysiological parameters when mature. These two clones were installed in a field trial and were submitted to two watering regimes during the dry seasons: non irrigated treatment (NI) corresponding to commercial production conditions, and irrigated treatment (IR). Foliar plasticity of 16-month-old trees was characterized at the morphological (mean leaf area, specific leaf area, leaf thickness), anatomic, physiological and molecular levels and was related to growth parameters. This study aimed to address the following questions: i) Which leaf traits are affected by seasonal water deficit? ii) Can we identify any common trends reflecting a specific strategy? iii) Do genotypes differ in terms of strategy, highlighting the genetic control of leaf plasticity?

2 Materials and Methods

2.1 Plant material

The two eucalyptus genotypes used in this study were selected on the basis of large differences in terms of growth and ecophysiological traits at maturity. Genotype 1-41, which was obtained by open pollination of *E. alba* (Delwaulle, 1985), was one of the first commercial clones to be used in Congolese plantations. Genotype 18-50 is an interspecific hybrid of *E. urophylla* and *E. grandis* obtained in the first generation of the Congolese breeding programs. At a mature stage, genotype 18-50 grows more rapidly than 1-41, with a mean annual increment of 28.5 *vs*. 25.1 $m^3.ha^{-1}.yr^{-1}$, respectively (Safou-Matondo *et al.*, 2005). It also has a higher water-use efficiency (4.3 *vs.* 2.2 $gC.kg^{-1}$ H_2O), and a higher specific leaf area (11.6 *vs.* 11.1 $kg.m^{-2}$; Y Nouvellon, personal communication). These two genotypes were propagated by cuttings, raised in a nursery for three months before planting.

2.2 Field experiment

The field experiment was set up in June 2007 in the coastal zone of Pointe-Noire, bordering the Atlantic Ocean (Yanika, 4°20'S, 11°38'E, 50 m a.s.l., Republic of Congo). The climate of this area is classified as tropical wet and dry, according to the Köppen classification, with a dry season extending from May to October. There is no rainfall and light availability is low during this period. Each genotype was subjected to two watering regimes during the dry season: one plot was irrigated (IR) so as to cover potential evapotranspiration losses, which were estimated at 3 mm per day, and the other plot was not irrigated (NI). Four plots consisting in 24 clonal replicates represented the two genotypes subjected to the two treatments. The measurements presented here were made at the end of the second dry season after planting (September 2008).

Soil water content (SWC) was measured weekly, from June 2007 to January 2009, by time domain reflectometry (TDR; Trase system, Soil moisture, Santa Barbara, CA). Four SWC profiles (four series of TDR probes installed at three depths: 0.15, 1 and 4 m) were analyzed for the four genotype x treatment conditions. Means were calculated for the replicates at each depth.

2.3 Growth and ecophysiological traits

We measured the height of all the trees in each plot (24 trees x 4 plots = 96 trees). For other measurements, we studied four to 11 trees with a height close to the median height of the plot.

For measurements of leaf dry weight, mean leaf area, specific leaf area and carbon discrimination, we harvested 11 trees per condition. For leaf and stem dry weight measurement, fresh leaves and stems were divided into different architectural units (Hallé, 1986; Barthelemy, 1991): architectural unit 1 corresponds to the trunk, architectural unit 2 corresponds to branches arranged along the trunk, and constitute the basic structure of architectural unit 3, etc. The total fresh mass of each architectural unit was determined by weighing in the field. Representative subsamples of each architectural unit were collected and weighed before and after drying at 65°C until a constant weight was reached. Water content was calculated for each of these subsamples and used to estimate total dry biomass for each architectural unit. Total leaf dry weight (LDW) and stem dry weight (SDW) were then obtained by adding the dry weights for each architectural unit.

Before drying, we scanned a sample of 25 leaves and analyzed the images with 'Scion Image' software (Scion Corporation, Frederick, MD, USA), to measure mean leaf area (MLA). Specific leaf areas (SLA) were then calculated from the measured leaf areas and the dry weight of the 25-leaf sample.

We selected 25 sun leaves per tree for analysis of carbon isotope composition. These leaves were dried at 65°C to constant weight and ground. A sample of 10 g was obtained by weighing and its isotopic composition ($\delta^{13}C$) was analyzed with an EA-2000 elemental analyzer (EuroVector SpA, Milan, Italy) coupled to an Elementar mass spectrometer (Isoprime Ltd, Cheadle, United Kingdom). Isotopic ratios were calculated as delta values in parts per thousand and not as carbon isotope discrimination (Farquhar and Richards, 1984), because the isotopic composition of carbon dioxide at the site was not measured.

We estimated stomatal density on a sample of eight leaves per tree, for five clonal replicates per condition. We used transparent nail varnish and tape to obtain impressions of the abaxial faces of leaves, as described by Ceulemans *et al.* (1995). These impressions were observed with an Olympus BM2 RFCA light microscope (Olympus America Inc., Melville, New York, USA), with the analysis of five microscopic fields per impression replicate at 20 x magnification (corresponding to representing 0.568 mm^2).

2.4 Leaf anatomic measurements

Four leaves of three to five trees per condition were harvested and immediately fixed in 100 mM phosphate buffer at pH 7.2, supplemented with 1% (v/v) glutaraldehyde, 2% (v/v) paraformaldehyde, 1% (w/v) caffeine, at room temperature (Schwendiman *et al.* 1988).

Leaves were then dehydrated and embedded in Technovit 7100 resin (Heraeus Kulzer), according to the manufacturer's recommendations.

We cut series of 2.5 µm transverse sections with glass knives on a Leica RM2065 rotary microtome. The collected sections were double-stained as described by Buffard-Morel *et al.* (1992). Polysaccharides were stained dark pink with periodic acid Schiff (PAS) reagent, and soluble proteins were stained blue with Naphthol Blue-Black stain (Sigma ref. 70490). Sections were then mounted in Mowiol. Slides were observed with a Leica DM6000 microscope (Leica, Germany), in bright-field conditions. Pictures were taken with a Retiga 2000R camera (QImaging Co.), and images were processed with Volocity 4.0.1 (Improvision, Lexington, MA, USA). Several measurements were made, as shown in figure 1. Some measurements were made only once per image, such as the thickness of the midrib, leaf blade, vascular bundle, upper and lower collenchyma and percentage of spongy parenchyma. Other measurements, such as the thickness of the upper and lower epidermis, of the upper and lower cuticle, of the palisade parenchyma and the diameter of xylem vessels, were measured five times and the mean value was then calculated.

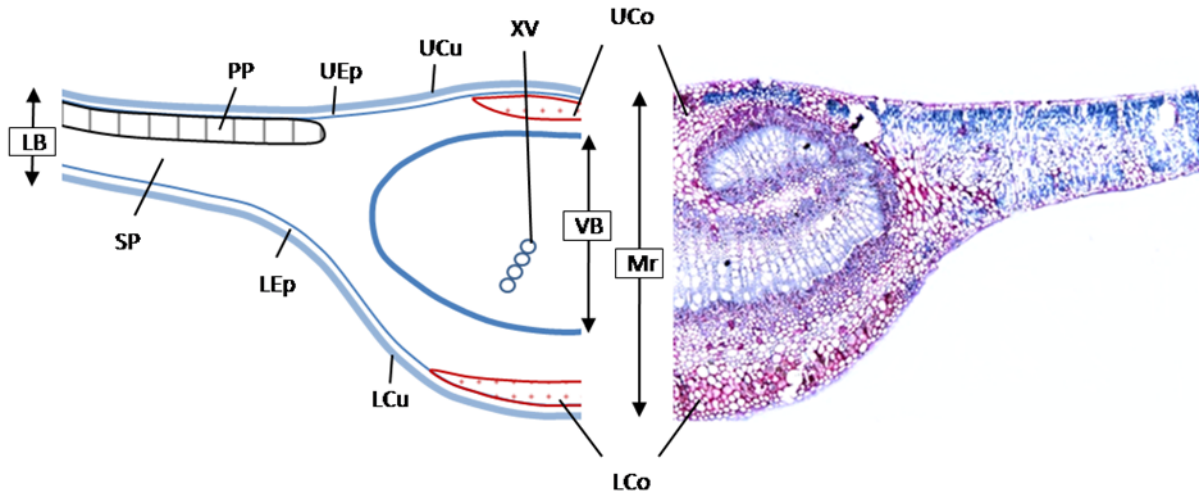


Figure 127 – Cross-section of eucalyptus leaf indicating the various parameters measured. Mr: midrib; LB: leaf blade; Lep: lower epidermis; Uep: upper epidermis; Lcu: lower cuticle; Ucu: Upper cuticle; SP: Spongy parenchyma; PP: palissade parenchyma; VB: vascular bundle; XV: xylem vessel diameter; Uco: upper collenchyma; Lco: lower collenchyma. Scale bar: 150 µm.

2.5 Determination and characterization of phenol content

We harvested 10 leaves per tree in the field, and immediately froze them in liquid nitrogen. A representative subsample was ground in liquid nitrogen, and 200 mg of the resulting powder was incubated twice, for 10 minutes each, with 10 ml of 70% aqueous acetone supplemented

with 2% formic acid. The extracts were combined, filtered, concentrated to dryness under vacuum, then dissolved in 5 ml of a mixture of methanol and water (50/50: v/v). Samples were passed through a filter with 0.45 µm pores (Millipore). Separation was carried out at 30°C, with a 250 mm × 4.6 mm, 5 µm particle size, endcapped reverse-phase ACE C-18 column (AIT, Houilles, France). For polyphenol quantification, HPLC analysis was carried out on a Dionex liquid chromatograph equipped with model P680 pumps, an ASI 100 autosampler and a UVD 340U diode array detector coupled to a HP ChemStation (Dionex, France). For polyphenol identification, HPLC/MS^n analysis was carried out on a SURVEYOR HPLC machine equipped with a UV6000LP DAD, model P4000 pumps and an AS3000 autosampler, coupled to an LCQ ion trap mass spectrometer fitted with an electrospray interface (Thermo Electron, San Jose, CA, USA). The LC/MS parameters used were described in a previous study (Mertz *et al.*, 2007). The solvents were 0.1% aqueous formic acid (solvent A) and acetonitrile/water/formic acid (80:19.9:0.1, v/v/v; solvent B). Gradient conditions were as follows: from 5 to 35% B in 50 min, from 35 to 50% B in 5 min, and from 50 to 80% B in 5 min, after which the column was washed and equilibrated for 15 min. The injection volume was 10 µl and the flow rate was 0.7 $ml.min^{-1}$. Detection was carried out at 280, 330 and 360 nm. Compounds were quantified by measuring peak areas and the results are expressed in $mAu.min^{-1}$.

2.6 Transcriptome analysis: 454-sequencing and RT-qPCR

Three half-runs of 454-sequencing were performed on shoot apices of the two genotypes subjected to the two treatments, as described by Villar *et al.* (2011), resulting in 12 sequencing sets. The read frequency for each sequencing set and for each contig was used to represent contig abundance. Digital differential expression (DE) was assessed with the DEGseq package in R (Wang *et al.*, 2010), and G, T and GxT effects were assessed by two-way analysis of variance in R (R Development Core Team). The log_2-transformed fold-change between genotypes 1-41 and 18-50 (G effect), between treatments IR and NI (T effect), and between IR and NI treatment for each genotype (GxT effect) were calculated with the DEGseq package. Bins were assigned to DE-contigs with Mercator (http://mapman.gabipd.org/web/guest/mercator). For each contig, fold-changes were plotted by Mapman (Usadel *et al.* 2009) on a pathway map representing symplastic and structural components, created from a bibliographic review.

We chose 20 genes to represent the symplastic or structural component and the expression of these genes was analyzed in mature leaves by RT-qPCR. Briefly, RNA was extracted as

described by Reid *et al.* (2006) and reverse transcribed with the ImpromII Reverse Transcription System (Promega USA, Madison, WI), according to the manufacturer's instructions. Primer pairs were designed from the 454-reads, with Primer3 software (Rozen and Skaletsky, 2000; see additional file 1). Real-time PCR was performed on 384-well plates, and monitored with a Light Cycler 480 (Roche Applied Science USA, Indianapolis, IN). Each reaction was performed in duplicate. The fluorescent signal was evaluated and starting concentrations were calculated with LinRegPCR (Ramakers *et al.* 2003; Ruijter *et al.* 2009). The relative abundance of transcripts was normalized with respect to the levels of mRNA for four constitutively expressed housekeeping genes —the elongation factor 1-α (*EF1-α*), elongation factor 2 (*eEF2*), glyceraldehyde 3P dehydrogenase (*G3PDH*) and cyclophilin (*Cyp*) genes — by calculating a normalization factor with GeNorm (Vandesompele *et al.*, 2002).

2.7 Statistical analysis

For all traits excepted 454-sequences abundance, means and standard errors were calculated. Plasticity index (PI) was calculated as follows: PI = (mean $_{NI}$ –mean $_{IR}$) / mean $_{IR}$ for each genotype. For those traits, we evaluated genotype (G), treatment (T) and genotype by treatment interaction (GxT) effects, by two-way ANOVA, according to the following model: $X_{ijk} = X_{mean} + aG_i + bT_j + cGxT_{ij} + \varepsilon_{ijk}$, where X_{ijk} is the trait measured in genotype i (1-14 or 18-50), treatment j (NI or IR) and for the replicate k. a, b and c are the regressors, and ε_{ijk} is the residual. ANOVA and principal component analysis (PCA) were performed with the XLSTAT statistical package of Excel.

3 Results

3.1 Watering treatment

In the field experiment, during the second dry season of the experiment rainfall ceased between May 22nd and October 20th 2008. During this dry season, the IR plots were irrigated from June 5th June to October 28th 2008, whereas no water was supplied to the NI plots for the two genotypes (1-41 and 18-50). In the NI condition, soil water content (SWC) at a depth of 15 cm reached a trough very rapidly after May 22nd . Minimum SWC at a depth of 1 and 4 m were reached in mid-July and t end of August, respectively (figure 2).

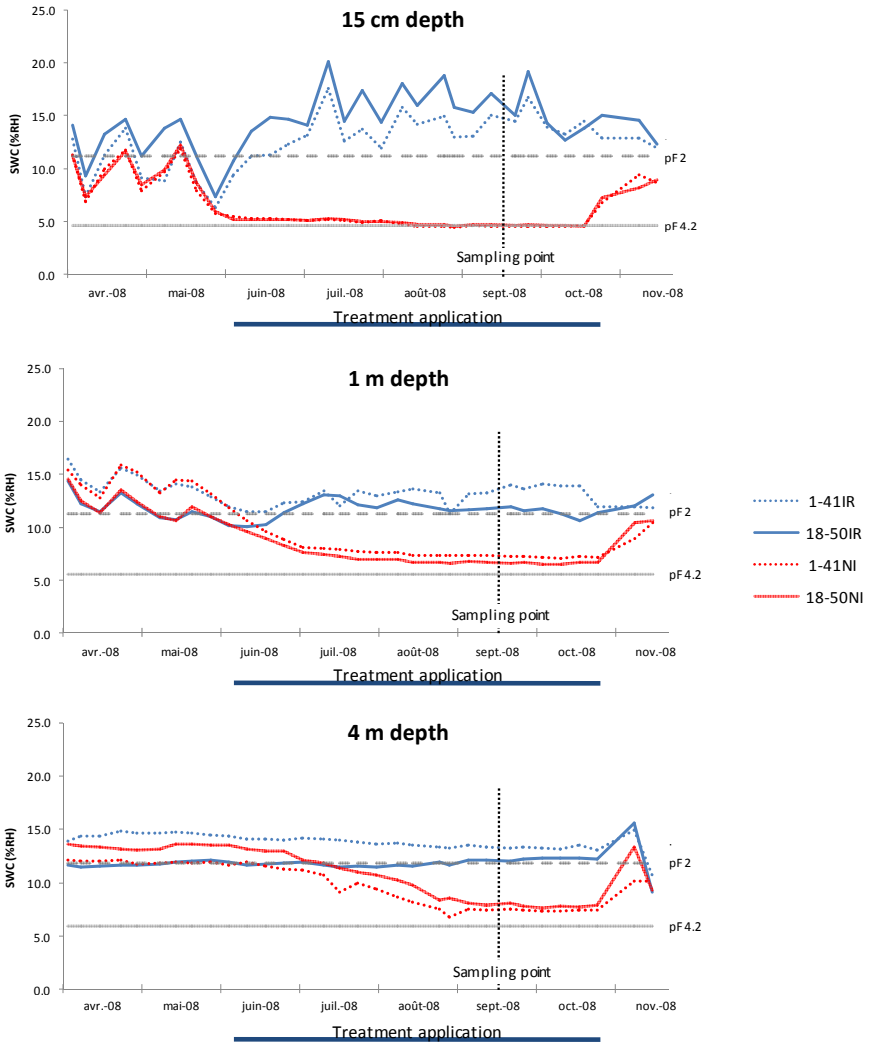


Figure 2 – Soil water content (SWC) at depths of 15 cm, 1 m and 4 m, from April 2008 to November 2008. Blue bars show the irrigation period for the irrigated treatment, for both genotypes (1-41IR, 18-50IR). The black dotted line indicates the sampling date.

Samples for phenotypic and molecular analysis were taken between September 15th and 22nd 2008, after four months without rainfall, when the trees were 16 months old. At the time of sampling, SWC for the IR treatment was up to field capacity (pF2, *i.e.* the moisture potential at field capacity); ranging from 11 to 17%, depending on the genotype and the depth considered. By contrast, SWC was lower in the NI treatment (from 4.5 to 8%), close to the wilting point (pF4.2, *i.e.* the moisture potential at which plants cease to be able to absorb soil water). For each treatment, and for the three depths, SWC did not differ significantly between genotypes.

3.2 Morphological and ecophysiological measurements

Table 1 summarizes the mean values for growth, morphological and ecophysiological traits for each genotype and each treatment, and the significance of G, T and GxT effects for each trait.

Table 1 – Means and standard errors, ANOVA p-values and plasticity index for the growth, morphological and ecophysiogical traits measured in this study for the two genotypes (1-41 and 18-50) subjected to the two treatments (IR and NI). Abbreviations are as follows: SDW, stem dry weight; LDW, leaf dry weight; MLA, mean leaf area; SLA, specific leaf area; $\delta^{13}C$, carbon isotopic composition; CC, carbon content.

	Mean ± Standard error				ANOVA p-values			Plasticity index	
	1-41NI	1-41IR	18-50NI	18-50IR	G effect	T effect	GxT effect	1-41 IR -> NI	18-50 IR -> NI
Height (m)	6.01 ± 0.68	6.95 ± 0.52	8.22 ± 0.85	9.87 ± 0.67	< 0.0001	< 0.0001	0.016	-14%	-17%
VOB (m^3)	0.009 ± 0.003	0.012 ± 0.004	0.012 ± 0.003	0.019 ± 0.005	< 0.0001	< 0.0001	0.006	-19%	-35%
SDW (kg)	5.83 ± 0.9	7.75 ± 1.05	9.6 ± 1.75	11.11 ± 1.23	< 0.0001	< 0.0001	0.605	-25%	-14%
LDW (kg)	1.67 ± 0.24	2.06 ± 0.26	2.31 ± 0.42	2.52 ± 0.36	< 0.0001	0.005	0.373	-19%	-8%
MLA (cm^2)	66.96 ± 5.68	69.47 ± 7.04	32.49 ± 2.5	35.2 ± 2.45	< 0.0001	0.083	0.947	-4%	-8%
SLA ($m^2.kg^{-1}$)	9.46 ± 0.72	10.47 ± 0.83	9.83 ± 0.9	11.78 ± 0.45	0.001	< 0.0001	0.048	-10%	-17%
Stomatal density (nb . mm^{-2})	114.79 ± 2.89	121.48 ± 4.49	112.68 ± 3.73	129.58 ± 19.19	0.518	0.019	0.276	-5%	-13%
δ13C (‰)	-30.55 ± 0.52	-31.64 ± 0.37	-29.54 ± 0.59	-30.58 ± 0.59	< 0.0001	< 0.0001	0.891	-3%	-3%

The two genotypes differed significantly in terms of growth: 18-50 was taller (9.04 m on average) than 1-41 (6.48 m), and produced more stem biomass (SDW=10.36 kg for 18-50 *vs.* 6.74 kg for 1-41) and more leaf biomass (LDW=2.42 kg and 1.85 kg for 18-50 and 1-41, respectively). Height (HT), stem dry weight (SDW) and leaf dry weight (LDW) were significantly lower for the NI treatment than for the IR treatment, for both genotypes (mean values for the two genotypes decreased from 8.4 m to 7.1 m for HT, 9.51 kg to 7.72 for SDW, and 2.30 to 1.99 for LDW). GxT effects were significant for height and to a less extent for SLA. The NI treatment had a stronger effect on 18-50 than on 1-41 for HT, with a plasticity index of -17% and -14%, respectively.

Mean leaf area (MLA) was lower for 18-50 (33.85 cm^2 per leaf) than for 1-41 (68.20 cm^2 per leaf), but specific leaf area (SLA) was higher for 18-50 (10.81 $m^2.kg^{-1}$ versus 9.97 $m^2.kg^{-1}$ for 1-41), suggesting that 18-50 had a lower leaf density than 1-41. Watering significantly increased SLA (11.16 $m^2.kg^{-1}$ for IR *vs.* 9.66 $m^2.kg^{-1}$ for NI) but not MLA (51.5 cm^2 per leaf for IR *vs.* 48.9 cm^2 per leaf for NI), suggesting that water deficit greatly increased leaf density. A significant GxT effect was detected for SLA, with NI treatment having a greater effect on SLA for 18-50 (PI of -17%) than for 1-41 (PI of -10%).

Carbon isotope composition ($\delta^{13}C$) was higher for 18-50 (-30.06‰) than for 1-41 (-31.07‰), suggesting that 18-50 may have a higher water use efficiency than 1-41. Carbon isotope composition was lower for the IR treatment than for the NI treatment (from -30.05‰ in NI to -31.09‰ in IR), but the difference between NI and IR conditions was similar for the two genotypes (PI= -3% for 18-50 and 1-41). Carbon content (CC) was significantly higher for genotype 1-41 than for 18-50, but no significant treatment effect was observed.

No differences in stomatal density (SD) were detected between genotypes, but there was a significant difference between NI (114 st. mm^{-2}) and IR (126 $st.mm^{-2}$) treatments. No significant GxT interaction was observed for this trait.

As expected, growth traits (HT, VOB and LDW) were strongly positively correlated (figure 3). In terms of morphological and physiological traits, SLA values were positively correlated with growth traits, whereas MLA, $\delta^{13}C$ and stomatal density displayed no correlation with growth traits. No clear relationship between MLA, $\delta^{13}C$ and stomatal density was revealed. MLA differed considerably between the two genotypes, and genotype 1-41 displayed considerable variation for this trait, whereas SLA was highly variable for genotype 18-50.

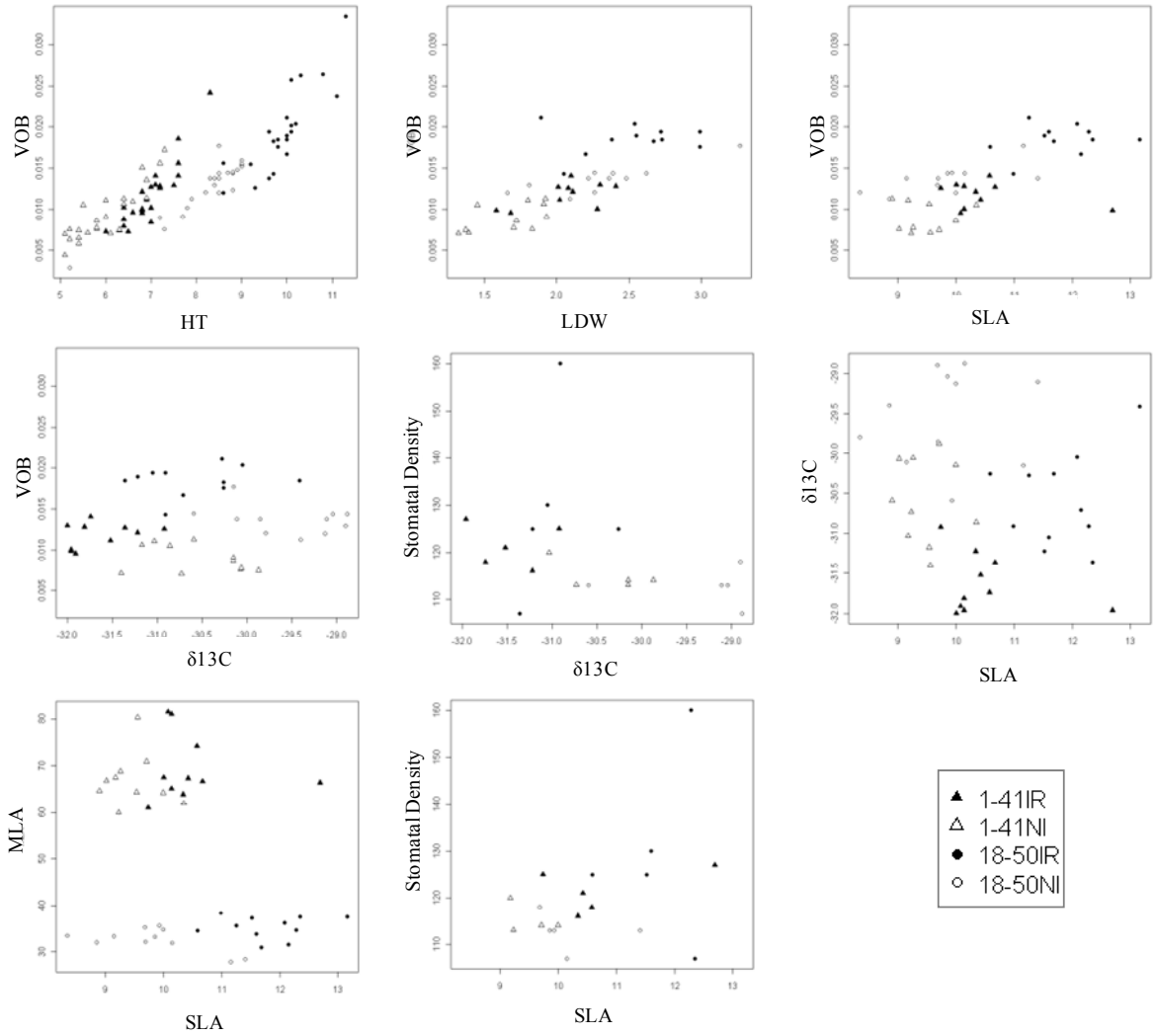


Figure 2 – Scatter plots of values for growth, morphological and physiological traits for both genotypes (1-41 shown as triangles, 18-50 as circles) subjected to the two treatments (IR closed symbols, NI open symbols). For volume over bark (VOB) and height (HT), 24 trees per condition were measured, for leaf dry weight (LDW), specific leaf area (SLA), mean leaf area (MLA) and carbon isotopic composition ($\delta^{13}C$), 10 to 11 replicates were measured, and 5 replicates were used for stomatal density.

Table 2 – Means and standard errors, plasticity index and ANOVA p-values for the anatomic traits measured in this study.

		Mean ± Standard error				ANOVA p-values			Plasticity index	
		1-41NI	1-41IR	18-50NI	18-50IR	G effect	T effect	GxT effect	1-41 IR -> NI	18-50 IR -> NI
Mr	Midrib thickness (µm)	1120 ± 299	1317 ± 161	804 ± 205	733 ± 158	0.001	0.566	0.234	-15%	10%
LB	Leaf blade thickness (µm)	406 ± 113	416 ± 35	324 ± 22	318 ± 49	0.019	0.952	0.815	-2%	2%
Lep	Lower epidermis thickness (µm)	11.7 ± 2	15 ± 2.3	10 ± 1.9	11.6 ± 0.5	0.020	0.022	0.380	-22%	-14%
Uep	Upper epidermis thickness (µm)	11 ± 2	16.2 ± 2.8	10.6 ± 2	10.8 ± 0.7	0.019	0.024	0.035	-32%	-2%
Lcu	Lower cuticle thickness (µm)	19.8 ± 4.8	14.6 ± 1.2	17.3 ± 3.8	8.9 ± 1	0.019	0.001	0.312	36%	94%
Ucu	Upper cuticle thickness (µm)	17.4 ± 3.9	11.8 ± 1.3	15 ± 3.7	8.7 ± 0.5	0.058	0.001	0.785	47%	72%
SP	Spongy parenchyma (% LB)	13.5 % ± 3.7	9.4 % ± 4.4	15.1 % ± 8.2	10.3 % ± 3	0.641	0.093	0.889	44%	47%
PP	Palisade parenchyma (%LB)	21.3 % ± 5.6	19.7 % ± 2.2	17.8 % ± 1.9	22.4 % ± 4.2	0.857	0.463	0.132	8%	-21%
VB	Vascular Bundle (%Mr)	47.8 % ± 2.4	49.8 % ± 3.7	47.9 % ± 7.5	45.6 % ± 4.3	0.386	0.945	0.361	-4%	5%
XV	Xylem vessels (%VB)	4.7 % ± 0.3	5.1 % ± 0.7	5 % ± 0.4	6 % ± 0.8	0.065	0.034	0.367	-8%	-17%
Uco	Upper collenchyma (%Mr)	35 % ± 4.1	39 % ± 1.8	44.4 % ± 6.4	37.3 % ± 2.8	0.066	0.431	0.012	-10%	19%
Lco	Lower collenchyma (%Mr)	30.4 % ± 3.9	32 % ± 4.9	31.1 % ± 1.8	33.6 % ± 2.6	0.567	0.301	0.838	-5%	-7%

3.3 Anatomic measurements

Pearson's correlation coefficients of anatomic measurements indicated that midrib, vascular bundle, leaf blade, xylem vessel diameters and collenchyma thicknesses were highly correlated (more than 75% at the 0.001 significance threshold, data not shown). We chose to use the most strongly correlated leaf growth traits (midrib, leaf blade or vascular bundle thickness) for normalization of the measured tissues thicknesses, to prevent size biases. We thus normalized parenchyma thicknesses with respect to leaf blade thickness, collenchyma and vascular bundle thicknesses with respect to midrib thickness, and xylem vessel diameter with respect to vascular bundle thickness.

Anatomic measurements (table 2) confirmed the trends observed in leaf morphology: genotype 1-41 had thicker leaves than 18-50, with a thicker midrib (1230 µm for 1-41 *vs* 764 µm for 18-50) and a thicker leaf blade (412 µm for 1-41 *vs* 320 µm for 18-50). No significant treatment effect was observed for these traits. The epidermis and cuticles were thicker in genotype 1-41 than in 18-50. The epidermis was thicker for the IR treatment than for the NI treatment, whereas the cuticle was thicker for the NI treatment. No significant difference was found between genotypes or treatments for palisade and spongy parenchyma, vascular bundle and lower collenchyma thicknesses. A significant treatment effect was observed for xylem vessel diameter, which was higher for the IR than for the NI treatment. A significant GxT interaction was observed for upper collenchyma thickness, which was greater for the IR for 1-41 (PI of -10%) and for the NI treatment for 18-50 (PI of 19%). The GxT interaction was also significant for upper epidermis thickness: the upper epidermis was thicker for the IR treatment for 1-41 (PI of -32%), whereas no significant difference between the two treatments was observed for 18-50 (PI of -2%).

Despite the small number of measurements, we were able to observe trends in multivariate analysis. In PCA (figure 4), the first two principal components accounted for more than 71% of the variation, and seemed to discriminate between genotypes and treatments, with the exception of two outliers. Consistent with the observations described above, genotype 1-41 was characterized by a thicker midrib, leaf blade and epidermis, whereas genotype 18-50 had higher proportion of palisade parenchyma and a larger xylem vessel diameter. NI treatment seemed to increase cuticle thickness and the proportion of spongy parenchyma, whereas IR treatment tended to increase the diameter of xylem vessels.

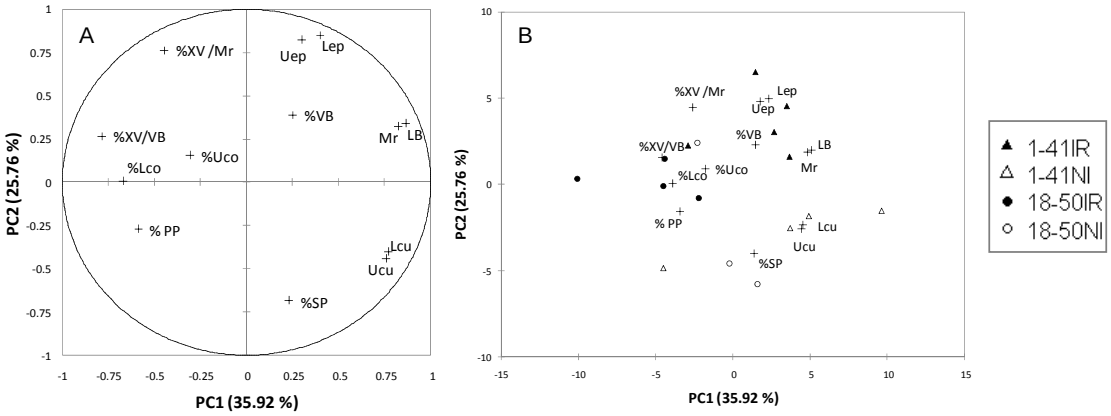


Figure 4 – Principal component analysis of anatomic measurements (PC1 and PC2). (A) correlation circle showing the position of the variables on the PC1-PC2 plane. (B) Main plane showing the positions of observations. Mr: midrib; LB: leaf blade; Lep: lower epidermis; Uep: upper epidermis; Lcu: lower cuticle; Ucu: Upper cuticle; SP: Spongy parenchyma; PP: palissade parenchyma; VB: vascular bundle; XV: xylem vessel diameter; Uco: upper collenchyma; Lco: lower collenchyma.

3.4 Polyphenol content

Polyphenols were extracted from leaves and separated by HPLC. Compounds were identified by mass spectrometry and UV spectroscopy, through comparisons with published data. Twenty peaks were detected and identified by number, with letters for overlapping or coeluting compounds (8, 9, 16, 18 and 19). Overall, 27 different compounds were identified. For peak 8 (retention time, Tr: 33.3 min), luteolin-C-hexose was identified in genotype 1-41, whereas the compound present in 18-50 remained unidentified. In peak 9, two compounds, rutine and apigenine C-hexoside, were identified only in genotype 1-41. In peak 16 (Tr: 42.8 min), two different compounds, quercetin rhamnoside and isorhamnetin pentoside were identified in genotype 1-41, whereas kaempferol hexoside and glucuronide were identified in 18-50. In peak 18 (Tr: 44.2 min), isorhamnetin glucuronide was identified in genotype 1-41, whereas quercetin-(galloyl) pentoside was identified in 18-50. Finally, in peak 19 (Tr: 44.5 min), quercetin hexoside was identified in genotype 1-41, whereas kaempferol pentoside was identified in 18-50. Table 3 summarizes the main results obtained for leaf polyphenol content. Three of the 27 compounds separated by HPLC remained unidentified, five were hydroxycinnamic acids, four were flavones, and 15 were flavonols. Only five compounds (A-NA; C-chlorogenic acid; D-coumaroyl quinic acid; N-quercetin glucuronide; O-quercetin hexoside) were present in both genotypes, reflecting broad qualitative differences between the genotypes.

Table 3– Characteristics of the polyphenol compounds analyzed and means, standard errors, plasticity index and ANOVA p-values for the quantitative values measured.

Peak	Retention time	Compound	Identification	Wavelength	Category	Mean ± Standard error (mAu.min^{-1})				ANOVA p-values			Plasticity index	
						1-41NI	1-41IR	18-50NI	18-50IR	G effect	T effect	GxT effect	1-41 IR -> NI	18-50 IR -> NI
1	19.6	A	NA	280	NA	137.7 ± 22.5	103.4 ± 27	261 ± 33.8	192.8 ± 57.8	0.001	0.037	0.442	33%	35%
2	20.8	B	coumaroyl quinic acid	330	hydroxycinnamic acids	15.7 ± 2.2	15.9 ± 2.9	0	0	< 0.0001	0.900	n.a.	-1%	n.a.
3	22.2	C	chlorogenic acid	330	hydroxycinnamic acids	11.7 ± 9.3	8 ± 5.4	44.5 ± 12.3	20.9 ± 3.8	0.000	0.009	0.039	46%	113%
4	28	D	coumaroyl quinic acid	330	hydroxycinnamic acids	3.9 ± 1.6	3.4 ± 1.3	23.9 ± 3.7	14.2 ± 4.1	< 0.0001	0.010	0.017	15%	68%
5	28.4	E	coumaroyl quinic acid	330	hydroxycinnamic acids	0	0	3.5 ± 3.4	2.8 ± 0.3	0.003	0.702	n.a.	n.a.	25%
6	32.2	F	luteolin-C-hexoside	330	hydroxycinnamic acids	10 ± 1.5	11.6 ± 0.2	0	0	< 0.0001	0.049	n.a.	-14%	n.a.
7	32.5	G	coumaroyl quinic acid	330	hydroxycinnamic acids	0	0	5 ± 1.2	3.7 ± 0.6	< 0.0001	0.075	n.a.	n.a.	35%
8	33.3	H	luteolin-C-hexoside	330	flavones	5 ± 1.2	5.7 ± 0.6	0	0	< 0.0001	0.326	n.a.	-12%	n.a.
8	33.3	I	NA	330	NA	0	0	1.4 ± 0.5	1.7 ± 0	< 0.0001	0.286	n.a.	n.a.	-18%
9	36.8	J	rutin	330	flavonols	12 ± 7.9	11 ± 5.3	0	0	< 0.0001	0.663	n.a.	9%	n.a.
9	36.8	K	apigenine C-hexoside	330	flavones	12 ± 7.9	11 ± 5.3	0	0	< 0.0001	0.663	n.a.	9%	n.a.
10	37.2	L	apigenin-C-hexoside	330	flavones	7.3 ± 1.2	6.7 ± 0.6	0	0	< 0.0001	0.441	n.a.	9%	n.a.
11	37.6	M	quercetin hexoside	330	flavonols	0	0	8 ± 3.3	10.3 ± 1.5	< 0.0001	0.234	n.a.	n.a.	-22%
12	38	N	quercetin glucuronide	330	flavonols	30.9 ± 2	34.7 ± 1.3	22.1 ± 3.3	23.7 ± 4.3	0.000	0.131	0.504	-11%	-7%
13	38.3	O	quercetin hexoside	330	flavonols	15.3 ± 0.9	17.4 ± 1.5	0.6 ± 0.2	1.6 ± 1.1	< 0.0001	0.025	0.384	-12%	-63%
14	38.8	P	kaempferol glucuronide	330	flavonols	7.8 ± 0.9	8.7 ± 0.7	0	0	< 0.0001	0.151	n.a.	-10%	n.a.
15	41	Q	quercetin pentoside	330	flavonols	0	0	17.2 ± 2.8	17.6 ± 2.2	< 0.0001	0.826	n.a.	n.a.	-2%
16	42.8	R	quercetin rhamnoside	330	flavonols	17.5 ± 6.2	18.7 ± 3	0	0	< 0.0001	0.738	n.a.	-6%	n.a.
16	42.8	S	isorhamnetin pentoside	330	flavonols	17.5 ± 6.2	18.7 ± 3	0	0	< 0.0001	0.738	n.a.	-6%	n.a.
16	42.8	T	kaempferol hexoside	330	flavonols	0	0	16.4 ± 7.1	13.9 ± 1.8	< 0.0001	0.497	n.a.	n.a.	18%
16	42.8	U	glucuronide	330	flavonols	0	0	16.4 ± 7.1	13.9 ± 1.8	< 0.0001	0.497	n.a.	n.a.	18%
17	43.5	V	isorhamnetin hexoside	330	flavonols	6 ± 1	5.9 ± 0.5	0	0	< 0.0001	0.852	n.a.	2%	n.a.
18	44.2	W	isorhamnetin glucuronide	330	flavonols	1.8 ± 0.4	1.7 ± 0.7	0	0	< 0.0001	0.801	n.a.	6%	n.a.
18	44.2	X	quercetin-(galloyl) pentoside	330	flavonols	0	0	5.2 ± 1.9	5.6 ± 1.1	< 0.0001	0.743	n.a.	n.a.	-7%
19	44.5	Y	quercetin hexoside	330	flavonols	6.2 ± 1.7	6.6 ± 0.5	0	0	< 0.0001	0.664	n.a.	-6%	n.a.
19	44.5	Z	kaempferol pentoside	330	flavonols	0	0	15.7 ± 2.8	16.7 ± 2.4	< 0.0001	0.641	n.a.	n.a.	-6%
20	48.3	AA	NA	280	NA	5.9 ± 1.1	5.5 ± 1.6	0	0	< 0.0001	0.724	n.a.	7%	n.a.

In quantitative analysis, these five compounds presented genotype effects: A-NA, C-chlorogenic acid and D-coumaroyl quinic acid were more abundant in 18-50 than in 1-41, whereas N-quercetin glucuronide and O-quercetin hexoside were more abundant in 1-41. Four of these five compounds presented treatment effects (the exception being N-quercetin glucuronide): A-NA, C-chlorogenic acid and D-coumaroyl quinic acid levels were increased by NI treatment, whereas O-quercetin hexoside levels were decreased by this condition. Only one of the 22 compounds present in one genotype only presented a treatment effect: F-luteolin-C-hexoside was more abundant in 1-41IR than in 1-41NI. A significant GxT interaction effect was observed for C-chlorogenic acid and D-coumaroyl quinic acid, with larger increases in response to NI treatment for genotype 18-50 (113% and 68% for compounds C and D, respectively) than for 1-41 (46% and 15%, respectively).

3.5 Overview of gene expression in shoot apices

We carried out RNA-Seq to analyze patterns of gene expression in shoot apices. Eight cDNA libraries, corresponding to two replicates of the four conditions (1-41IR; 18-50IR; 1-41NI and 18-50NI) were sequenced, with 3 half runs of 454-sequencing (Villar *et al.*, 2011). We used $\log_2$-transformed ratios to compare contig abundances for genes encoding structural and symplastic components. According to ANOVA, G, T and GxT effects were attributed to 1,445 contigs. Only 38 of these contigs corresponded to structural components, classified as genes related to cellulose, pectate lyases, lignins and fatty acids synthesis, and other cell wall-related genes. Conversely 374 contigs were classified as symplastic components. Primary metabolism was represented by genes related to the metabolism of proteins, carbohydrates and lipids. Secondary metabolism was represented by genes corresponding to isoprenoids, phenylpropanoids and flavonoids. The other categories considered for symplastic components were photosynthesis, ATP synthesis, redox regulation and transport-related genes (figure 5, additional file 2). A comparison of contig abundance between 1-41 and 18-50 showed that much of the variance could be accounted for by G effect (figure 5A). For most gene categories, there seemed to be a mosaic of genes overexpressed in one genotype or another. Just to give an example, protein degradation seemed to be driven principally by serine proteases in genotype 1-41 and by ubiquitins in genotype 18-50. However, two categories of genes seemed to be more strongly expressed in 1-41: 38 of 44 photosynthesis-related contigs and six of nine contigs corresponding to other cell wall-related genes. Very few contigs displayed differential expression as a function of treatment (figure 5B).

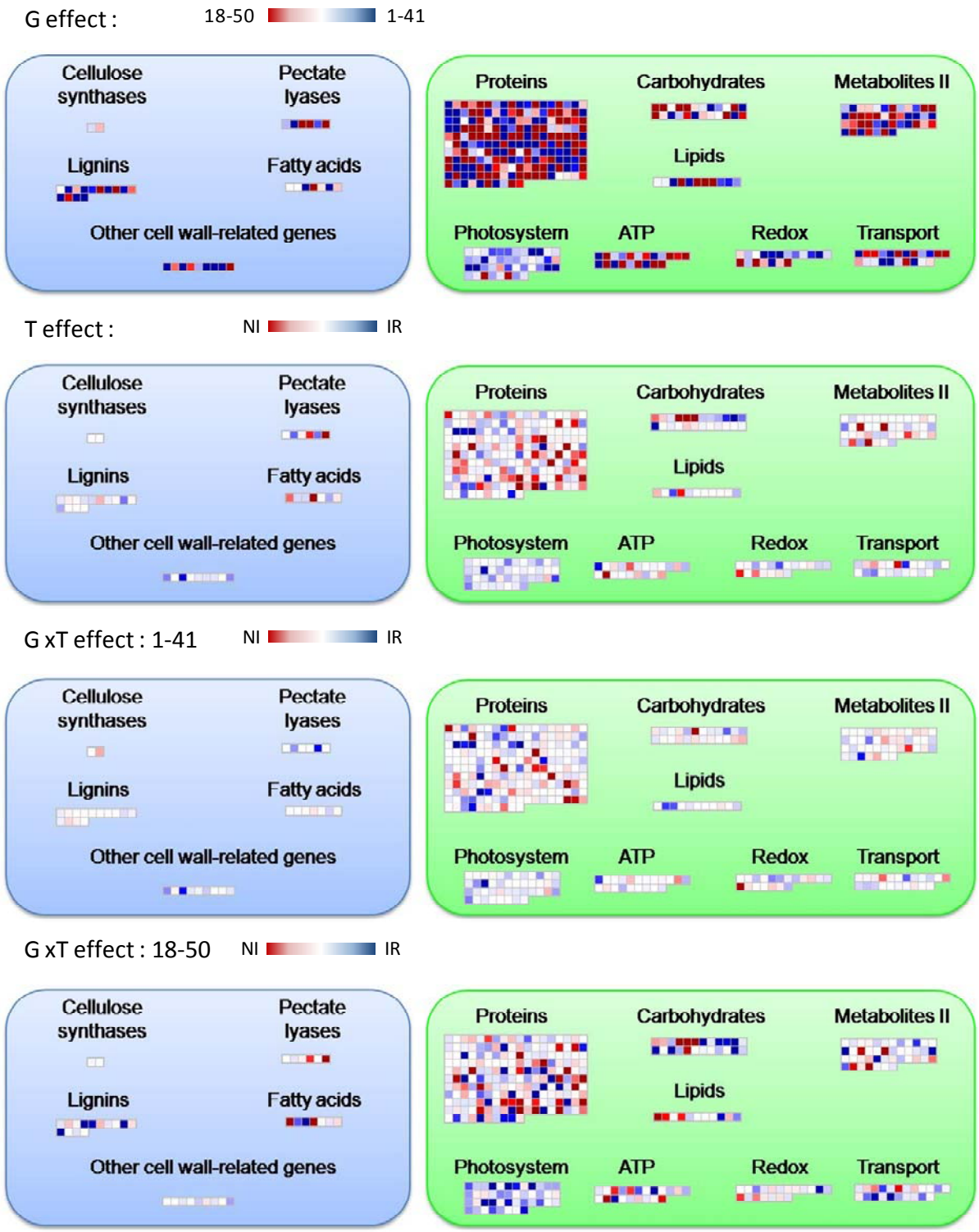


Figure 5 – Comparison of gene expression between A) genotypes (G effect), B) treatments (T effect) and between treatments for each genotype separately (GxT effect; C) 1-41 and D) 18-50). Genes related to structural components are shown in the blue box, whereas genes related to symplastic components are shown in the green box. Each square represents the log_2-transformed ratio between the two conditions compared.

Photosystem components, transport and cell wall-related genes were globally overexpressed in the IR condition. Carbohydrate metabolism genes were overexpressed due to starch degradation in the NI treatment, whereas genes related to the metabolism of minor carbohydrates were overexpressed in the IR treatment. A comparison of treatment effects for each genotype separately (figure 5C and 5D) suggested that 18-50 had a more marked response to treatment than 1-41. Some categories showed similar patterns between genotypes, but different degrees of plasticity: i.e. scale interaction effect, as described by Lynch and Walsh (1998). Thus, photosystem, cell wall and lignin biosynthesis-related genes were overexpressed in 18-50IR, whereas few differences between both conditions were observed for 1-41. Similarly starch degradation-related genes were overexpressed in the NI treatment for 18-50, but no clear differences between conditions were observed for 1-41. Conversely, pectate lyases and glycolysis-related genes had divergent patterns of expression in 1-41 and 18-50: pectate lyase genes were overexpressed in 18-50NI and in 1-41IR, and glycolysis-related genes were overexpressed in 18-50IR and in 1-41NI.

3.6 Expression of selected genes in mature leaves

We studied the expression of 10 genes representing structural components and 12 genes representing symplastic components by RT-qPCR (table 4).

Table 4 – Means and standard errors of expression levels, plasticity index and ANOVA p-values for the expression of 22 genes measured by RT-qPCR

		Means ± SE ($x10^9$)				ANOVA p-values			Plasticity index	
Gene name	Gene code	1-41NI	1-41IR	18-50NI	18-50IR	G effect	T effect	GxT effect	1-41 IR -> NI	18-50 IR -> NI
4-Coumarate: CoA ligase	4CL	86 ± 18	141 ± 82	91 ± 22	62 ± 36	0.096	0.559	0.059	-39%	48%
Caffeoyl-CoA 3-O-methyltransferase 1	CCoAOMT1	224 ± 70	110 ± 62	114 ± 29	78 ± 61	0.012	0.009	0.143	104%	46%
Caffeoyl-CoA 3-O-methyltransferase 2	CCoAOMT2	146 ± 91	184 ± 133	154 ± 49	114 ± 88	0.476	0.978	0.360	-21%	35%
Cuticular 1; catalytic	CER6	122 ± 107	307 ± 282	54 ± 16	119 ± 46	0.075	0.082	0.392	-60%	-55%
Cellulose sythase 1	CesA1	44 ± 48	187 ± 228	79 ± 50	72 ± 38	0.469	0.222	0.181	-77%	9%
Caffeic acid O-methyl transferase	COMT2	98 ± 31	198 ± 70	62 ± 32	74 ± 59	0.002	0.024	0.070	-51%	-17%
Glycosyl hydrolase family 3 protein	GHP3	13 ± 11	17 ± 15	33 ± 16	40 ± 15	0.003	0.405	0.809	-23%	-18%
Transcription factor	LIM1	58 ± 69	92 ± 92	121 ± 42	93 ± 39	0.271	0.924	0.284	-37%	31%
BURP domain-containing protein / polygalacturonase	POLG	0.414 ± 0.25	1.28 ± 1.86	126 ± 72	79 ± 28	0.000	0.203	0.187	-68%	60%
S-adenosyl methionine synthase	SAMS	50 ± 34	38 ± 19	89 ± 57	88 ± 50	0.030	0.708	0.776	34%	2%
1-deoxy-d-xylulose 5-phosphate reductoisomerase	DXR	32 ± 20	112 ± 74	21 ± 5	79 ± 47	0.287	0.003	0.590	-71%	-73%
Adenosylmethionine decarboxylase family protein	DCAM	136 ± 46	524 ± 203	52 ± 25	246 ± 183	0.007	0.000	0.108	-74%	-79%
Delta tonoplast integral protein	dTIP	102 ± 80	513 ± 287	37 ± 25	124 ± 79	0.004	0.002	0.028	-80%	-70%
Early light inducible protein / chlorophyll binding	ELIP	242 ± 92	161 ± 64	43 ± 26	131 ± 102	0.003	0.915	0.023	50%	-67%
Glutathione-disulfide reductase	GDR	85 ± 43	59 ± 20	56 ± 25	59 ± 45	0.360	0.466	0.365	44%	-5%
Nonspecific lipid transfer protein 1	LP1	5562 ± 3823	10394 ± 3849	3087 ± 1182	3115 ± 1604	0.002	0.092	0.096	-46%	-1%
Proton gradient regulation 5	PGR5	97 ± 50	129 ± 105	13 ± 11	56 ± 52	0.012	0.212	0.848	-25%	-77%
Alpha-glucan phosphorylase 2	PHS2	1.22 ± 1.07	0.07 ± 0.056	125 ± 92	62 ± 64	0.001	0.214	0.231	1648%	101%
Plasma membrane intrinsic protein 1;2	PIP1B	63 ± 43	68 ± 65	16 ± 9	22 ± 18	0.017	0.772	0.946	-6%	-30%
60S acidic ribosomal protein P1	RPP1A	0.081 ± 0.103	0.281 ± 0.331	0.059 ± 0.04	0.315 ± 0.525	0.686	0.201	0.600	-71%	-81%
Hydrolase, hydrolyzing O-glycosyl compounds	SIP2	493 ± 199	500 ± 239	283 ± 115	267 ± 223	0.022	0.961	0.900	-1%	6%
Ubiquinol-cytochrome C reductase complex	UCR	93 ± 45	75 ± 9	195 ± 80	153 ± 70	0.002	0.271	0.651	23%	27%

Fourteen genes displayed a genotype effect, being mostly overexpressed in genotype 1-41 (9 of the 14 genes). Most of the genes corresponding to symplastic components displayed a G effect (10 of 12), whereas few structure-related genes presented a G effect (4 of 10), suggesting that the two genotypes differed principally in symplastic functioning rather than establishment of structure. Four of the five genes displaying a T effect (*COMT2*; *DCAM*; *δTIP*; *DXR*) were more strongly expressed in the IR treatment. Only caffeoyl-CoA 3-O-

methyltransferase 1 (*CCoAOMT1*) was overexpressed in the NI treatment. In terms of GxT interaction effects, a gene encoding an aquaporin (*δTIP*) displayed a slighty higher underexpression in NI than in IR conditions for 1-41 (PI of -80%) than for 18-50 (PI of -70%). More interestingly, a second gene, involved in photosynthesis (*ELIP1*), was overexpressed in 1-41 (PI of +50%) subjected to NI treatment, whereas it was underexpressed in 18-50 (PI of -67%).

A scatterplot of the PCA for gene expression revealed no general trend, indicating that gene expression could not be fully accounted for by conditions analyzed (additional file 3). Only the three replicates of 18-50 for the NI treatment were clustered, suggesting that expression was most strongly conserved between replicates for this genotype in this condition. The overexpression of *PHS2*, *POLG*, *UCR* and *SAMS* seemed to be related to 18-50NI, possibly corresponding to a specific molecular adjustment in response to water deficit.

4 Discussion

In this study, we compared responses to water deficit during the dry season of two genotypes known to differ in terms of productivity. Despite the genotype 18-50 had a higher cumulated growth whatever the treatment (VOB = 0.016 m^3 per tree in average *vs* 0.010 m^3 for 1-41), it displayed a higher difference of growth between the two treatments (PI of -35% for VOB in 18-50 vs PI of -19% for 1-41). Those differences may be accounted for by different strategies used by the both genotypes, that would be described in the following sections.

4.1 In NI treatment, the more productive genotype reduced photosynthesis, primary metabolism and growth

In our study, carbon isotope composition ($\delta^{13}C$) was lower in the NI condition for both genotypes (strong T effect, p-value < 10^{-4}). There may be two reasons for this decrease (Farquhar *et al.,* 1989): i) lower levels of CO_2 diffusion through stomata, and/or ii) lower levels of CO_2 incorporation due to a lower efficiency of enzymatic reactions. These two factors involve different structural and symplastic components observable at different scales of analysis.

A morphological analysis of leaves showed that MLA was not affected in the NI condition (no significant T or GxT effect), although it did differ considerably between genotypes, due to the parental species used to produce the hybrids studied. NI treatment decreased stomatal

density in the two genotypes, due to a decrease in the number of stomata per leaf in each genotype. This variation of the number of stomata may account for the lower $\delta^{13}C$ levels, due to a decrease in mean CO_2 diffusion at the scale of a leaf. In addition to this structural plasticity, symplastic components, such as stomatal closure, may also account for these differences in carbon isotope discrimination but were not investigated here.

CO_2 diffusion through stomata is a prerequisite for photosynthesis, so isotope carbon discrimination may be an indicator of photosynthetic activity (Farquhar *et al.*, 1989). In terms of anatomy, the relative thicknesses (% of leaf blade) of the tissues supporting photosynthesis (palisade and spongy parenchyma) revealed no significant G, T or GxT interactions. There was therefore no plasticity in the structural components of the tissues supporting photosynthesis. By contrast, transcriptome analysis showed that symplastic components associated genes were mostly underexpressed in the NI condition, in both shoot apices and mature leaves, suggesting a decrease in photosynthesis. This is illustrated by the expression levels in mature leaves of *PGR5* in both genotypes. *PGR5* (a gene encoding a proton gradient-regulating protein) has been shown to regulate cyclic electron flow around photosystem I in response to environmental stimuli in *Arabidopsis thaliana* (Nandha *et al.*, 2007). It decreased in the NI condition, and this decrease was stronger for 18-50 than for 1-41 (significant GxT effect). A similar abundance profile was observed in the shoot apex for *PGR5*. An other gene, *ELIP* (encoding a chlorophyll-binding protein) was underexpressed in genotype 18-50 subjected to the NI treatment, which is consistent with the putative role of the encoded protein in chlorophyll synthesis suggested by Rossini *et al.* (2006). However, it was found to be overexpressed in genotype 1-41. Osipenkova *et al.* (2010) showed that *ELIP* could be influenced by several exogenous or endogenous stimuli, including light in particular, so the differences in expression pattern between both genotypes may be due to differences in the perception of stress signals.

As a consequence of the lower level of photosynthesis, global cell activity levels seemed to be lower in water deficit conditions, this trend appearing more marked for genotype 18-50 than for 1-41. Primary metabolism-associated genes related to myoinositol synthesis were underexpressed in the NI condition, particularly in 18-50. Patra *et al.* (2010) showed that transgenic tobacco plants accumulating higher levels of inositol have a higher competence for growth and photosynthesis. Genes related to lipid metabolism and cell walls were also downregulated in the NI treatment, as were most water transport-related genes (e.g. *dTIP*, a gene encoding an aquaporin involved in water transport). A similar trend was observed in

mature leaves, particularly in genotype 1-41. These results suggest a global slowing of cell metabolism in shoot apices under conditions of water deficit that may affect genotype 18-50 more strongly than 1-41, which agrees with a higher pronounced decrease of growth in 18-50 between the IR and NIR conditions.

4.2 The more productive genotype engage leaf structures for preserving turgor

MLA and leaf thickness did not differ between the two watering regimes, but the NI treatment decreased SLA in both genotypes. This trend in SLA variation in response to water deficit is consistent with a significant increase in leaf density. Seasonal variations of SLA have already been described in several species of *Eucalyptus* (Turner *et al.*, 2008), particularly under the climatic conditions studied here (Nouvellon *et al.,* 2010). There may be several reasons for this decrease in SLA: the selective retention of low-SLA leaves (Ishii *et al.,* 2002), a decrease in the ratio of young to old leaves (Sala *et al.*, 1994), variations in leaf nutrient or non structural carbohydrate contents (Misson *et al.*, 2006) and retranslocation processes (Sala *et al.*, 1994). In the following sections, we will described modifications of leaf structures concerning dermal, support and conducting tissues that could accounted for by decrease of SLA in NI treatment.

Dermal tissue

The increase in leaf density may be accounted for by several anatomic changes: simultaneous increases in cuticle thickness (a heavy layer) and decreases in epidermal thickness (a light layer). A decrease in epidermis cell size has already been reported in drought-stressed olive plants (Bosabalidis and Kofidis, 2002) and has been shown to increase resistance to cell collapse (Oertli *et al.*, 1990). The leaf cuticle consists of lipids, mostly cutin compounds interleaved with waxes, conferring resistance to water diffusion (Riederer and Schreiber, 2001). Increase in leaf cuticle thickness has been widely described as a drought tolerance mechanism in several species, including succulents (Cristescu *et al.,* 2010), herbaceous plants (Kosma *et al.*, 2009) and trees (Bacelar *et al.*, 2004). For example, comparing two provenances of *Eucalyptus cloeziana*, Ngugi *et al.* (2003) showed that trees from dry lands had thicker cuticles than those from wet lands. Unfortunately, in the shoot apex transcriptome survey, we found no significant T or GxT effects for genes related to wax synthesis, whereas, in other species, these genes are considered to be candidate genes for drought tolerance improvement. These genes include the *WAX2* homologous genes in rice (Gao *et al.*, 2010) and

gl1 in maize (Sturaro *et al.*, 2006). Cuticle reinforcement-related genes may not yet be expressed in the immature tissues of shoot apices, or they may be expressed at a level below the detection threshold.

Support tissues

The decrease in SLA in response to water deficit was higher in genotype 18-50 than in 1-41, possibly due to an increase in the amount of upper collenchyma at the midrib in this genotype. The development of collenchyma cells in response to water deficit has already been described in transgenic rice plants (Hanba *et al.,* 2004). This tissue may be involved in the response to water deficit, with its thick cell walls, consisting mostly of cellulose and pectin, serving as an apoplastic reservoir for ion storage and ensuring turgor regulation (Satter and Galston, 1981). In this study, we observed a global overexpression of genes related to pectate lyases in the shoot apices of genotype 18-50 subjected to water deficit. In mature leaves of 18-50, a polygalacturonase-related gene (*POLG*) was expressed, but the expression of this gene was not detected in 1-41, and no differential expression of this gene as a function of irrigation treatment was observed. These results may reflect both a need for pectin degradation for shoot apex development (taking various stages of leaf development into account) and greater needs in conditions of water deficit due to the early establishment of collenchyma tissue in the leaf primordium.

Conducting tissue

Finally, the structure of conducting tissues was also involved in the response to water deficit, with a decrease of xylem vessel diameters observed in the NI condition, particularly in genotype 18-50. According to the Hagen-Poiseuille law, a decrease in vessel diameter may improve hydraulic conductance (Tyree and Zimmermann, 2002). Similarly, Hao *et al.* (2010) showed that *Ficus* species with narrower petiole vessels were more able to persist under drought stress than those with wider petiole vessels. Decreases in vessel thickness may indicate that genotypes adapt their conductive structures to improve water delivery in leaves, and that this process occurred to a greater extent in genotype 18-50.

4.3 The more productive genotype synthesized more foliar compounds related to antioxidant activity

Chemical characterization of the polyphenol content of leaves showed 27 polyphenol compounds in mature leaves with a significant genotype effect. Three compounds from the

three polyphenol categories identified (hydroxycinnamic acids, flavonols and flavones) displayed T effects and two presented T and GxT effects. These polyphenols were mostly hydroxycinnamic acids, overrepresented in the NI condition (4 of 5-fold increase). Chlorogenic acid and coumaroyl quinic acid (compounds with a significant GxT effect) accumulated to a greater extent in genotype 18-50 than in 1-41, and twice as much of these compounds was present in the NI than in the IR condition for 18-50. Chlorogenic acid and coumaroyl quinic acid are intermediates in phenylpropanoid, flavonoid and stilbenoid biosynthesis (Weisshaar and Jenkins, 1998), and are known to be involved in defense mechanisms (Dixon and Paiva, 1995). Moreover, the accumulation of phenylpropanoids and flavonoids during water deficit has been widely described: in *Arabidopsis* (Bray, 2002), grape (Grimplet *et al.*, 2009) and maize (Vincent *et al.*, 2005), with the accumulation of these compounds in various organs, including leaves and berries, accounted for by their antioxidant function. For example, Grace and Logan (2000) showed that chlorogenic acid acts as a hydrogen-donating antioxidant and a reductive substrate for peroxidase under conditions of carbohydrate accumulation or excess light energy. At the chemical level genotype 18-50 showed more plasticity in response to water deficit for two antioxidants compounds than the 1-41, potentially conferring more efficient defense mechanisms. This hypothesis is supported by the overexpression of genes involved in terpene biosynthesis (secondary metabolites with antioxidant properties) in the shoot apices of 18-50 in the NI condition (data not shown).

5 Conclusion

In this study, we found that the more productive genotype (18-50) displayed greater plasticity of foliar traits than genotype 1-41 during the dry season. We first observed a decrease in SLA in the NI condition, this decrease being greater for genotype 18-50 than for genotype 1-41 and potentially resulting from several physiological adjustments. A transcriptome survey suggested a possible decrease in photosynthesis and primary metabolism, making it possible to reallocate resources for the establishment of protective structures or agents. We identified mechanisms for limiting water loss by cuticle reinforcement and a decrease in stomatal density, for preserving leaf turgor by increasing the thickness of the collenchyma layer, and for enhancing hydraulic conductance by decreasing xylem vessel diameter. Moreover, the accumulation of antioxidant compounds, such polyphenols, may help to preserve leaf functions. The accumulation of complex protective compounds, such as lipids, lignins and phenols, entails a high cost (Wright *et al.*, 2004). As larger decreases in SLA are associated

with larger increases in construction costs, it seems likely that genotype 18-50 makes a greater investment than genotype 1-41 when faced with water deficit. This could accounted for by the decrease between IR and NI conditions were higher for 18-50 in terms of growth. This "high-cost" strategy of genotype 18-50 may confer advantages in terms of leaf resistance to biotic and abiotic stresses (Wright *et al.*, 2004). The higher growth potential of this genotype may demonstrate the benefits from these long-term investments, with leading leaves able to recover high photosynthesis performance as soon as it starts to rain again. However, this costly strategy may not remain adaptive if the dry season becomes longer, as predicted, due to climate change.

Acknowledgments:

This article forms part of the PhD thesis of Emilie Villar, funded by CIRAD. This work was funded by CIRAD (ATP "Plasticité Phénotypique des pérennes sous contrainte hydrique au champ") and *Conseil Régional d'Aquitaine* ("ABIOGEN" project, no. Presage 32973).

We would like to thank Chafia Belaidi and Fabienne Montes for anatomic studies (UMR DAP, CIRAD), Alicia Maillard for phenol characterization (UMR Qualisud, CIRAD), Antony Reychler for stomatal density measurements (UMR BioGeCo, INRA), and Pascal Tillard for delta ^{13}C measurements (UMR IBIP, CIRAD).

Some of the experiments (RT-qPCR) presented here were performed at the Genome-Transcriptome facility of Bordeaux (grants from the *Conseil Régional d'Aquitaine* nos. 20030304002FA and 20040305003FA and from the European Union, FEDER no. 2003227). We would like to thank Annabel Porté and Alain Franc (UMR BioGeCo) for interesting discussions about plant physiology and statistical analysis.

Additional file 1–Primers used for the RT-qPCR analysis

Gene name	Gene code	Forward primer (5'-3')	Reverse primer (5'-3')
4-Coumarate: CoA ligase	4CL	CGATGAAGGATGAGGTTGC	AACACGACCTGCTTCGAGAT
Caffeoyl-CoA 3-O-methyltransferase 1	CCoAOMT1	GTATGTGAGGTACTACAGGG	CTCTCTTTTTCGTGGGTTCT
Caffeoyl-CoA 3-O-methyltransferase 2	CCoAOMT2	GCAAGATGAGAAGAACCATG	AGCACGAAGTCCCGGTAGTA
Cuticular 1; catalytic	CER6	TCCATATTCCCGAAGTGGTC	CGCAATCCACCTATCAAACC
Cellulose sythase 1	CesA1	TAAAGTGCTTGCTGGCATTG	TGAGAAGCGATGTCCATTTG
Caffeic acid O-methyl transferase	COMT2	GTGGATATGCCATGACTGGA	TGACCAAGAACTTGCACACC
Glycosyl hydrolase family 3 protein	GHP3	GGGTTGCAGACGTTCTGTTT	ATGAGGATCCCCAACATTCA
Transcription factor	LIM1	AGAGGGAAACAATGGCTCCT	GCTTACAACCAACCCGAGAA
BURP domain-containing protein / polygalacturonase	POLG	CGAAACCAAGATGTGTGTGG	CGATCAAGACCTCCTGCTTC
S-adenosyl methionine synthase	SAMS	TTCATACCGTGCTGATCTCG	TGATGATCTTACGGCCAGTG
1-deoxy-d-xylulose 5-phosphate reductoisomerase	DXR	TGAGATCATTCCAGGGGAAC	GCTATGTCCTTGCCTGCTTC
Adenosylmethionine decarboxylase family protein	DCAM	ATCCAGAGGGTGTTGGTCTG	CTCCAGCACTGTACCCCTTC
Delta tonoplast integral protein	dTIP	CCGTGGTCAGTGGAGACTTC	CATTGACAAGAGGTGCATGG
Early light inducible protein / chlorophyll binding	ELIP	TCCCTCTTCATCGCCTACAC	GTCTGCCATTGATCCTCTCC
Glutathione-disulfide reductase	GDR	TAAGTGTGGTGCAACGAAGG	AACCCGTCTTGTGACTGACC
Nonspecific lipid transfer protein 1	LP1	AGGGGATCGACTTCAACCTC	GAGCTCGTGGACCTCACTTC
Proton gradient regulation 5	PGR5	TGGCTTGTTGCTAGGGTTTC	ATTGATCGCGAGCACTTTTC
Alpha-glucan phosphorylase 2	PHS2	TTCTAGCTTCCCTCCCTTCC	GTTAGGCGGAGTGCATTTGT
Plasma membrane intrinsic protein 1;2	PIP1B	TTCTGCAACATGGCTACACC	ATTCCTCTTGGCATCAGTGG
60S acidic ribosomal protein P1	RPP1A	CACCAATGTTGGTTCTGGTG	AGACCAAAGCCCATGTCATC
Hydrolase, hydrolyzing O-glycosyl compounds	SIP2	CGATTTGGCGCTTACTCTTC	ACCGGCCATCTGTACATCTC
Ubiquinol-cytochrome C reductase complex	UCR	GCTGTCTCCTGACGTTTGTG	CCGGCTATTTCGACTCCTTC

REFERENCES

Bacelar, E. A., C. M. Correia, *et al.* (2004). "Sclerophylly and leaf anatomical traits of five field-grown olive cultivars growing under drought conditions." Tree Physiology **24**(2): 233-239.

Barthélémy, D. (1991). "Levels of organization and repetition phenomena in seed plants." Acta Biotheoretica **39**(3): 309-323.

Bosabalidis, A. M. and G. Kofidis (2002). "Comparative effects of drought stress on leaf anatomy of two olive cultivars." Plant Science **163**(2): 375-379.

Bradshaw, A. D. and E. W. C. a. J. M. Thoday (1965). Evolutionary Significance of Phenotypic Plasticity in Plants. Advances in Genetics, Academic Press. **Volume 13:** 115-155.

Bray, E. A. (2002). "Classification of genes differentially expressed during water-deficit stress in Arabidopsis thaliana: an analysis using Microarray and differential expression data." Annals of Botany **89**: 803-811.

Buffard-Morel, J., J. L. Verdeil, *et al.* (1992). "Somatic Embryogenesis of Coconut-Palms (Cocos-Nucifera-L.) from Leaf Explants - Histology." Canadian Journal of Botany-Revue Canadienne De Botanique **70**(4): 735-741.

Carter, J. L. and D. A. White (2009). "Plasticity in the Huber value contributes to homeostasis in leaf water relations of a mallee Eucalypt with variation to groundwater depth." Tree Physiology **29**(11): 1407-1418.

Caruso, C. M., H. Maherali, *et al.* (2006). "Plasticity of physiology in Lobelia: Testing for adaptation and constraint." Evolution **60**(5): 980-990.

Ceulemans, R., L. van Praet, *et al.* (1995). "Effects of CO2 Enrichment, Leaf Position and Clone on Stomatal Index and Epidermal Cell Density in Poplar (Populus) " New Phytologist **131**(1): 99-107.

Cristescu, M., D. Anton, *et al.* (2010). "Study of leaves structures that determine the resistance to dryness at succulent plants." Journal of Horticulture, Forestry and Biotechnology **14**(1): 60-64.

Delagrange, S. (2011). "Light- and seasonal-induced plasticity in leaf morphology, N partitioning and photosynthetic capacity of two temperate deciduous species." Environmental and Experimental Botany **70**(1): 1-10.

Delwaulle, J. C. (1985). "Clonal Plantations of Hybrid Eucalyptus in the Congo." Bois et Forets des Tropiques(208): 37-42.

DeWitt, T. J., A. Sih, *et al.* (1998). "Costs and limits of phenotypic plasticity." Trends in Ecology & Evolution **13**(2): 77-81.

Dixon, R. A. and N. L. Paiva (1995). "Stress-Induced Phenylpropanoid Metabolism." Plant Cell **7**(7): 1085-1097.

Farquhar, G. D., J. R. Ehleringer, *et al.* (1989). "Carbon Isotope Discrimination and Photosynthesis." Annual Review of Plant Physiology and Plant Molecular Biology **40**: 503-537.

Franks, P. J., P. L. Drake, *et al.* (2009). "Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: an analysis using Eucalyptus globulus." Plant, Cell and Environment **32**(12): 1737-1748.

Gao, G., J. Zou, *et al.* (2010). "Tissue speciality and stress responses in expression of three WAX2 homologous genes in rice." Acta Agronomica Sinica **36**(8): 1336-1341.

Grace, S. C. and B. A. Logan (2000). "Energy dissipation and radical scavenging by the plant phenylpropanoid pathway." Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **355**(1402): 1499-1510.

Grimplet, J., M. D. Wheatley, *et al.* (2009). "Proteomic and selected metabolite analysis of grape berry tissues under well-watered and water-deficit stress conditions." Proteomics **9**(9): 2503-2528.

Grulke, N. E. (2010). "Plasticity in physiological traits in conifers: Implications for response to climate change in the western US." Environmental Pollution **158**(6): 2032-2042.

Halle, F. (1986). "Modular Growth in Seed Plants." Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **313**(1159): 77-87.

Hanba, Y. T., M. Shibasaka, *et al.* (2004). "Overexpression of the barley aquaporin HvPIP2;1 increases internal CO2 conductance and CO2 assimillation in the leaves of transgenic rice plants." Plant and Cell Physiology **45**(5): 521-529.

Hao, G. Y., L. Sack, *et al.* (2010). "Differentiation of leaf water flux and drought tolerance traits in hemiepiphytic and non-hemiepiphytic Ficus tree species." Functional Ecology **24**(4): 731-740.

Ishii, H., E. D. Ford, *et al.* (2002). "Variation in specific needle area of old-growth Douglas-fir in relation to needle age, within-crown position and epicormic shoot production." Tree Physiology **22**(1): 31-40.

Kleczewski, N. M., D. A. Herms, *et al.* (2010). "Effects of soil type, fertilization and drought on carbon allocation to root growth and partitioning between secondary metabolism and ectomycorrhizae of Betula papyrifera." Tree Physiology **30**(7): 807-817.

Kosma, D. K., B. Bourdenx, *et al.* (2009). "The impact of water deficiency on leaf cuticle lipids of Arabidopsis." Plant Physiology **151**(4): 1918-1929.

Lusk, C. H., P. B. Reich, *et al.* (2008). "Why are evergreen leaves so contrary about shade?" Trends in Ecology & Evolution **23**(6): 299-303.

Maherali, H., C. M. Caruso, *et al.* (2010). "Adaptive Value and Costs of Physiological Plasticity to Soil Moisture Limitation in Recombinant Inbred Lines of Avena barbata." American Naturalist **175**(2): 211-224.

Mertz, C., V. Cheynier, *et al.* (2007). "Analysis of phenolic compounds in two blackberry species (Rubus glaucus and Rubus adenotrichus) by high-performance liquid chromatography

with diode array detection and electrospray ion trap mass spectrometry." Journal of Agricultural and Food Chemistry **55**: 8616-8624.

Misson, L., K. P. Tu, *et al.* (2006). "Seasonality of photosynthetic parameters in a multi-specific and vertically complex forest ecosystem in the Sierra Nevada of California." Tree Physiology **26**(6): 729-741.

Nandha, B., G. Finazzi, *et al.* (2007). "The role of PGR5 in the redox poising of photosynthetic electron transport." Biochimica Et Biophysica Acta-Bioenergetics **1767**: 1252-1259.

Ngugi, M. R., D. Doley, *et al.* (2003). "Leaf water relations of Eucalyptus cloeziana and Eucalyptus argophloia in response to water deficit." Tree Physiology **23**(5): 335-343.

Nouvellon, Y., J. P. Laclau, *et al.* (2010). "Within-stand and seasonal variations of specific leaf area in a clonal Eucalyptus plantation in the Republic of Congo." Forest Ecology and Management **259**(9): 1796-1807.

Oertli, J. J., S. H. Lips, *et al.* (1990). "The Strength of Sclerophyllous Cells to Resist Collapse Due to Negative Turgor Pressure." Acta Oecologica-International Journal of Ecology **11**(2): 281-289.

Osipenkova, O. V., M. S. Odintsova, *et al.* (2010). "The influence of light, hormonal, and carbohydrate signal systems on ELIP gene expression in gun-mutants Arabidopsis thaliana." Applied Biochemistry and Microbiology **46**(3): 331-338.

Patra, B., S. Ray, *et al.* (2010). "Enhanced salt tolerance of transgenic tobacco plants by co-expression of PcINO1 and McIMT1 is accompanied by increased level of myo-inositol and methylated inositol." Protoplasma **245**(1-4): 143-152.

Poorter, H., S. Pepin, *et al.* (2006). "Construction costs, chemical composition and payback time of high- and low-irradiance leaves." Journal of Experimental Botany **57**(2): 355-371.

Ramakers, C., J. M. Ruijter, *et al.* (2003). "Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data." Neuroscience Letters **339**(1): 62-66.

Ramirez-Valiente, J. A., D. Sanchez-Gomez, *et al.* (2010). "Phenotypic plasticity and local adaptation in leaf ecophysiological traits of 13 contrasting cork oak populations under different water availabilities." Tree Physiology **30**(5): 618-627.

Rehfeldt, G. E., W. R. Wykoff, *et al.* (2001). "Physiologic plasticity, evolution, and impacts of a changing climate on Pinus contorta." Climatic Change **50**(3): 355-376.

Reid, K. E., N. Olsson, *et al.* (2006). "An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development." BMC Plant Biology **6**(27): (14 November 2006).

Riederer, M. and L. Schreiber (2001). "Protecting against water loss: analysis of the barrier properties of plant cuticles." Journal of Experimental Botany **52**(363): 2023-2032.

Rossini, S., A. P. Casazza, *et al.* (2006). "Suppression of both ELIP1 and ELIP2 in Arabidopsis does not affect tolerance to photoinhibition and photooxidative stress." Plant Physiology **141**(4): 1264-1273.

Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-86.

Ruijter, J. M., C. Ramakers, *et al.* (2009). "Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data." Nucleic Acids Research **37**(6).

Safou-Matondo, R., P. Delaporte, *et al.* (2005). "Hybrid and clonal variability of nutrient content and nutrient use efficiency in Eucalyptus stands in Congo." Forest Ecology and Management **210**(1-3): 193-204.

Sala, A., S. Sabate, *et al.* (1994). "Canopy Structure within a Quercus-Ilex Forested Watershed - Variations Due to Location, Phenological Development, and Water Availability." Trees-Structure and Function **8**(5): 254-261.

Sampedro, L., X. Moreira, *et al.* (2010). "Genetics, phosphorus availability, and herbivore-derived induction as sources of phenotypic variation of leaf volatile terpenes in a pine species." Journal of Experimental Botany **61**(15): 4437-4447.

Satter, R. L. and A. W. Galston (1981). "Mechanisms of Control of Leaf Movements." Annual Review of Plant Physiology and Plant Molecular Biology **32**: 83-110.

Schlichting, C. D., M. Pigliucci, *et al.* (1998). "Phenotypic evolution: A reaction norm perspective." Phenotypic evolution: A reaction norm perspective: xii+387p.

Sturaro, M., F. Salamini, *et al.* (2006). "Characterization of gl1, a maize gene that affects cuticular wax accumulation." Maize Genetics Cooperation Newsletter(80): 2-3.

Sultan, S. E. and H. G. Spencer (2002). "Metapopulation structure favors plasticity over local adaptation." American Naturalist **160**(2): 271-283.

Sultana, M., T. Asaeda, *et al.* (2010). "Morphological plasticity of submerged macrophyte Potamogeton wrightii Morong under different photoperiods and nutrient conditions." Chemistry and Ecology **26**(3): 223-232.

Trentacoste, E. R., V. O. Sadras, *et al.* (2011). "Effects of the source:sink ratio on the phenotypic plasticity of stem water potential in olive (Olea europaea L.)." Journal of Experimental Botany.

Turner, N. C., E. D. Schulze, *et al.* (2010). "Growth in two common gardens reveals species by environment interaction in carbon isotope discrimination of Eucalyptus." Tree Physiology **30**(6): 741-747.

Turner, N. C., E. D. Schulze, *et al.* (2008). "Annual rainfall does not directly determine the carbon isotope ratio of leaves of Eucalyptus species." Physiologia Plantarum **132**(4): 440-445.

Tyree, M. T. and M. H. Zimmermann (2002). Xylem Structure and the Ascent of Sap, Springer.

Usadel, B., F. Poree, *et al.* (2009). "A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize." Plant Cell and Environment **32**(9): 1211-1229.

Valladares, F. and U. Niinemets (2008). "Shade Tolerance, a Key Plant Feature of Complex Nature and Consequences." Annual Review of Ecology Evolution and Systematics **39**: 237-257.

van Tienderen, P. H. (1991). "Evolution of Generalists and Specialists in Spatially Heterogeneous Environments." Evolution **45**(6): 1317-1331.

Vandesompele, J., K. De Preter, *et al.* (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biology **3**(7).

Villar, E., C. Klopp, *et al.* (2011). "RNA-Seq reveals genotype-specific molecular responses during seasonal water deficit in eucalyptus." Submitted to BMC Genomics.

Vincent, D., C. Lapierre, *et al.* (2005). "Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation." Plant Physiology **137**(3): 949-960.

Vitasse, Y., C. C. Bresson, *et al.* (2010). "Quantifying phenological plasticity to temperature in two temperate tree species." Functional Ecology **24**(6): 1211-1218.

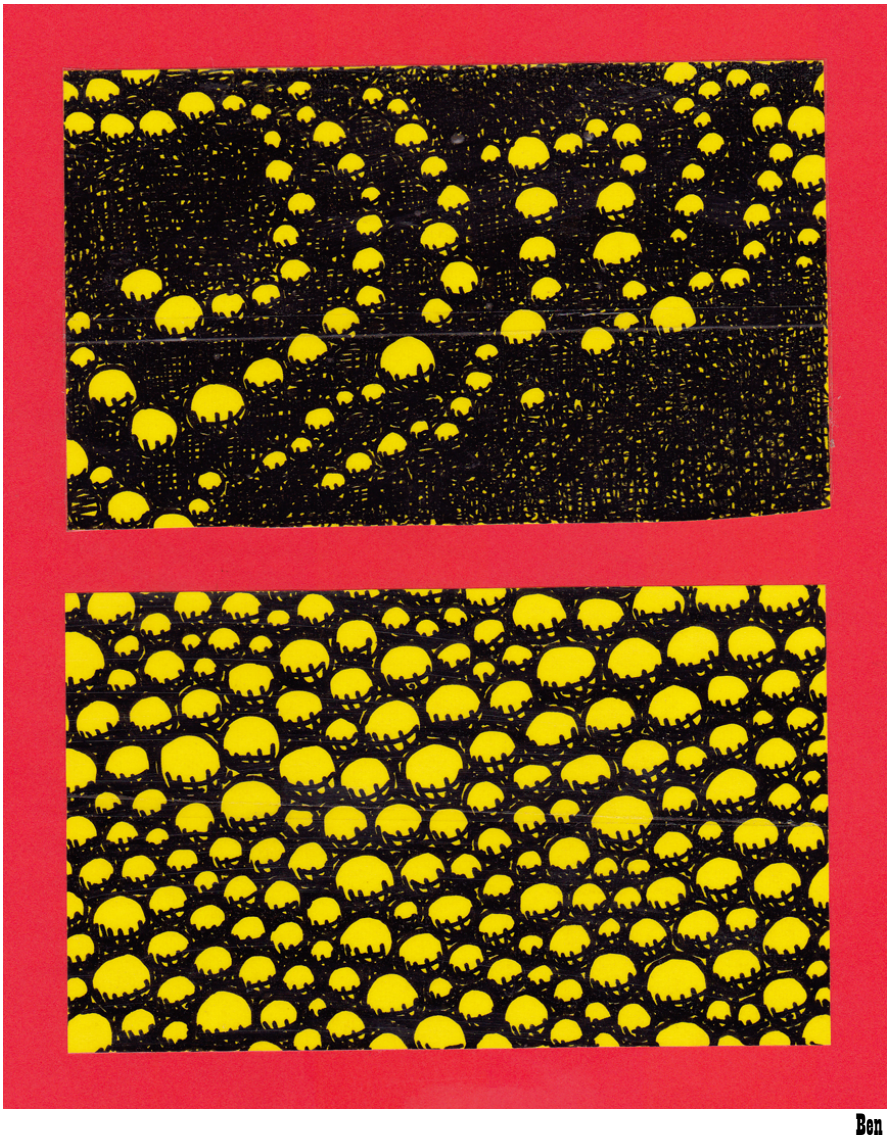
Wang, L. K., Z. X. Feng, *et al.* (2010). "DEGseq: an R package for identifying differentially expressed genes from RNA-seq data." Bioinformatics **26**(1): 136-138.

Warren, C. R., E. Dreyer, *et al.* (2006). "Ecotype adaptation and acclimation of leaf traits to rainfall in 29 species of 16-year-old Eucalyptus at two common gardens." Functional Ecology **20**(6): 929-940.

Weisshaar, B. and G. I. Jenkins (1998). "Phenylpropanoid biosynthesis and its regulation." Current Opinion in Plant Biology **1**(3): 251-257.

Wright, I. J., P. B. Reich, *et al.* (2004). "The worldwide leaf economics spectrum." Nature **428**(6985): 821-827.

Chapitre V - Analyse de la plasticité moléculaire



RNA-Seq reveals genotype-specific molecular responses during seasonal water deficit in eucalyptus

Emilie Villar[1,2,3], Christophe Klopp[4], Céline Noirot[4], Evandro Novaes[5], Matias Kirst[5], Christophe Plomion[2,6], Jean-Marc Gion[1,2§]

[1]CIRAD, UPR39 Genetic Diversity and Breeding of Forest Tree Species, Campus de Baillarguet TA 10C, F–34398 Montpellier Cedex 5, France

[2]INRA, UMR1202 BIOGECO, F-33610 Cestas, France

[3]CRDPI, BP1291, Pointe Noire, République du Congo

[4] Plateforme bioinformatique Genotoul, UR875 Biométrie et Intelligence Artificielle, INRA, 31326 Castanet-Tolosan, France

[5]School of Forest Resources and Conservation, University of Florida, PO Box 110410, Gainesville, USA

[6]Université de Bordeaux, UMR1202 BIOGECO, F-33610 Cestas, France

[§] Corresponding author

E-mail addresses:

EV: villar@cirad.fr

CK: Christophe.Klopp@toulouse.inra.fr

CN: Celine.Noirot@toulouse.inra.fr

EN: evandro@ufl.edu

MK: mkirst@ufl.edu

CP: plomion@pierroton.inra.fr

JMG: gion@cirad.fr

1 Abstract

1.1 Background

In a context of climate change, phenotypic plasticity provides long-lived species, such as trees, with a means of adapting to environmental variations occurring within a single generation. In eucalyptus plantations, water availability is a key factor limiting production. Despite many physiologically investigations, the molecular mechanisms underlying the adaptation of eucalyptus to water shortage remain unclear. In this study, we compared the molecular responses to water deficit during the dry season of two eucalyptus hybrids used in industrial plantations and differing in productivity.

1.2 Results

RNA-Seq, with long reads provided by 454 pyrosequencing of RNA extracted from shoot apices, gave extensive transcriptome coverage, providing a catalog of 129,993 unigenes (49,748 contigs and 80,245 singletons) from the initial sequence of 398 Mb corresponding to 1.14 Mreads. These 454-sequencing data enriched considerably the *Eucalyptus* EST collection with 36,985 unigenes that did not match with previously published *Eucalyptus* EST. Digital analysis of read abundance in 14,460 contigs with more than 10 reads identified 1,280 contigs displaying differential expression between the two genotypes, 155 contigs showing differential expression between treatments (irrigated vs. non irrigated conditions), and 274 contigs with significant genotype x treatment interaction (GxT) effects. The more productive genotype displayed a larger set of genes with a more marked molecular response than the less productive genotype. Stress signal transduction seemed to involved different pathways between the genotypes, suggesting that water shortage induces different cellular stress cascades, perceived differently by the two genotypes. Consecutively, response of structural proteins was also widely different between genotypes: the more productive genotype decreased expression of genes related to photosystem, transport and secondary metabolism, whereas genes related to primary metabolism and cell organisation were over-expressed.

1.3 Conclusions

For the more productive genotype, the ability to express a larger set of genes with higher expression differences appears a key element in the maintenance of biomass during the dry season. The strategy developed by this genotype may involved a decrease of photosynthetic

activity during the dry season associated with resources reallocation through major changes in the expression of primary metabolism associated genes. Further efforts will be needed to assess the adaptive nature of the genes highlighted in this study.

2 Background

Planted forests constitute only 7% of the global forest area, but contribute to a much higher proportion of overall forest goods and services (e.g. up to 35% of industrial roundwood supply). In the context of climate change, the adaptation of planted forests is thus a key concern for sustainability of the forestry sector. The adaptation of industrial plantations to present and future environmental conditions (including extreme weather events) depends on a number of factors, including the genetic diversity of the material used for reforestation and the phenotypic plasticity of individual genotypes. Genetic diversity ensures that forest trees can survive, adapt and evolve under changing environmental conditions [1, 2], whereas phenotypic plasticity constitutes a shorter term response to environmental changes at the individual level of particular importance in long-lived organisms, such as trees [2, 3].

Eucalyptus is one of the key genera among planted trees. This genus is one of the most important hardwood fibre crops planted worldwide (19 Mha according to [4]). Its wood is used for a range of industrial applications: timber, firewood, and pulpwood being the main products. This tree grows rapidly and is highly adaptable. These properties have led to its introduction at many sites worldwide, at latitudes extending from southern Europe to South Africa. In its natural range (Australia), it is also found in a diverse spectrum of ecological niches. The genetic diversity of *Eucalyptus* has been studied extensively, particularly in commercial species, with neutral markers [5-11] and genes potentially involved in adaptive traits [12-14]. The authors of these studies reported a remarkable level of nucleotide diversity, a prerequisite for adaptability. However, a certain degree of phenotypic plasticity is likely to ensure the better adaptation of individual genotypes to uncertain environmental conditions [15]. This adaptive component is of particular importance in the context of clonal forestry, in which *Eucalyptus* is a major player.

Ecophysiological studies have shown that water is the principal factor limiting stem growth in *Eucalyptus* (e.g. [16, 17]). Moreover, some studies have reported that eucalyptus genotypes differ in terms of their capacity for phenotypic modification in response to water deficit [18-20]. Several physiological mechanisms for coping with drought have been described in these

species: i) the regulation of transpiration to decrease water loss [21], ii) resource reallocation from the shoot to the root, to increase water uptake [17], and iii) adjustment of osmotic potential [22] or protection against reactive oxygen species, to prevent damage due to stress [20]. At the molecular level, Tsuchihira *et al.* [23] have shown that drought tolerance is correlated with the aquaporin content of leaves. Shvaleva *et al*. [20] also highlighted the importance of glutathione reductase, ascorbate peroxidase and superoxide dismutase for reactive oxygen species detoxification in roots and of the carotenoids involved in protecting leaves against drought stress. Drought tolerance mechanisms have been described in detail at the molecular level for model plants, such as *Arabidopsis* (for annuals, e.g. [24-26]) and *Populus* (for perennials, e.g. [27-29]), but little is known about the molecular basis of drought tolerance in *Eucalyptus*, particularly in field conditions.

Next-generation sequencing (NGS) technologies are providing new opportunities for studies of the molecular plasticity in response to water deficit. The tremendous capacity of NGS methods makes it possible to sequence a large proportion of the transcriptome, equivalent to several thousand expressed sequence tags (ESTs), which is particularly useful in non model organisms for which few genomic resources are available [30]. Moreover, NGS appears to be suitable for transcript profiling, combining the high throughput of serial analysis of gene expression (SAGE) with the accuracy of EST sequencing [31]. These techniques have been widely used for transcriptome profiling, particularly for studies of biotic [32] and abiotic [25] stress responses, and the characterisation of developmental processes [33]. These techniques provide considerable sequencing depth, which, together with quantitative analysis, makes it possible to visualise variations in expression for a large proportion of the transcriptome [30] and to identify key molecular players responding to environmental variation.

In plants, the shoot apical meristem (SAM) is a key organ in stem development. The SAM initiates phytomers and regulates shoot growth by integrating several signals, such as hormones (ABA, auxins, cytokinins) and transcription (e.g. homeobox) [34]. When plants are subjected to environmental stimuli, the leaf developmental network is adjusted by changes in shoot apex activation [35]. In *Eucalyptus*, EST resources have been developed for various tissues, such as roots, leaves and wood-forming tissues [36-39], but a limited number of genomic resources are available for shoot apices, despite the important role of this organ in plant organogenesis.

In this study, we compared transcript profiles in the shoot apices of two eucalyptus genotypes used in industrial plantations under two watering regimes — irrigated (IR) *versus* non-irrigated (NI) — during the dry season, in a field trial in the Republic of Congo. These two genotypes differ in their growth rates and ecophysiological characteristics at maturity, with one genotype more productive and water use-efficient than the other. We used 454-Roche FLX Titanium methods to sequence non normalised cDNA libraries constructed from shoot tip mRNA. After checking for technical reproducibility, we addressed the following questions: i) Are there molecular differences between genotypes, reflected in the contrasting phenotypes, and do these differences affect specific pathways or have a random effect on the transcriptome? ii) Can we detect molecular plasticity in the response to water shortage during the dry season, and which pathways are affected? iii) Does this plasticity differ between genotypes (i.e. is there any genotype-by-environment interaction?), and which genes or pathways reflect these differences?

3 Methods

3.1 Plant material

We compared the response to water shortage during the dry season of two eucalyptus genotypes, 1-41 (NCBI Taxonomy ID: 764271) and 18-50 (NCBI Taxonomy ID: 765255) used in industrial plantations in the Republic of Congo. Hybrid 1-41 was obtained by open pollination of *E. alba* (the male parent is unknown). Hybrid 18-50 was obtained by controlled pollination of *E. urophylla* (14-36) by *E. grandis* (9-10). These two genotypes differ in their growth rates and water use efficiency (WUE) at maturity (personal communication, Dr. Y. Nouvellon, UMR80, CIRAD, France), 18-50 being superior than 1-41.

3.2 Field experiment

Trees were vegetatively propagated by rooted cuttings and established in a field experiment in Yanika, Republic of Congo (4°20'S, 11°38'E, 50 m above sea level) in June 2007. Trees were planted in blocks of 64 cuttings per genotype and per treatment, including a buffering zone. Two watering regimes were used during the dry season: no-irrigation (NI) and irrigation (IR). Trees were watered with sprinklers, to cover potential evapotranspiration losses, estimated at 3 mm per day. Plant material was sampled, to evaluate the effect of water deficit on above-ground biomass and molecular plasticity, in September 2008, 16 months after the trees were

planted. As the dry season began at about May 15th in 2008, the trees under NI treatment had been subjected to four months without rainfall at this specific sampling point.

3.3 Soil water content

Soil water content (SWC) was measured by time domain reflectometry (TDR; Trase system, Soil moisture, Santa Barbara, CA). Four series of TDR probes per genotype and per treatment were installed horizontally, at six depths (0.15, 0.5, 1, 2, 3 and 4 m). Mean values were calculated from the four replicated measurements at each depth.

3.4 Biomass production

We harvested 11 trees per treatment and divided them into the following compartments: stem, dead branches, living branches and leaves. Each compartment was weighed in the field. Representative subsamples of each compartment were then harvested, and weighed before and after drying at 65 °C to constant weight. Water content was calculated for each of these subsamples and used to estimate total dry biomass for each compartment. Total above-ground biomass (the sum of all the compartment) was analysed by two-way ANOVA, with R (R Development Core Team), according to the following model: $Y = G + T + GxT$, where Y is above-ground biomass, G is the genotypic effect, T, the treatment effect and GxT the interaction between genotype and treatment. Bonferroni tests were used for the comparison of means.

3.5 cDNA synthesis

The experimental design, from tissue sampling to library construction and sequencing, is illustrated in figure 1. Shoot apices were collected from three trees for each genotype and treatment and immediately frozen in liquid nitrogen. Two RNA extractions of three apices per tree were then performed, as previously described by Reid *et al.* [40]. RNA samples were treated with Turbo RNase-Free DNase (Ambion, Austin, TX, USA) and purified with the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). RNA concentration and quality were analysed with an Agilent Technologies 2100 Bioanalyser (Agilent Technologies, Mississauga, ON, USA) and an ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE, USA). The three RNA preparations per treatment (corresponding to three trees), were pooled in equal proportions, providing templates for cDNA libraries S1-S8 (figure 1). Full-length cDNA was obtained from 1 µg of RNA, with the Smart cDNA Library Construction Kit (Clontech, Mountain View, CA, USA), according to the manufacturer's instructions. We amplified the cDNA with PCR Advantage II Polymerase (Clontech, Mountain View, CA, USA), over 16

cycles (7 s at 95 C, 20 s at 66 °C, and 4 min at 72 °C). This cDNA amplification procedure was repeated several times for each sample, with pooling to give a total of 6 µg of cDNA fragments of more than 1,000 bp quantified with an Agilent Technologies 2100 Bioanalyser. Eight cDNA libraries (S1-S8) were constructed, giving two biological replicates for the two genotypes (1-41, 18-50) subjected to the two watering regimes (IR, NI).

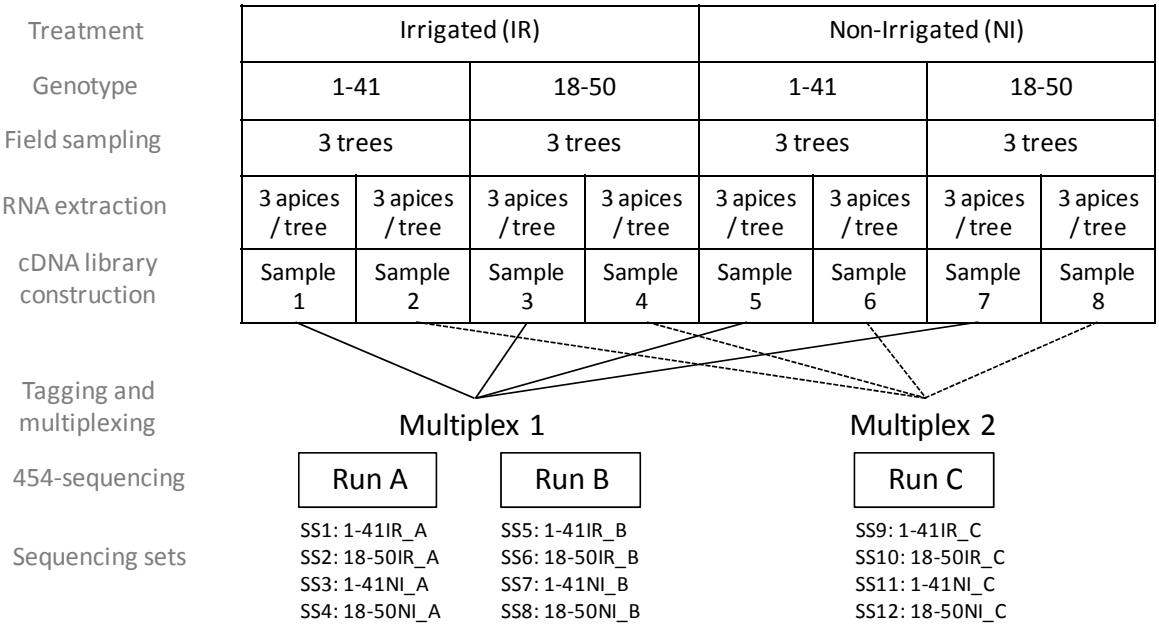


Figure 1 - Procedure used, from tissue sampling to sequencing
Two genotypes (1-41 and 18-50) were subjected to two watering regimes (IR and NI). Shoot apices from three trees per genotype and per treatment were collected in the field. Total RNA was extracted from three apices per tree. Two replicate RNA extractions were carried out for the construction of two independent replicate cDNA libraries per genotype and per treatment (resulting in 8 templates for cDNA library construction). For sequencing, each cDNA library was tagged, and two independent multiplexes were created by pooling one sample for each genotype and treatment combination. Multiplex #1 was sequenced with two 454-Roche FLX Titanium half-runs, resulting in eight sequencing sets, whereas multiplex #2 was sequenced with one half-run and resulted in four sequencing sets.

3.6 Library construction, 454-sequencing, sequence cleaning and assembly

We nebulised 5 µg of each cDNA sample to a mean fragment size of 650 bp and ligated it to an adaptor, according to standard procedures [41]. Each cDNA library was tagged with MID adaptators, and two independent multiplexes were created by pooling one sample for each genotype and treatment. Multiplex #1 comprised samples S1, S3, S5 and S7, whereas muliplex #2 comprised samples S2, S4, S6 and S8. One half-run (run A) of sequencing was initially carried out for multiplex #1 on a GS-FLX Titanium platform (454 Life Science, Brandford, CT, USA) at Cogenics (Meylan, France). Two half-runs of sequencing for multiplex #1 (run B) and multiplex #2 (run C) were then performed by Agencourt (Beverly,

MA, USA) on a GS-FLX Titanium sequencer. Base calling with GS-FLX software generated 353,344 high-quality reads for the first half-run and in 785,322 reads for the second complete run. Sequences were deposited in the NCBI short-read archive (SRA) under accession number SRA012867.2 (figure 2).

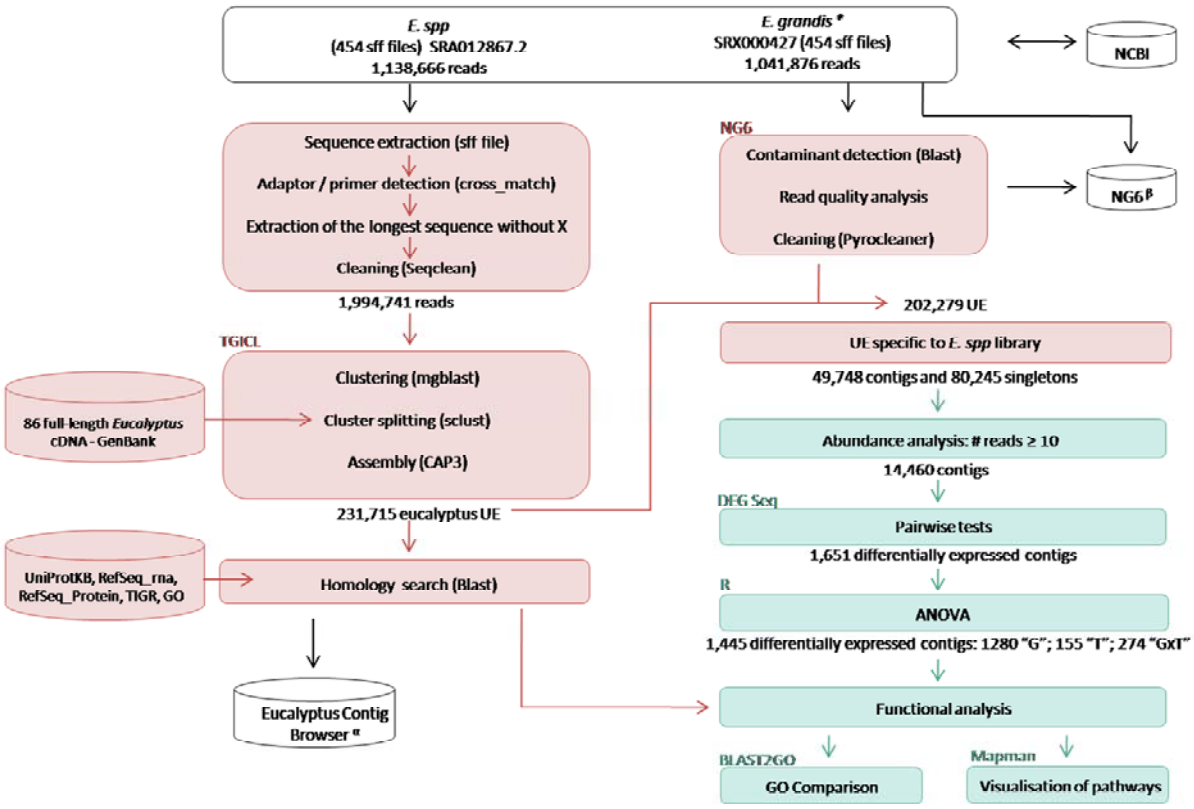


Figure 2 – Pipeline used for sequence analysis
White boxes correspond to sequence generation, red boxes correspond to library assembly and annotation and green boxes correspond to abundance and functional analysis. The software suite or program used is indicated in the upper part of the box and the results of each step of the analysis is shown under the box. UE: unigene element; "G": genetic effect, "T": treatment effect and "GxT": genotype by treatment interaction effect.
**E. grandis* .sff files were obtained from Novaes *et al*. (2008).
α: available at http://genotoul-contigbrowser.toulouse.inra.fr:9092/Eucalyptus/index.html
β: available at http://vm-bioinfo.toulouse.inra.fr/ng6/, user = euca ; password = euca33bio!

Sequences were assembled as previously described [42], with TGICL [43], using the 12 sets of sequencing data from this study and the four sets of sequencing data obtained for *E. grandis* [accession number SRA001122] by Novaes *et al.* [13]. The 454-sequencing reads (1,138,666 from this study and 1,041,876 from Novaes *et al.* [13]) were screened by cross_match (http://bozeman.mbt.washington.edu/phrap.docs/phrap.html) for primers and adaptors and then masked. For each 454-sequencing read, the longest non-masked region was extracted and further cleaned up with SeqClean (http://compbio.dfci.harvard.edu/tgi/). The

shorter regions were discarded to eliminate potential chimeras. In parallel, all reads were stored in the NG6 system (http://vm-bioinfo.toulouse.inra.fr/ng6/, user name: euca, password: euca33bio!) and three kinds of analysis were performed for each of the 16 sequencing sets, as previously described [42]: Blast search for *E. coli*, phage and yeast contaminants, read quality analysis, removal of sequences that were too long or too short (more than two standard deviations away from the mean) and of dirty sequences (sequences with more than 4% N), low-complexity sequences and duplicated reads, with the Pyrocleaner program. Only unigene elements (UE) resulting from sequences generated in this experiment (the *E.spp* sequencing set) were used for digital gene expression analysis.

3.7 Digital gene expression analysis

Contigs with less than 10 reads for the 12 sequencing sets generated in this study were eliminated from further statistical analyses. For the 14,460 remaining contigs, the numbers of reads per sequencing set and per contig were used to assess gene abundance. Two types of statistical analysis were performed. First, pairwise comparisons were carried out between genotypes (18-50 vs. 1-41 sequencing sets, irrespective of treatment) and between treatments (IR vs. NI sequencing sets, irrespective of genotype). Four additional comparisons were carried out for each genotype and each treatment, as follows: IR vs. NI for genotypes 1-41 and 18-50, 1-41 vs. 18-50 for treatments IR and NI. Statistical tests, based on the use of the MARS method in the DEGseq package [44] were performed to assess differential expression [44]. Second, two-way analysis of variance (ANOVA) were performed on contigs, making use of the three replicates (run A, B and C) per treatment to test genotype (G), treatment (T) and genotype-treatment interaction (GxT) effects. Gene abundances were normalised by dividing by sequence length and the total number of sequences in each sequencing set. Contigs with a q-value <0.05 in the DEGseq test [45] and with p-values <0.05 after ANOVA were considered to be differentially expressed and were extracted for further analysis. The 14,460 genes analysed were classified into four classes: not significantly differentially expressed (NS), and showing genotype (“G” contigs), treatment (“T” contigs) or genotype x treatment (“GxT” contigs) effects. For the comparison of expression levels, we used log2-transformed ratios between contig abundances in the various contrasts used in the DEGseq analysis.

3.8 Functional annotation of differentially expressed genes

Contigs were assigned a putative function by BLASTX [46], using various public databases: UniProtKB/Swiss-Prot Release 57.1 from April 14th 2009, RefSeq Protein Release 34 from March 6th 2009, Pfam Release 23.0 from July 20th 2008, with an e-value cut-off value of of $10e^{-5}$, and TIGR Arabidopsis_thaliana release_14; TIGR Helianthus_annuus release_6; TIGR Populus release_4; TIGR Picea release_3; TIGR Vitis_vinifera release_6 with an e-value cut-off of $10e^{-2}$. Gene Ontology terms were assigned via the UniprotKB accession and clustered with Blast2GO [47]. The differential distributions of each class of effect (T, G, GxT and NS) between Biological Processes, Molecular Functions and Cellular Components were assessed in Fisher's exact tests, with the significance threshold set at 0.05. Pathway analysis was carried out with Mapman [48]. Differentially expressed contigs were assigned to functional categories (or bins) by Mercator (http://mapman.gabipd.org/web/guest/mercator). A dedicated pathway map was created to represent most of these contigs. The Wilcoxon rank sum test was used to identify differentially regulated bins.

4 Results

4.1 Monitoring of soil water content

A factorial design including two genotypes (1-41 and 18-50) and two water regimes (irrigated IR vs. non-irrigated NI) was established in a field trial in one of the main areas of eucalyptus plantations in the Republic of Congo, and was followed over a period of two years. Soil water content (SWC) was monitored throughout the experiment, at six depths, from 0.15m to 4 m, to follow water availability in the four sets of conditions studied. In this study, we focus on the effect of the second dry season (ie. four months without rainfall) on biomass production and molecular response.

In the NI treatment, SWC varied from 4.5% to 8% and no significant difference was found between the two genotypes (figure 3). SWC values were close to wilting point (pF 4.2), at which plants cease to be able to absorb soil water. In the IR treatment, SWC was much higher (ranging from 11% to 17%), and exceeded field capacity (pF 2), indicating that water was not a limiting factor for tree growth. SWC was also higher in the area surrounding genotype18-50 than in that surrounding 1-41, except at two depths (1 m and 4 m). These data suggest that these two genotypes absorb water from different depths and have root systems that have developed differently.

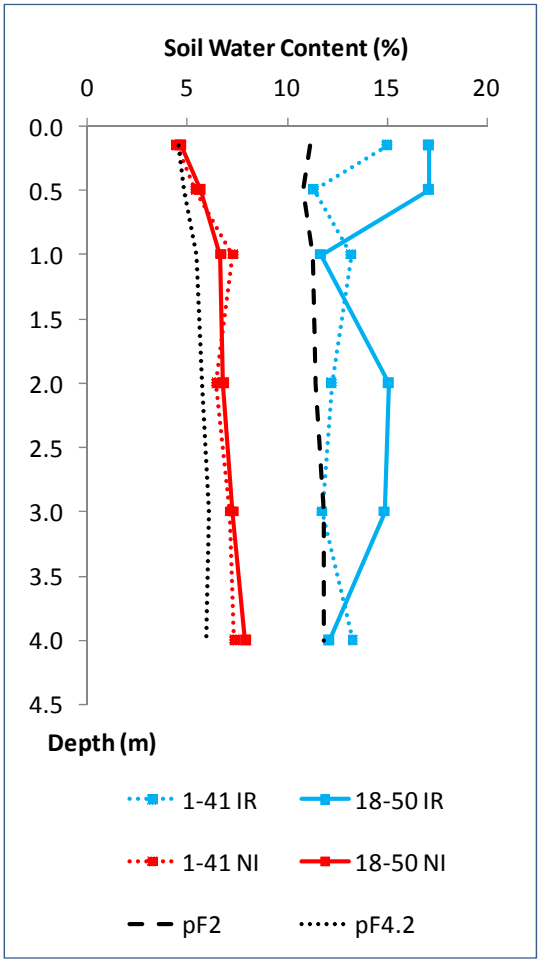


Figure 3 – Profile of soil water content (SWC)
SWC at six depths, for both genotypes (1-41 and 18-50), subjected to two water regimes (irrigated: IR and not irrigated: NI). SWC at wilting point (pF4.2) and at field capacity (pF2) were calculated from data obtained at a site close to the experimental field (Laclau, personal communication).

4.2 Effect of water deficit on biomass production

Genotype ($p<0.0001$) and treatment ($p<0.001$) had significant effects on above-ground biomass production (stem, dead branches, living branches and leaves) (figure 4). Mean biomass was higher for genotype 18-50 (12.8 kg) than for genotype 1-41 (8.6 kg) confirming earlier findings at maturity. Unsurprisingly, mean biomass was higher for the IR treatment (11.8 kg) than for the NI treatment (9.7 kg). Two-way ANOVA showed that the GxT interaction effect was not significant (Additional file 1), but relative biomass loss was nonetheless lower for genotype 18-50 (12%) than for genotype 1-41 (24%). This hypothesis was supported by Bonferroni pairwise comparison tests, which showed that the difference in biomass between the IR and NI treatments was significant only for genotype 1-41. These

results suggest that genotype 18-50 can continue to grow during the dry season, whereas the growth of genotype 1-41 is affected by water shortage.

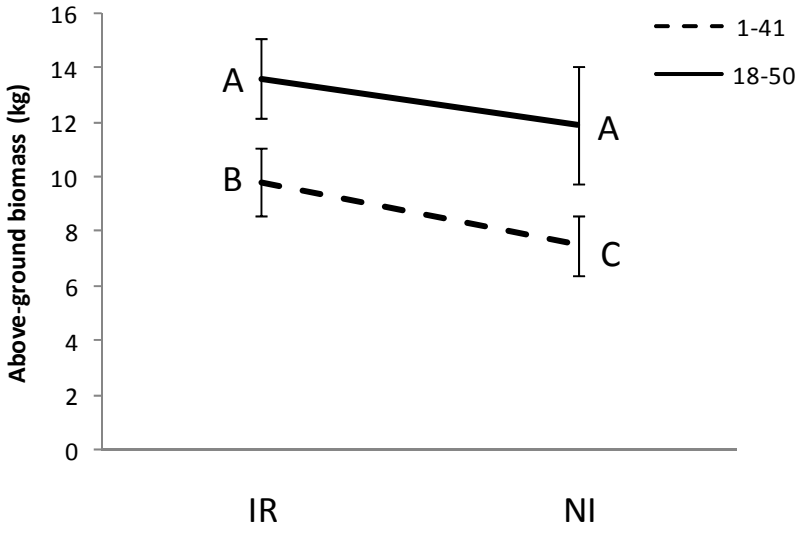


Figure 4 - Above-ground biomass
Means and standard errors of above-ground biomass were calculated for 11 trees per treatment, for both genotypes (1-41 and 18-50), subjected to two water regimes (irrigated: IR and not irrigated: NI). Letters indicate the groups obtained in Bonferroni tests for multiple pairwise comparisons.

4.3 Sequencing of the *Eucalyptus* shoot apex transcriptome

We chose the shoot apex as the main organ for studies of the molecular response of the two eucalyptus genotypes to water deficit, due to its role in shoot organogenesis. Shoot apices were pooled, as described in the methods section, from three trees in four sets of conditions (2 genotypes x 2 treatments). Three half-runs (A, B, C) of 454-Roche FLX Titanium sequencing provided a total of 1,138,666 reads (398 Mb) corresponding to 353,344 (A), 405,223 (B) and 380,099 (C) reads. The mean read length was 334 bp for run A, 369 bp for run B and 344 bp for run C (Table 1).

Reads were slightly shorter for run A, with a major read-length class of 400 to 420 bp, whereas runs B and C were characterised by reads of 460-480 bp in length (Additional file 2). Contig lengths were maximised to optimise annotation, by assembling all the reads (1,138,666) generated in this study (*E. spp* sequencing set) with other GS-20 and GS-FLX 454-reads (1,041,876 reads) for various organs of *E. grandis* genotypes (*E. grandis* sequencing set; [13]).

Table 1 - Summary statistics for the three 454-sequencing half-runs

Sample	Genotype	Treatment	Sequencing set	Runs	# of Reads	# of Reads	# of bp	Average length of reads (bp)	Average length of reads (bp)
1	1-41	IR	SS1: 1-41IR_A	run A	353,344	58,921	19,313,318	328	334
3	18-50	IR	SS2: 18-50IR_A			92,165	30,365,092	329	
5	1-41	NI	SS3: 1-41-NI_A			95,500	31,350,808	328	
7	18-50	NI	SS4: 18-50NI_A			106,758	37,126,644	348	
1	1-41	IR	SS5: 1-41IR_B	run B	405,223	139,137	51,085,203	367	369
3	18-50	IR	SS6: 18-50IR_B			112,051	41,110,040	367	
5	1-41	NI	SS7: 1-41-NI_B			59,907	21,833,839	364	
7	18-50	NI	SS8: 18-50NI_B			94,128	35,374,789	376	
2	1-41	IR	SS9: 1-41IR_C	run C	380,099	93,651	32,338,261	345	344
4	18-50	IR	SS10: 18-50IR_C			90,511	31,595,445	349	
6	1-41	NI	SS11: 1-41NI_C			83,530	27,929,481	334	
8	18-50	NI	SS12: 18-50NI_C			112,407	38,802,862	345	
TOTAL				1.5 run GS-FLX Titanium	1,138,666	1,138,666	398,225,782	350	350

Figure 2 summarises the various stages in sequence analysis, corresponding to data cleaning, read assembly, annotation and abundance analysis. After the cleaning step (removal of vectors and adaptors), 1,994,741 reads were available for assembly.

The assembly of these sequences with TGICL generated 231,715 unigene elements (UE) comprising 80,854 contigs and 150,861 singletons. Further searches of low-quality sequences and duplicated reads on the NG6 platform resulted in a total of 202,279 UE (69,584 contigs and 132,690 singletons), from which 129,993 UE (49,748 contigs and 80,245 singletons) were identified in the *E. spp* sequencing set (this study) and used for digital gene expression analysis (Table 2).

Table 2 – Assembly statistics from TGICL (I), and figures obtained after Pyrocleaner analysis (II), for the set of sequences reported here (III).

	Tgicl (I)	**Pyrocleaner (II)**	***E. spp* (III)**
# Contigs	80,854	69,584	49,748
# reads in contigs	1,843,806	1,386,859	851,751
Average length of all contigs (bp)	552	608	734
# Large Contigs >500 bp	34,076	33,962	32,694
Average length of large contigs	900	901	912
# Singletons	150,861	132,695	80,245
# UE	231,715	202,279	129,993

4.4 Contribution of the three half-run replicates

In total, 90,579 UE (70% of *E. spp* UE) did not match the *E. grandis* sequence reads from Novaes *et al.* [13]. Most of these sequences were singletons (71,761), although some were contigs (18,818), corresponding to 27% of the *E. spp* contigs. A Blast homology search (cut-off: 10^{-50}) of published eucalyptus databases (ESTs from GenBank released in April 2010;

454-ESTs generated by Novaes *et al*., [13]; 454-ESTs from JGI from *E. globulus* xylem and leaf tissues; Illumina contigs generated by Mizrachi *et al*., [39]) showed that 36,985 UE (comprising 5,596 contigs and 31,389 singletons) did not match any known sequence. These new genomics resources in *Eucalyptus* therefore extend the list of genes known to be expressed in *Eucalyptus*, which can be considered as a significant resource at the of the eucalyptus genome sequence.

Due to the smaller number of reads in run A, the total number of UE generated from run A was also smaller (58,763) than that generated from the other two runs (67,467 in run B and 67,756 in run C). Runs A, B and C accounted for 72%, 78% and 78% of *E. spp* contigs, respectively. Each supplementary half-run produced between 16 and 21% new contigs for a second half-run, and between 5 and 7% for a third half-run (Additional file 3). Vega-Arreguin *et al.* [49] reported similar trends in maize, with a plateau of gene representation reached after the third successive GS-20 454-sequencing run. The number of reads generated was therefore considered sufficient to sample the most expressed genes.

The Pearson's correlation coefficients (figure 5) obtained for read frequencies in the 12 sequencing sets from *E. spp* (87% on average) and the four sequencing sets from *E. grandis* (86% on average) were similar. However, the mean correlation was much weaker (52%) between the *E. spp* and *E. grandis* sequencing sets, suggesting that different fractions of the transcriptome had been sampled from these two studies. As expected, correlations were stronger between replicates (92%; illustrated by squares in figure 5) than between different samples (86%**)** in *E. spp* sequencing sets, suggesting a high level of technical repeatability for this experiment.

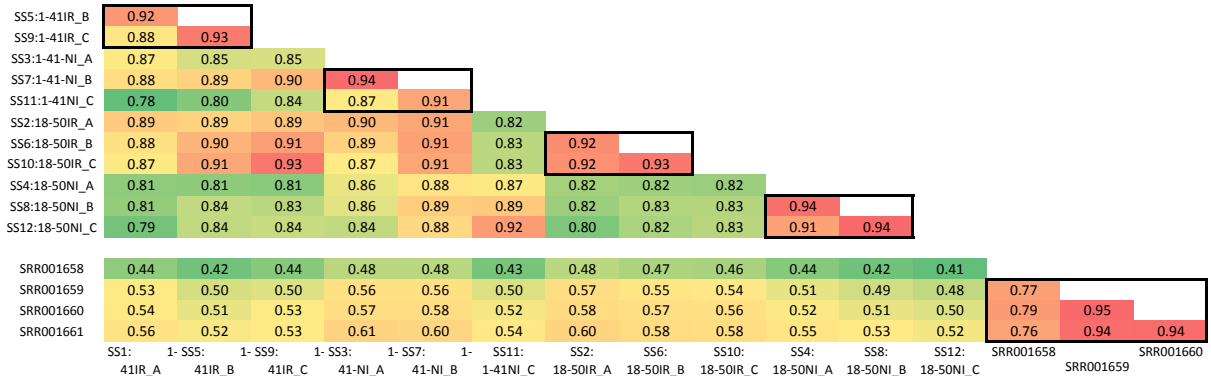


Figure 5 – Pearson correlation coefficient for read frequencies (within contigs) between sequencing sets. SS1-SS12 are sequencing sets from *E.spp*; SRR001658-SRR001661 are sequencing sets from *E. grandis* (Novaes *et al.,* 2008). The colour scale indicates the strength of the correlation, from red (strongest correlations) to green (weakest correlations). All the correlation coefficients were significant, with a p-value <0.001.

Within conditions, correlations between the 1-41 and 18-50 sequencing sets were strong and similar between treatments: 90% for the IR treatment and 88% for the NI treatment, suggesting that these two genotypes displayed similar patterns of gene expression when placed in the same environmental conditions. Correlations between the IR and NI sequencing sets were slightly weaker and a stronger contrast was observed between genotypes: on average, 86% for genotype 1-41 and 82% for genotype 18-50. The weaker correlations obtained for 18-50 suggest that this genotype had a more pronounced response to water deficit than genotype 1-41.

4.5 Homology search

BLAST search results are summarised in table 3.

Table 3 – Annotation results for protein hits, nucleic acid hits, and Gene Ontologies (GO): Biological Process (BP), Cellular Component (CC) and Molecular Function (MF)

	E.spp **Sequencing set**	**Protein hits**	**Nucleic acid hits**	**GO-BP**	**GO-CC**	**GO-MF**
# UE	129,993	66,135 (51%)	44,652 (34%)	32,835 (25%)	30,836 (24%)	33,222 (26%)
# Contigs	49,748	34,951 (70%)	37,383 (75%)	18,091 (36%)	17,242 (35%)	18,355 (37%)
# Singletons	80,245	31,184 (39%)	7,269 (9%)	14,744 (18%)	13,594 (17%)	1,4867 (19%)

% of sequences related to E. spp. libraries is shown in brackets

A functional annotation was obtained more frequently for contigs (70% of the contigs harboured sequences similar to those in protein databases, and 75% displayed similarity to sequences in nucleic acid databases) than for singletons (39% and 9%, respectively). These differences were expected, given the longer mean length of the contigs (734 bp) than of singletons (319 bp), and their lower frequency. The best nucleic acid sequence matches were obtained with *Arabidopsis* (55% of annotated sequences), followed by *Vitis* (25%), and *Populus* (6%). The greater similarity to *Arabidopsis* genes may be due to the closer phylogenetic relationship of *Eucalyptus* to *Arabidopsis* (both belonging to eurosid II) than to *Populus* (eurosid I, [50]). The *Arabidopsis* genome has also been annotated in a more detailed manner than the *Populus* genome. Interestingly, the similar characteristics of eucalyptus and poplar in terms of their tallness, longevity and secondary growth, do not seem to play a major role in the similarity of the sequences transcribed.

According to Gene Ontology (GO) classification, 38,190 UE (25% of the *E.spp* sequencing set UE) were associated with at least one biological process (BP), molecular function (MF) or cellular component (CC). The proportions of UE annotated in each category were generally similar to those obtained in *Arabidopsis* (Additional file 4), suggesting that the *E.spp*

sequencing sets studied here will guaranty the analysis of gene expression for a broad range of functional categories.

4.6 Transcript abundance analysis

After the trimming of contigs represented at least by 10 reads in all the *E.spp* sequencing sets, 14,460 contigs remained for abundance analysis. Two statistical tests were performed in series to detect differences in the expression levels of these 14,460 contigs, distributed in four sequencing sets(1-41IR, 1-41NI, 18-50IR and 18-50NI), with three replicates per sequencing set. First, the DEG-seq test [44], which provides powerful pairwise comparisons between sequencing sets, highlighted 1,651 differentially expressed contigs (FDR ≤ 5%). Two-way ANOVA was then performed to assess the effects of the two main factors (G and T) and their interaction (GxT) on contig abundance. This analysis identified 1,445 contigs with at least one significant effect (p-value ≤ 5%; figure 6, additional file 5).

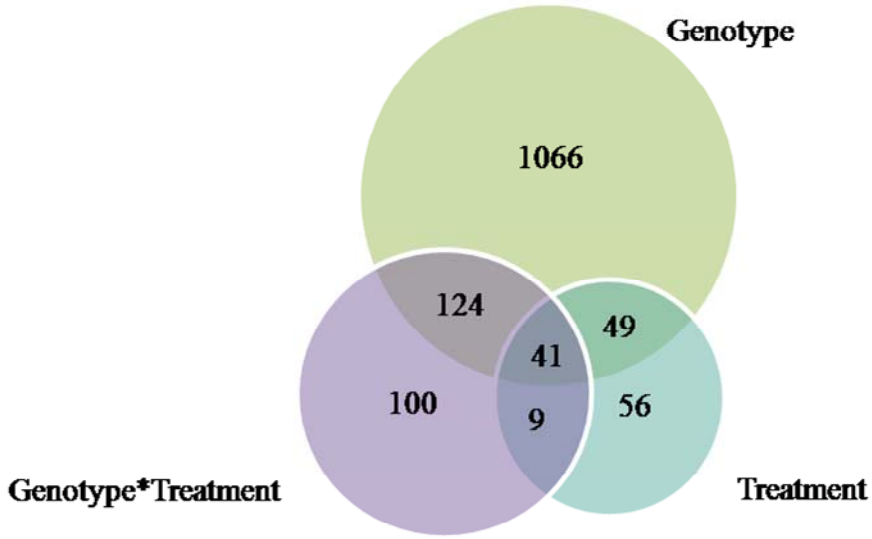


Figure 6 – Differentially expressed contigs
Venn diagram indicating the number of differentially expressed contigs showing a G, T and/or GxT effect, for the 14,460 contigs analysed.

Most of the differentially expressed contigs (1,280) showed a genotype effect ("G contigs"), with 624 "G contigs" overexpressed in genotype 1-41 vs. 656 overexpressed in genotype 18-50 (figure 7A). Of the 656 contigs overexpressed in genotype 18-50, 289 (44%) appeared to be expressed only in 18-50 (with no corresponding reads in the 1-41 sequencing sets) whereas only 55 contigs (9%) showed the reverse trend, suggesting that 18-50 may express of a larger set of genes or different splicing variants.

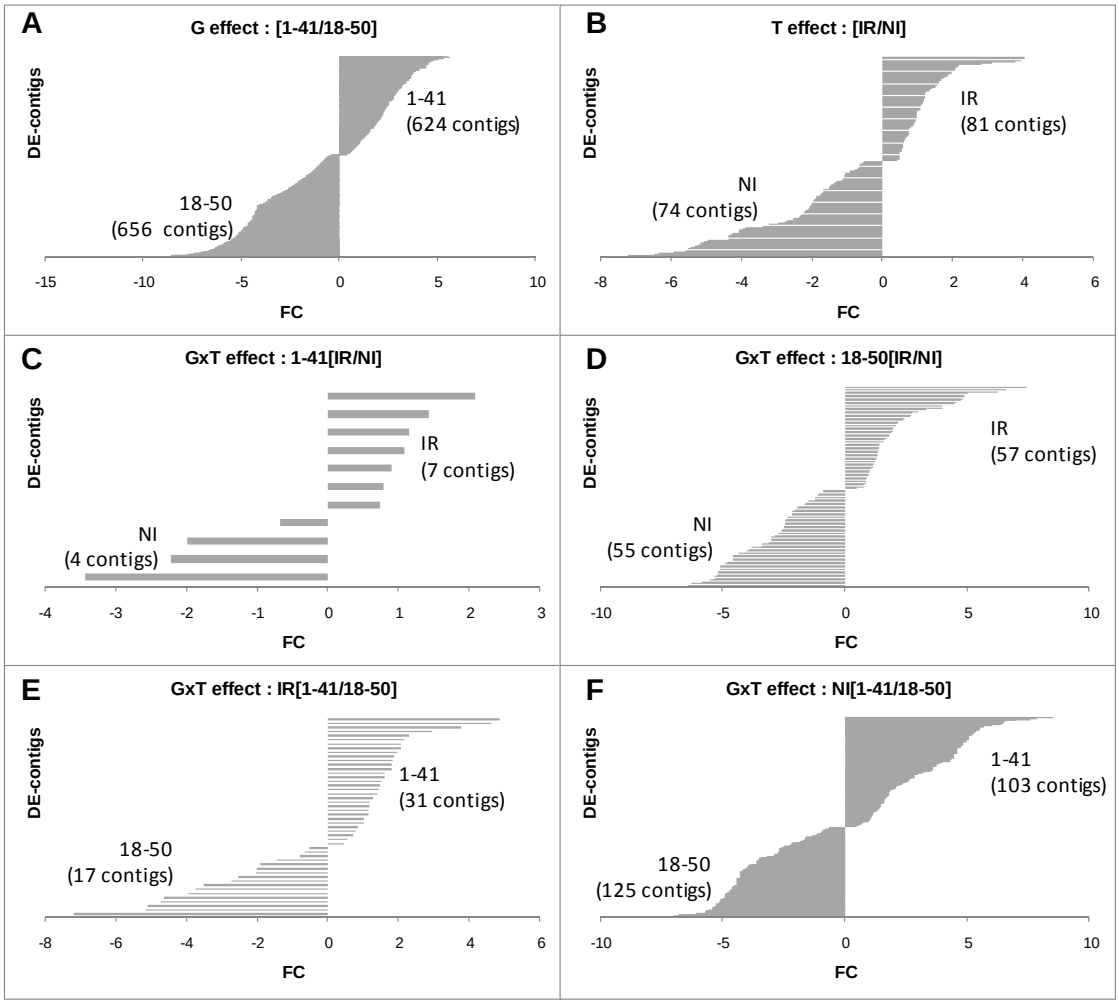


Figure 7 – Fold-change (FC) distribution
Log2-transformed FC were calculated with DEGseq and plotted for each class of effect. (A) Genotype effect: contigs overexpressed in genotype 1-41 are presented with positive values (+), whereas contigs overexpressed in genotype 18-50 have negative values (-). Similarly, (B) treatment effects: IR (+) and NI (-); (C) genotype x treatment effect for genotype 1-41: IR (+) and NI (-); (D) genotype x treatment effect for genotype 18-50: IR (+) and NI (-); (E) genotype x treatment effect for treatment IR: 1-41 (+) and 18-50 (-); (F) genotype x treatment effect for treatment NI: 1-41 (+) and 18-50 (-). The number of contigs overexpressed is shown in brackets.

A total of 155 "T contigs" (figure 7B) were found, with similar numbers overexpressed in the two treatments (81 in IR and 74 in NI). Thirteen "T contigs" were expressed only in NI conditions, whereas all the "T contigs" overexpressed in IR conditions were also found in NI

sequencing sets, suggesting that few "specific genes" are upregulated in response to water deficit but that the set of genes expressed in favourable conditions is also expressed at a lower level in stressed plants.

Finally, 274 contigs were "GxT" contigs. The larger number of "GxT" contigs than of "T contigs" suggests that some of the observed molecular plasticity is under genetic control. Only 11 "GxT contigs" displayed significant differential expression between the IR and NI conditions in genotype 1-41 (figure 7C), whereas 112 "GxT contigs" displayed such behaviour in genotype 18-50 (figure 7D), suggesting a more pronounced response in 18-50. Similarly, 48 "GxT contigs" were differentially expressed between the two genotypes in IR conditions (figure 7E), whereas 228 "GxT contigs" were differentially expressed between the two genotypes for the NI treatment (figure 7F). These results suggest that, despite the rather similar expression patterns for the two genotypes in IR conditions, water deficit induced a molecular response specific to each genotype, reflecting different strategies for dealing with water shortage during the dry season.

4.7 Blast2GO functional analysis

Gene Ontology (GO) analysis was performed on 9,058 contigs (of 14,460 contigs containing more than 10 reads). These contigs were assigned to biological processes (BP), with 11 main subcategories (7,593 contigs), cellular compartments (CC), with five main subcategories (7347 contigs), and molecular functions (MF), with eight main subcategories (7,638 contigs). Consistent with the number of annotations per category (BP,CC,MF), two subcategories were found to be strongly represented: > 65% of the contigs were assigned to cellular and metabolic processes for BP, more than 40 and 85% were assigned to cells and organelles for CC, and > 40% were assigned to binding and catalytic activity for MF. Figure 8 shows the distribution of contigs between these subcategories as a function of the four classes of effects ("NS", "G", "T", and "GxT" contigs).

The homogeneity of the relative abundance of contigs between the "significant" classes ("G", "T", or "GxT") and the "not significant" class ("NS") in each GO category was assessed with Fisher's exact tests. In BP, "G" contigs were overrepresented in four subcategories (response to stimulus, developmental process, death and multiorganism process). "T" contigs were overrepresented in only one subcategory (response to stimulus). Finally, "GxT" contigs were overrepresented in three subcategories (response to stimulus, death and developmental process). These differences in relative abundance suggests that genes related to defence

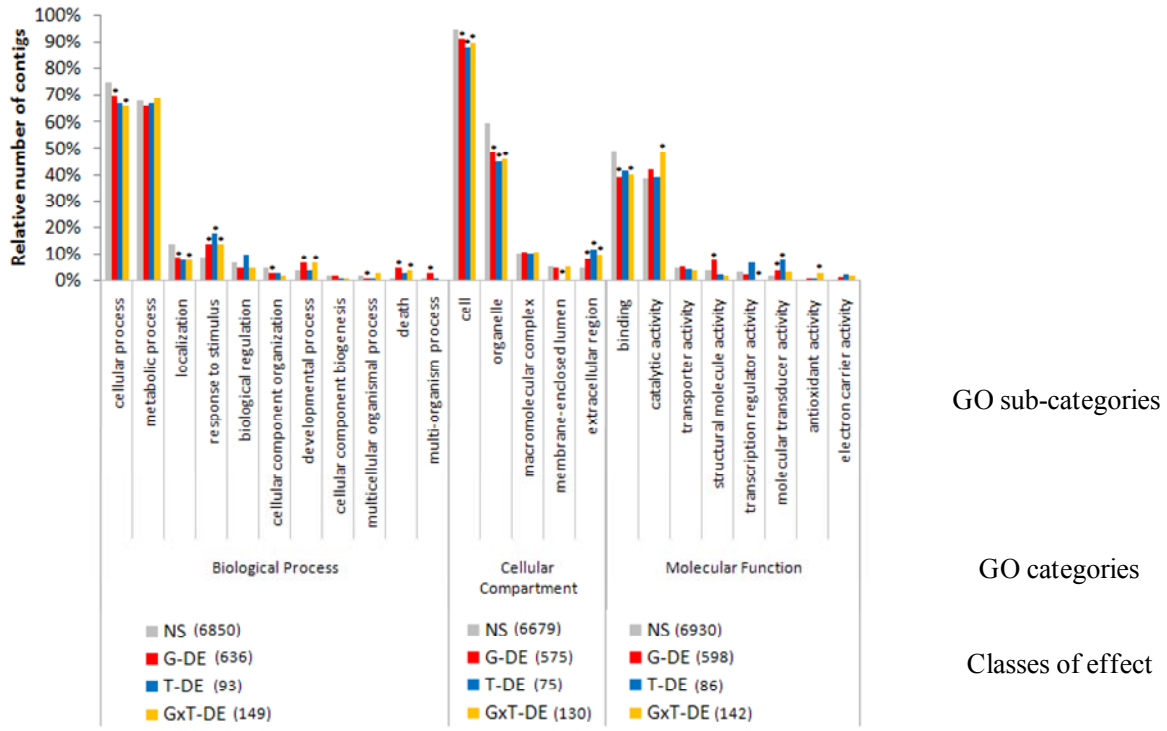


Figure 8 – Distribution of Gene Ontology categories between effect classes
Percentage of contigs annotated for Biological Process, Cellular Component and Molecular Function GO category (level 2). Non-significant contigs (NS) and contigs displaying genotype (G), treatment (T), or genotype x treatment (G*T) effects are shown. The numbers in brackets are the numbers of annotated contigs per class of contigs and GO category. * indicates distributions significantly different between NS and other classes of effect, in Fisher's exact tests with a threshold of 0.05.

reactions are the main contributors to differences between significant and "NS" contigs. For CC, only one subcategory (extracellular region) presented a higher relative abundance for all three types of effect. For MF, "G" contigs were overrepresented in two subcategories (structural molecule activity and molecular transducer activity), whereas "T" contigs were overrepresented in only one subcategory (molecular transducer activity); for"GxT" contigs, two other subcategories (catalytic activity and antioxidant activity) presented higher relative abundances than "NS".

4.8 Analysis of metabolic pathways with MapMan

Of the 1,445 contigs displaying significant differential expression, the 1,280 "G contigs" did not enable characterisation of specific molecular processes (i.e. showing clear coregulation for genes of the same biosynthesis pathway). For 95% of these contigs, many different genes from different molecular processes were activated, depending on the genotype concerned

(Additional file 6). In some instances (5% of "G" contigs), some bins presented specific overexpression in one genotype (Additional file 7). For example, contigs related to ethylene biosynthesis and cell organisation were overexpressed principally in genotype 18-50, whereas contigs related to photosynthesis, nitrilases, calcium signalling and pathogenesis-related protein bins were overexpressed in genotype 1-41. Serine proteases (9 contigs) were expressed more strongly in genotype 1-41 (p-value of 0.038), whereas ubiquitin E3-encoding proteins (23 contigs) were expressed more strongly in genotype 18-50 (p-value of 0.096), suggesting that proteolysis occurred via different pathways in the two genotypes.

Analysis of the metabolic pathways for "T" contigs was limited because of the small number of contigs (155), which were distributed in several bins (Additional file 8). However, the expression of genes related to carbohydrate degradation and ethylene biosynthesis was found to be stronger for the NI treatment, whereas the expression of genes related to ribosomal protein synthesis and cell development appeared to be stronger for the IR treatment (Additional file 7).

Different patterns were observed for "GxT" contigs (figure 9): i) Some pathways (25% of "GxT" contigs) displayed similar patterns in the two genotypes, but with responses of different magnitudes (i.e. scale plasticity, as defined by Lynch and Walsh [51]). Photosystem components tended to be overrepresented in genotype 18-50 in IR conditions, whereas few differences were observed between conditions for genotype 1-41. Conversely, genes related to cell organisation and PR-proteins were more likely to be overexpressed in NI conditions in genotype 18-50 than in genotype 1-41, ii) Most of the pathways (75% of "GxT"contigs) displayed opposite trends in the two genotypes (i.e. a re-ranking interaction effect [51]). For example, in the NI treatment, genes related to phenylpropanoid biosynthesis, auxin biosynthesis, heat stress, and light signalling were overexpressed in genotype 1-41, but underexpressed in genotype 18-50. Conversely, genes related to global primary metabolism

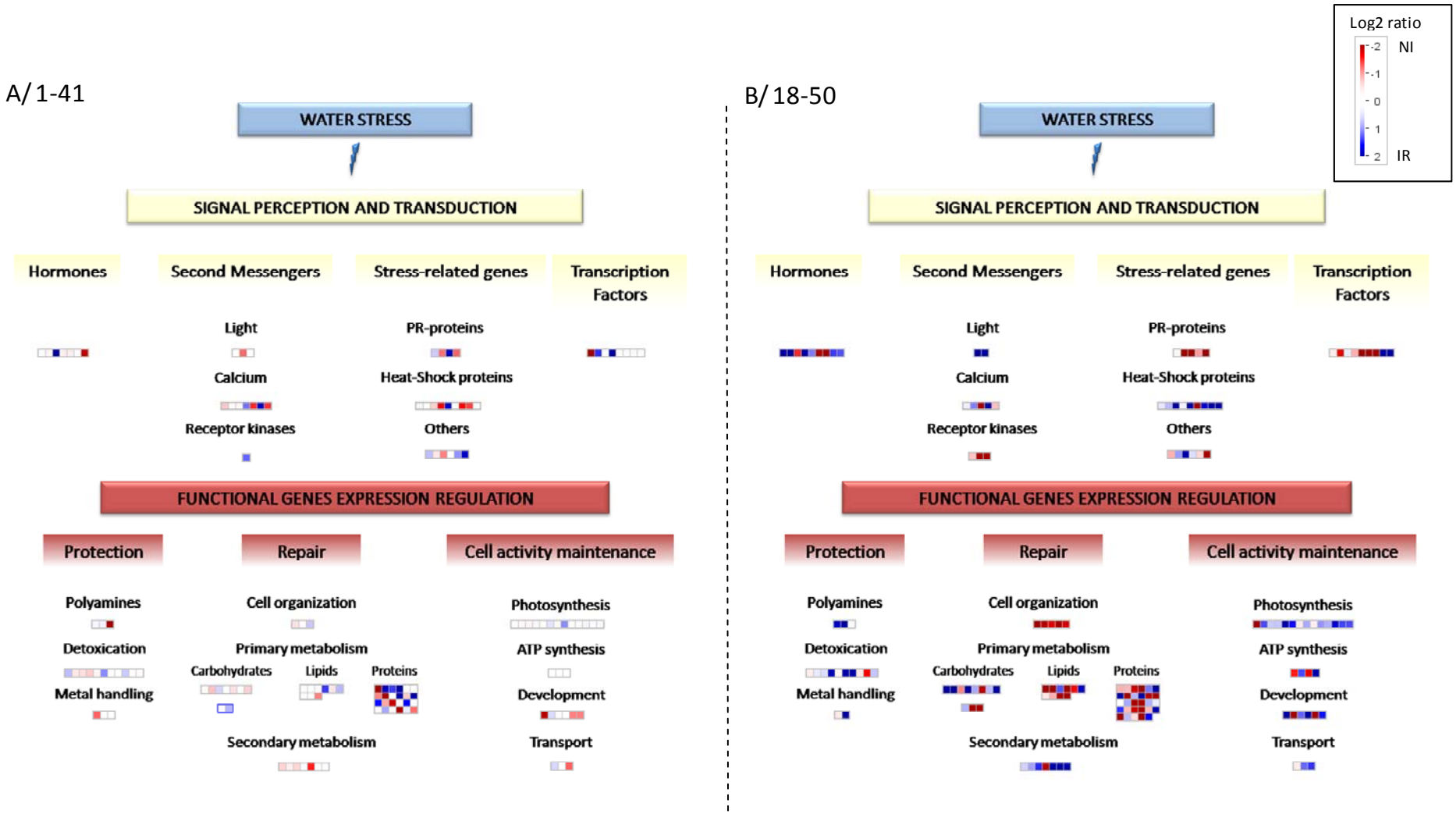


Figure 9 – Distribution of "GxT" contigs between metabolic pathways
Each square represents the log2-transformed ratio of abundance between irrigated (IR) and non-irrigated (NI) treatments for one contig: for genotype 1-41 (A) and 18-50 (B). Contigs in blue were overexpressed for the IR treatment and contigs in red were overexpressed for the NI treatment.

(particularly starch degradation and ribosomal protein synthesis) and receptor kinases were overexpressed in genotype 18-50 but underexpressed in genotype 1-41 in NI conditions (additional file 7). This result strongly suggests that different metabolic pathways and genes were activated in these two genotypes in response to water shortage. Thus, these two genotypes exhibit different molecular strategies to cope with water deficit during the dry season.

5 Discussion

5.1 Gene discovery and expression analysis by NGS

Next-generation sequencing (NGS) technologies are becoming the method of choice for large-scale transcriptome analysis, even for non-model species (e.g. [30, 42], reviewed in [52]). Several technologies have been developed, differing essentially in the number of reads generated and read length (reviewed in [53]), and these techniques make it possible to catalogue the genes expressed and to monitor gene expression.

In *Eucalyptus*, Mizrachi *et al.* [39] generated 3.93 Gb of short reads (36-60 bp) with sequencing-by-synthesis technology from Illumina, and assembled this information *de novo* into 18,894 contigs (Illumina-contigs) of more than 200 bp each (22.1 Mb in total). In this study, we obtained 0.398 Mb of sequence with longer reads (mean of 350 bp), which we assembled into 48,950 contigs (454-sequencing contigs) of more than 200 bp each (36.5 Mb in total). We were thus able to assemble many more reads, corresponding to 9.2% of the sequencing set, much greater than the 0.56% reported by Mizrachi *et al.* [39]. BLAST searches for sequence similarities between the two datasets showed that 86% of the contigs were common to both studies (42,550 454-sequencing contigs matched 16,278 Illumina contigs). However, each Illumina contig matched a mean of five 454-sequencing contigs, indicating that the alternative splicing forms detected with our approach were probably confounded in the short-read assembly. Thus, 454 technology seems to outperform Illumina SBS for cataloguing of the eucalyptus transcriptome, confirming the difficulty involved in assembling short reads into transcriptional units [54]. Finally, we found that 36,985 454-UE did not match any previously described eucalyptus ESTs, and that 43% of these UE (18% of the 454-sequencing contigs) displayed no match with any nucleic acid or protein sequence published for any other species. Our 454-sequencing data have therefore considerably enriched the *Eucalyptus* EST collection.

RNA-Seq is also an interesting approach to obtaining a comprehensive digital gene expression profile for specific tissues, cell types or developmental processes. In this study, the high degree of repeatability observed for the three replicates made it possible to test G, T and GxT effects with a statistical support. We were able to monitor the expression of 14,460 UE without the need for prior knowledge and to identify 1,445 UE displaying at least one significant effect. Some technical biases, such as non-linear amplification and a lack of sequencing depth may have resulted in a lack of precision in the prediction of gene expression by 454-sequencing. Short read-based sequencing approaches provide ample read coverage and generally give better predictions of gene expression [55, 56]. Thus, a combination of long and short reads may be seen as a reasonable strategy for the analysis of gene expression [54, 57, 58], in which long-read sequencing can be used to establish a comprehensive catalogue of transcriptional units accounting for splice variants, with the mapping of short reads onto this assembly providing greater sequencing depth, improving predictions of abundance.

5.2 Behaviour of variance components at the phenotypic and molecular levels

That phenotypic and molecular variations are accounted for principally by genotypic differences is a first major result of the present study that deserves discussion.

Both above-ground biomass and contig abundance seemed to be influenced principally by genotype: above-ground biomass was, on average, 49% higher for 18-50 than for 1-41, and most of the differentially expressed contigs (1,280 of 1,455) highlighted in this study presented a genotypic effect, 1,066 displaying a pure “G” effect. The number of contigs overexpressed in one or other of the genotypes was found to be similar (624 and 656 contigs overexpressed in 1-41 and 18-50, respectively). No particular genotypic signature in terms of functional categories or pathways was detected. The two genotypes differed strongly in phenotype (not only it terms of growth potential, but also in terms of leaf morphology, stomatal distribution and water use efficiency), but it remains unclear whether these differences in transcript abundance were responsible for trait variation, neutral, or simply involved in reproductive isolation between parental species. Indeed, differences in gene expression between species have been reported in the field of ecological genomics [59, 60] and interpreted as a mechanism of speciation. In our study, both genotypes were hybrid combinations between different species. This may have increased the number of differences between their transcriptomes. Further investigations of the role of gene expression in

ecological speciation may be required in the near future, particularly for eucalyptus, in which species complexes are common [61].

The difference in terms of variance accounted for by genotype-by-environment interaction (GEI) at the phenotypic and molecular level is a second important result of this study.

The NI treatment resulted in a significantly lower above-ground biomass, and this difference was greater for the least productive of the two genotypes, 1-41 (24%), than for 18-50 (12%). The G and T effects accounted for 56% and 13% of the above-ground biomass variation, respectively, whereas GxT effects accounted for only 0.2% of the variance. At the molecular level, we also found a higher proportion of genes displaying G effects (1,280 contigs, 8.8% of the contigs screened) but, surprisingly, we found fewer genes displaying T effects (155 contigs, 1.1% of the contigs screened) than GxT effects (274 contigs, 1.9% of the contigs screened).

While only 11 contigs were differentially expressed between the two treatments in genotype 1-41, 112 contigs showed differential expression in genotype 18-50 (4 contigs displayed differential expression for both genotypes). Moreover, when the whole *E. spp* sequencing set was screened, genotype 18-50 presented a larger number of specific contigs (10.5% of the *E.spp* sequencing set) than 1-41 (5.5% of the *E.spp* sequencing set). These results suggest that a larger set of genes is activated in genotype 18-50, leading to the triggering of specific responses to water deficit. This higher molecular sensitivity of genotype 18-50 may confer advantages ultimately resulting in a greater capacity to cope with water deficit during the dry season and, therefore, in stronger growth.

5.3 Genes displaying genotype-by-environment interaction (GEI) effects reflect differences in signal perception and response strategy

Contigs displaying GEI effects could be classified into two groups according to the function of the proteins encoded [62]: i) regulatory proteins responsible for drought signal transduction and response triggering, and ii) structural proteins involved in cell protection, damage repair and the maintenance of cell activity.

As far as regulatory genes were concerned, the genes involved in the biosynthesis of hormones, such as ethylene and auxins in particular (aldo/keto reductase, proteins of the ethylene-responsive family, 2-oxoglutarate-dependent dioxygenase) were mostly overexpressed in 1-41 and underexpressed in 18-50, in NI conditions. Surprisingly, only a few

genes related to ABA biosynthesis were highlighted in this study, despite this hormone being the first to be identified as involved in drought stress signalling [62, 63]. Interactions between hormones, including the antagonist interaction between ethylene and ABA in particular [64, 65], mostly influence stomatal closure and leaf growth [66].

Genes acting as second messengers in the transduction of hormonal signal to stomatal guard cells [67, 68] also displayed GEI effects: during the dry season, genes involved in calcium signalling were predominantly overexpressed in 1-41, whereas the response of 18-50 to drought preferentially involved receptor kinases. We also identified other signal transducers, such as light-induced proteins and heat-shock proteins, which may be related to other types of stress induced by water deficit, including osmotic stress due to pH variations and oxidative stress due to the accumulation of reactive oxygen species (ROS) [69, 70]. These pathways were mostly overexpressed in 1-41 and underexpressed in 18-50 during the dry season. Pathogenesis-related (PR) protein genes were overexpressed in 18-50, but displayed a less clear-cut pattern of expression in 1-41. PR proteins were initially reported to be induced by hormones or ROS in response to biotic stress [71, 72], but they have also been shown to be involved in other abiotic stresses [73]. Lee *et al*. [72] suggested that PR proteins may be used as storage proteins when growth is limited by environmental factors.

Some transcription factors have also been shown to respond strongly to water shortage in genotype 18-50. Two, in particular, encoded factors homologous to *AtMYB12* and *AtMYB85*, which have been shown to regulate secondary metabolism (flavonoid and lignin biosynthesis, respectively) in *Arabidopsis* [74]. These results suggest that water shortage induces different cellular stress cascades, perceived differently by the two genotypes.

Stress signal transducers interact to trigger the regulation of functional gene expression for the maintenance of three main functions: cell protection, damage repair and the maintenance of cell activity. Our results suggest that more genes related to cell protection were involved in the response to water shortage in genotype 1-41 than in genotype 18-50. Protection against osmotic stress seems to involve mostly carbohydrates, with 11 contigs displaying GEI effects, and, to a lesser extent, polyamines, which may modulate some ion channels [75]. Similarly, Warren *et al.* [76] reported that sugars appear to be the main metabolites involved in osmotic regulation in *Eucalyptus marginata*, having a greater effect than amino-acid metabolites, such as proline. Protection against oxidation involves redox-related proteins [77], and metal-handling proteins.

By contrast to the trends observed for genes relating to cell protection, more genes related to damage repair seem to be expressed during the dry season in genotype 18-50 than in genotype 1-41, particularly for genes related to cell organisation. The overexpression, during NI treatment, of genes related to primary metabolism, including carbohydrate, lipid and protein synthesis and degradation, suggests that resources are reallocated for the repair of cell structures or the formation of new structures during water shortage. The patterns of expression of secondary metabolism genes differed between the two genotypes. As an example, 1-41 displayed a higher number of genes related to terpenoids and flavonoids synthesis (that may protect against oxidative stress) overexpressed in the NI treatment compared to genotype 18-50. However, the contrasts between NIR and IR treatments were much higher for 18-50. Conversely, genes related to lignin biosynthesis (e.g. *CCoAOMT*) were overexpressed in the IR condition only for 18-50.

Gene related to photosynthesis were found to be under-expressed in 18-50 subjected to NI treatment, whereas no variation were found in 1-41. Other metabolic processes, such cell development and transport, controlled by genes encoding water or sugar channels, have been shown to be decreased by NI treatment, particularly in 18-50. These results confirm the trends observed by Alexandersson *et al.* [78] in *Arabidopsis*. These authors studied the expression of 18 genes encoding aquaporins and showed that most of these genes were downregulated in leaves subjected to a gradual water deficit. Interestingly, cell activity seemed to be more reduced at transcriptional level for 18-50, although this genotype grew more strongly. It is possible that, during water deficit, genotype 18-50 reduces its rate of photosynthesis and reallocates resources (as suggested by changes in primary metabolism) to preserve its cell structures and ability to resume growth when conditions become more favourable.

5.4 Evolutionary implication behind GEI

We found that 31 of the 274 contigs displaying a GxT effect were absent from the 1-41 sequence set (11.3%), whereas only two such contigs were absent from the 18-50 sequence set (0.7%). Unfortunately, of 31 contigs absent from 1-41 sequence set, 16 could not been assigned to an homolog gene in *Arabidopsis*. The others corresponded to genes related to cell organisation (ankyrins), ethylene synthesis, protein metabolism, PR proteins and receptor kinases. These genes may be considered non-essential for tree development, and are therefore unlikely to be subject to selection constraints. Landry *et al.* [79] showed an overrepresentation of non-essential genes (the deletion of which does not lead to death) among genes displaying GEI in *Saccharomyces cerevisiae*. They proposed two hypotheses to account for the activity

of these genes being compensated in cells: i) metabolic buffering: non-coded metabolites may be rerouted through the metabolic network, and ii) genetic buffering: paralogous genes may supply the missing function. We have shown in the results section (Additional file 5) that differences between genotypes may be accounted for by the preferential expression of different members or splicing forms of genes from the same family. This observation may confirm the hypothesis of genetic buffering.

Scale plasticity was observed for 146 of the 274 differentially expressed contigs: genotype ranks were conserved between treatments but one genotype reacted more or less strongly to the environmental variation. Conversely, 90 contigs showed a change in ranking between genotypes (rank plasticity). Landry *et al.* (2006) hypothesised that these two types of GEI would have different effects on the evolution of plastic traits: in the case of scale plasticity, whatever the environment, selection would result in the same favoured genotype, whereas in the case of rank plasticity, different genotypes would be selected in different environments. In this study, a similar proportion of differentially expressed genes presented scale plasticity (62%) and rank plasticity (38%), a potentially good basis for the selection of either specialist or generalist genotypes.

6 Conclusions

In this study, we have shown that next-generation sequencing is a powerful tool for transcriptome screening: with 398 Mb of sequence, we were able to assemble 69,584 contigs, with remaining 80,245 singletons, and to determine the relative abundance of 14,460 contigs each comprising more than 10 reads. Large differences between genotypes, in terms of both the associated phenotype and the transcriptome, were observable at the juvenile stage. Differences in gene expression between the two genotypes appear to affect the whole transcriptome, rather than specific pathways. The genotype-specific response to water shortage (i.e. GxT effect) was more pronounced than the response common to both genotypes (i.e. T effect). The genes displaying genetically controlled plasticity were found to belong to a number of different pathways essentially related to signal transduction and primary metabolism. The more productive genotype, 18-50, seemed to express a larger set of genes, leading to the triggering of specific molecular responses. Moreover, GxT interaction results principally from a lack of molecular response in genotype 1-41, together with a strong response of genotype 18-50 (table 4). The ability to activate more strongly its transcriptome

appears a key element in the maintenance of biomass in conditions of water deficit for the genotype 18-50.

Finally, although this study provides clues to the way in which different genotypes activate their transcriptomes when subjected to water deficit, more information is required for a broader view of the molecular mechanisms involved during the dry season. First, we need to describe the distribution of genetic variation for such reaction norms in breeding populations [80]. Second, epigenetics or post-transcriptional regulation mechanisms are well known to interfere with abiotic stress responses [81, 82]. Indeed, evaluation of the adaptive nature of specific gene products remains a key issue in molecular breeding [83]. A combination of several approaches, such as reverse and forward approaches [84] with the integration of a maximum of phenotypic traits, is required to identify relevant candidate genes for breeding.

Table 4 – Summary of phenotypic and molecular plasticity evidenced for the two studied genotypes (1-41 and 18-50), between irrigated (IR) and non irrigated (NI) treatments.

		1-41 **IR → NI**	**18-50** **IR → NI**
Above-ground biomass		-	0
Transcript abundance	Hormones	+	--
	Secondary messengers	+ (light signalling)	-- (light signalling) ++ (receptor kinases)
	Other stress related genes	+	-- (heat-shock proteins) ++ (PR proteins)
	Transcription factors	0	++
	Cellular protection	+	--
	Damages repair	+ (carbohydrates, secondary metabolism)	++ (cell organisation, lipids, proteins) -- (secondary metabolism)
	Cellular activity maintenance	0	-- (photosynthesis, transport)

Abbreviations: +/- means that in a given category more than 50% of the genes are over/underexpressed in NI.++/-- means that in a given category more than 50% of the genes are over/underexpressed in NI with a log2 tranformed fold change <-0.8 or >0.8.

Authors' contributions

EV: participated in the field work, performed the molecular work, carried out the statistical analysis and drafted the manuscript. CK and CN performed the bioinformatic work (assembly and annotation). EN and MK provided *E. grandis* sequencing sets and participated in the statistical analysis. CP participated in co-ordination of the molecular, bioinformatic and statistical work and in drafting the manuscript. JMG participated in the design of the study, co-ordination of the field, molecular, bioinformatic and statistical work and drafting the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgements

We thank Andréas Ndeko, André Mabiala, Joël Polidori and Dr. Aubain R. Saya (CRDPI, Rep. of Congo) for setting up the experimental trial. This article is part of the PhD thesis of Emilie Villar, supervised by Jean-Marc Gion and Christophe Plomion. EV was supported by CIRAD. This work was supported by grants from CIRAD, ATP project "Plasticité phénotypique des pérennes sous contrainte hydrique au champ", and the *Conseil Régional d'Aquitaine* ("ABIOGEN" project no.° Presage 32973). Some of the experiments (RT-qPCR) presented here were performed at the Genome Transcriptome facility of Bordeaux (grants from the *Conseil Régional d'Aquitain* nos. 20030304002FA and 20040305003FA and from the European Union, FEDER no. 2003227).

SUPPORTING INFORMATION

1 Methods

The transcript abundance of 36 genes was estimated by RT-qPCR for three types of sample: i) cDNA samples (obtained with the Smart cDNA Library Construction kit) used for sequencing, ii) RNAs used for library construction and sequencing, and iii) RNAs extracted from independent samples. The extracted RNAs obtained in (ii) and (iii) were reverse-transcribed with the ImpromII Reverse Transcription System (Promega, Madison, WI, USA) according to the manufacturer's instructions. Primer pairs were designed from the 454-reads with Primer3 software (Rozen and Skaletsky, 2000) (see supporting file 1). Real-time PCR was performed on 384-well plates, and was monitored with a Light Cycler 480 (Roche Applied Science, Indianapolis, IN, USA) at the Genome and Transcriptome Facility of Bordeaux (France, *https://www4.bordeaux-aquitaine.inra.fr/pgtb),* France. Each reaction mixture contained 3 µl of cDNA preparation (diluted 1/50 for Improm reverse-transcribed cDNA and 1/1000 for SMART reverse-transcribed cDNA), 5 µl of 2 x ABsolute QPCR SYBR Green Mix (ABgene, Epsom, UK) and 70 nM gene-specific primer in a total volume of 10 µl. The cDNA samples and master mix were pipetted with a Hamilton workstation (Hamilton Robotics, Reno, NV, USA). Each reaction was performed in duplicate. The fluorescent signal was evaluated and starting concentrations were calculated with LinRegPCR (Ramakers *et al.* 2003; Ruijter *et al.* 2009). The relative abundance of transcripts was normalised against the mRNA levels for four constitutively expressed housekeeping genes: elongation factor 1-α (EF1-α), elongation factor 2 (eEF2), glyceraldehyde 3P dehydrogenase (G3PDH) and cyclophilin (Cyp), by calculating a normalisation factor with GeNorm (Pattyn *et al.*, 2003). Log2-transformed ratios between treatments were generated for each genotype, to facilitate comparison of the results with those of digital expression profiling.

2 Results

We selected a subset of 36 contigs for validation of their expression levels by RT-qPCR. Sixteen displayed stable levels of expression ($0.5<FC<2$), and 21 showed a contrasted expression pattern ($FC<0.5$ or $FC>2$) in the digital expression analysis, between genotypes or treatments, or displaying GxT effects. Four stably expressed genes were chosen with GeNorm for data normalisation: these genes encoded elongation factor 1-alpha (*EF1a*), eukaryotic

translation elongation factor 2 (*eEF2*), cyclophilin (*CYP*), and glycerol-3-phosphate dehydrogenase (*G3PDH*). RT-qPCR was carried out on three sets of RNA/cDNA samples: i) 8 cDNA samples (used for 454-sequencing) obtained by SMART amplification of eight RNA samples corresponding to the four conditions with two replicates (set #1), ii) eight new cDNA samples obtained by linear reverse transcription (RT) of the same RNA samples (set #2), iii) eight new cDNA samples obtained by RT of eight new RNA samples (set #3).

Gene expression, as analysed by RT-qPCR, was strongly correlated in sets #2 and #3 (significant Spearman correlation of 93.4%), suggesting a high degree of homogeneity in the pattern of expression between biological replicates. Correlations between gene expression in these two sets and in set #1 were weaker, with a mean correlation of 68.3%. The non-linear SMART amplification may account for these differences in expression patterns and suggests that caution is required when using this technique for expression validation. We therefore chose not to use this method in subsequent analyses. Spearman's correlations between digital expression profiles (454-results) and the two sets of RT-qPCR results were significant, but relatively low: 35.3% (set #2), 28.7% (set #3). The weakness of these correlations may be accounted for by the successive biases of the method: non-linear amplification by SMART and low sequencing depth, generating only a limited number of reads for digital expression analysis. Nevertheless, DEGseq analysis of the 11 contigs with significant fold-changes between IR and NI treatments for the two genotypes in 454-sequencing data, confirmed the expression patterns of seven genes in set #1, ten genes in set #2, and seven genes in set #3 (supporting file 2). The reaction norms of molecular variations cannot be fully described by 454 quantification (Spearman's correlations between the results obtained with the two methods vary from 28.7% to 35.3%), the pattern of molecular plasticity of significant contigs was similar for the different types of expression quantification.

Supporting file 1 – Primers used for validation by RT-qPCR

Gene Code	Gene Name	Contig Name	Primer Forward (5'-3')	Primer Reverse (5'-3')
SHINE1	DNA binding / transcription factor	F0F0A3T01A00IG.l.eu.2	TCAAGAAATGCAGCAAGACG	GACCCAGTTGGAGTCTGAGC
LP1	nonspecific lipid transfer protein 1	F0F0A3T01A00NX.l.eu.2	AGGGGATCGACTTCAACCTC	GAGCTCGTGGACCTCACTTC
SIP2	hydrolase, hydrolyzing O-glycosyl compounds	F0F0A3T01A032B.l.eu.2	CGATTTGGCGCTTACTCTTC	ACCGGCCATCTGTACATCTC
LIM1	transcription factor	F0F0A3T01A03RQ.l.eu.2	AGAGGGAAACAATGGCTCCT	GCTTACAACCAACCCGAGAA
RD26	NAC domain containing protein 2; transcription factor	F0F0A3T01A0DJA.l.eu.2	CAACACACCTCCGCAATATG	TCCCAACCTTCCAACTTCAC
DCAM	adenosylmethionine decarboxylase family protein	F0F0A3T01A0ESR.l.eu.2	ATCCAGAGGGTGTTGGTCTG	CTCCAGCACTGTACCCCTTC
CHIAB	basic chitinase	F0F0A3T01A0IP5.l.eu.2; F0F0A3T01A3DVE.l.eu.2	CAAGGGCTTCTACACTTACG	AGCAGTATCCCCACGCATAG
ELIP1	(EARLY LIGHT-INDUCABLE PROTEIN); chlorophyll binding	F0F0A3T01A0S92.l.eu.2	TCCCTCTTCATCGCCTACAC	GTCTGCCATTGATCCTCTCC
PHS2	ALPHA-GLUCAN PHOSPHORYLASE 2	F0F0A3T01A0SF1.l.eu.2	TTCTAGCTTCCCTCCCTTCC	GTTAGGCGGAGTGCATTTGT
CCoAOMT2	caffeoyl-CoA 3-O-methyltransferase, putative	F0F0A3T01A0Z5S.l.eu.2; F0F0A3T01BLVGJ.l.eu.2	GCAAGATGAGAAGAACCATG	AGCACGAAGTCCCGGTAGTA
UCR	ubiquinol-cytochrome C reductase complex, putative	F0F0A3T01A1SQD.l.eu.2	GCTGTCTCCTGACGTTTGTG	CCGGCTATTTCGACTCCTTC
DELTA-TIP	delta tonoplast integral protein	F0F0A3T01A2Q8I.l.eu.2	CCGTGGTCAGTGGAGACTTC	CATTGACAAGAGGTGCATGG
CesA1	CELLULOSE SYNTHASE 1	F0F0A3T01A31VD.l.eu.2	TAAAGTGCTTGCTGGCATTG	TGAGAAGCGATGTCCATTTG
PIP1B	plasma membrane intrinsic protein 1;2	F0F0A3T01A344C.l.eu.2	TTCTGCAACATGGCTACACC	ATTCCTCTTGGCATCAGTGG
CCoAOMT1	caffeoyl-CoA 3-O-methyltransferase, putative	F0F0A3T01A4G7A.l.eu.2	GTATGTGAGGTACTACAGGG	CTCTCTTTTTCGTGGGTTCT
LHY1-2	LATE ELONGATED HYPOCOTYL	F0F0A3T01A4GAO.l.eu.2	GAAAGAGGATAATGGGCCAAG	CCTCGGGAAATTATGAGCAG
RBCL	large subunit of RUBISCO	F0F0A3T01A5ANA.l.eu.2	GCTAAGAACTACGGTAGAGC	TTCACCTGTTTCAGCCTGTG
ERD15	EARLY RESPONSIVE TO DEHYDRATION 15	F0F0A3T01A5C8L.l.eu.2	ATATTTCGCCGTCGTTCAAG	ATAGAGGGGCATCTGGGTTC
PGR5	PROTON GRADIENT REGULATION 5	F0F0A3T01A5OGX.l.eu.2 ; F0F0A3T01A1FPP.l.eu.2	TGGCTTGTTGCTAGGGTTTC	ATTGATCGCGAGCACTTTTC
GDR1	GLUTATHIONE-DISULFIDE REDUCTASE	F0F0A3T01A5Q34.l.eu.2	TAAGTGTGGTGCAACGAAGG	AACCCGTCTTGTGACTGACC
2-ODD	2-oxoglutarate-dependent dioxygenase	F0F0A3T01A8608.l.eu.2	CTCCAGATCATGTCCAACGA	CAAACAGGCTTTCACAATCG
HYP1	HYPOTHETICAL PROTEIN 1	F0F0A3T01A8IGK.l.eu.2	CCGGTGATGTGAGAGAAACC	AGCGGTATACGCGACTTGAG
RPP1A	60S acidic ribosomal protein P1	F0F0A3T01AFPEY.l.eu.2	CACCAATGTTGGTTCTGGTG	AGACCAAAGCCCATGTCATC
GHP3	glycosyl hydrolase family 3 protein	F0F0A3T01AJF18.l.eu.2	GGGTTGCAGACGTTCTGTTT	ATGAGGATCCCCAACATTCA
DXR	1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE	F0F0A3T01AJOYE.l.eu.2	TGAGATCATTCCAGGGGAAC	GCTATGTCCTTGCCTGCTTC
HEN4	HUA ENHANCER 4; nucleic acid binding	F0F0A3T01AYA64.l.eu.2	GAGAGGAAGGTGGATGCTTG	TGAGCTGCATGAGTTTGCTC
HVA22	abscisic acid-responsive HVA22 family protein	F0F0A3T01B0RI8.l.eu.2	TCGAGAGCAAATCACCAATG	CTCTCACCGTCTGCACTCTG
LHY1-1	LATE ELONGATED HYPOCOTYL	F0F0A3T01B9JR1.l.eu.2	CGTTGAAAGAGGGAGTTCAGG	GAGCAGTTTGCATCTCATCTG
POLG	polygalacturonase, putative	F0F0A3T01BFXHG.l.eu.2	CGAAACCAAGATGTGTGTGG	CGATCAAGACCTCCTGCTTC
CER6	CUTICULAR 1; catalytic	F0F0A3T01EOHZ5.l.eu.2	TCCATATTCCCGAAGTGGTC	CGCAATCCACCTATCAAACC

Supporting file 2 – Comparison of fold-change (FC) ratio between RT-qPCR and digital analysis for 30 genes. Expression assessed by determining the amount of cDNA obtained with the SMART construction kit (Set #1) and in two biological replicates (Set #2 and Set #3) by linear reverse transcription (RT) was compared with the results for 454-sequencing. The mean log2-transformed FC values for the two biological replicates are indicated

				Log2 FC : [IR/NI]								Fold Change Similarity		
				454		Set 1 (SMART)		Set 2 (RT)		Set 3 (RT)				
Gene Code	Gene Name	Contig Name	nb reads	1-41	18-50	1-41	18-50	1-41	18-50	1-41	18-50	Set #1/ 454	Set #2/ 454	Set #3/ 454
LP1	nonspecific lipid transfer protein 1	F0F0A3T01A00NX.l.eu.2	2751	1.42*	-1.6*	1.36	-0.6	3.12	n.a. (-)[c]	n.a. (+)	-2.58	Y	Y	Y
ELIP1	(EARLY LIGHT-INDUCABLE PROTEIN); chlorophyll binding	F0F0A3T01A0S92.l.eu.2	317	-0.07	3.32*	-0.77	-0.34	n.a. (+)[b]	0.02	4.32	0.51	N	Y	Y
SIP2	hydrolase, hydrolyzing O-glycosyl compounds	F0F0A3T01A032B.l.eu.2	463	0	1.98*	-1.13	-0.81	2.89	2.35	4.02	-0.64	N	Y	N
LHY1-2	LATE ELONGATED HYPOCOTYL	F0F0A3T01A4GAO.l.eu.2	275	0.62	4*	2.19	3.13	4.46	0.62	4.17	2.15	Y	Y	Y
DCAM	adenosylmethionine decarboxylase family protein	F0F0A3T01A0ESR.l.eu.2	392	0.34	1.88*	0.85	1.59	3.6	0.26	3.22	1.3	Y	Y	Y
DELTA-TIP	delta tonoplast integral protein	F0F0A3T01A2Q8I.l.eu.2	637	0.02	1.24*	-0.02	2.21	4.12	2.28	2.39	2.26	Y	Y	Y
LHY1-1	LATE ELONGATED HYPOCOTYL	F0F0A3T01B9JR1.l.eu.2	75	0.77	4.93*	1.89	0.09	0.38	0.52	0.49	0.21	Y	Y	Y
PIP1B	plasma membrane intrinsic protein 1;2	F0F0A3T01A344C.l.eu.2	177	0.06	4.43*	1.19	1.1	2.13	-2.05	n.a. (+)	n.a. (-)	Y	N	N
PHS2	ALPHA-GLUCAN PHOSPHORYLASE 2	F0F0A3T01A0SF1.l.eu.2	43	0.87	-2.53*	-0.35	-0.4	0.35	-0.68	-0.22	-0.61	N	Y	N
PGR5	PROTON GRADIENT REGULATION 5	F0F0A3T01A5OGX.l.eu.2	209	0.49	2.14*	1.25	3.45	2.26	2.59	2.14	3.7	Y	Y	Y
CHIAB	basic chitinase	F0F0A3T01A0IP5.l.eu.2	290	-1.11*	-1.07*	-1.41	-1.12	-7.5	-5.71	-6.97	-10.93	Y	Y	Y
UCR	ubiquinol-cytochrome C reductase complex, putative	F0F0A3T01A1SQD.l.eu.2	303	0.45	1.04	0.23	0.21	n.a. (+)	0.83	2.88	-0.08	Y	Y	N
HEN4	HUA ENHANCER 4; nucleic acid binding	F0F0A3T01AYA64.l.eu.2	22	-0.13	-3.48	0.65	-0.82	2.24	-0.34	8.99	-0.17	Y	N	N
POLG	polygalacturonase, putative	F0F0A3T01BFXHG.l.eu.2	21	n.a.[a]	-2.11	-0.51	-1.8	n.a.	-2.02	n.a.	-1.89	Y	Y	Y
SHINE1	DNA binding / transcription factor	F0F0A3T01A00IG.l.eu.2	205	-0.07	-0.65	-0.14	0.55	n.a.	-3.69	n.a.	-2.12	N	Y	Y
2-ODD	2-oxoglutarate-dependent dioxygenase	F0F0A3T01A8608.l.eu.2	46	1.99	-1.48	0.75	0.68	2.99	-2.05	3.34	-2.88	N	Y	Y
HYP1	HYPOTHETICAL PROTEIN 1	F0F0A3T01A8IGK.l.eu.2	23	-0.54	-2.14	0.78	0.59	5.26	0.33	4.7	2.11	N	N	N
HVA22	abscisic acid-responsive HVA22 family protein	F0F0A3T01B0RI8.l.eu.2	18	3.19	3.78	-0.71	-0.91	4.35	-6.26	4.54	-1.97	N	N	N
DXR	1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE	F0F0A3T01AJOYE.l.eu.2	68	1.31	1.26	0.94	4.21	2.45	1.14	2.95	3.62	Y	Y	Y
GHP3	glycosyl hydrolase family 3 protein	F0F0A3T01AJF18.l.eu.2	55	0.46	-1.15	1.45	-1.48	1.65	-0.59	0.93	-1.42	Y	Y	Y
ERD15	EARLY RESPONSIVE TO DEHYDRATION 15	F0F0A3T01A5C8L.l.eu.2	162	-0.13	0.59	-0.23	-0.1	2.84	0.19	0.66	0.36	N	N	N
RPP1A	60S acidic ribosomal protein P1	F0F0A3T01AFPEY.l.eu.2	15	-3.3	-2.03	-0.59	0.08	1.73	1.94	-0.55	0.26	N	N	N
LIM1	transcription factor	F0F0A3T01A03RQ.l.eu.2	156	-0.99	-0.4	-4.92	1.9	n.a.	-2.24	n.a.	-3.06	N	Y	Y
RD26	NAC domain containing protein 2; transcription factor	F0F0A3T01A0DJA.l.eu.2	51	-0.13	0.88	0.94	3.98	3.01	-1.01	1.79	n.a.	N	N	Y
CCoAOMT2	caffeoyl-CoA 3-O-methyltransferase, putative	F0F0A3T01A0Z5S.l.eu.2	114	-0.57	0.68	-0.12	0.52	-0.04	0.35	-0.04	-0.27	Y	Y	N
GDR1	GLUTATHIONE-DISULFIDE REDUCTASE	F0F0A3T01A5Q34.l.eu.2	114	-0.58	0.59	-0.3	1.14	4.31	1.85	5.08	-0.48	Y	N	N
CesA1	CELLULOSE SYNTHASE 1	F0F0A3T01A31VD.l.eu.2	169	-0.94	0.21	-0.41	-1.5	2.65	-0.43	n.a. (+)	-3.6	N	N	N
CCoAOMT1	caffeoyl-CoA 3-O-methyltransferase, putative	F0F0A3T01A4G7A.l.eu.2	55	0.41	-0.44	-0.07	0.19	0.14	0.07	0.27	0.09	Y	Y	Y
RBCL	large subunit of RUBISCO	F0F0A3T01A5ANA.l.eu.2	233	-0.1	0.06	0.89	-0.21	0.97	-0.06	1.17	-0.26	N	N	N
CER6	CUTICULAR 1; catalytic	F0F0A3T01EOHZ5.l.eu.2	16	-0.39	0.56	-0.63	-3.03	2.11	-0.32	4.23	-3.13	N	N	N

* Significant fold-change according to DEGseq analysis, at a q-value threshold of 0.05.

a: fold-change could not be calculated because no reads were found for the genotype considered in either the IR or the NI sequencing set.

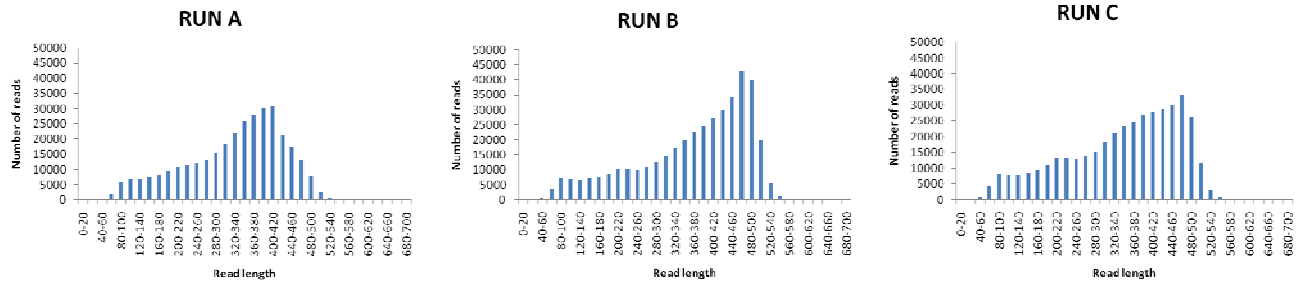
b: fold-change could not be calculated because no expression was reported for the NI treatment, for the genotype considered.

c: fold-change could not be calculated because no expression was reported for the IR treatment, for the genotype considered.

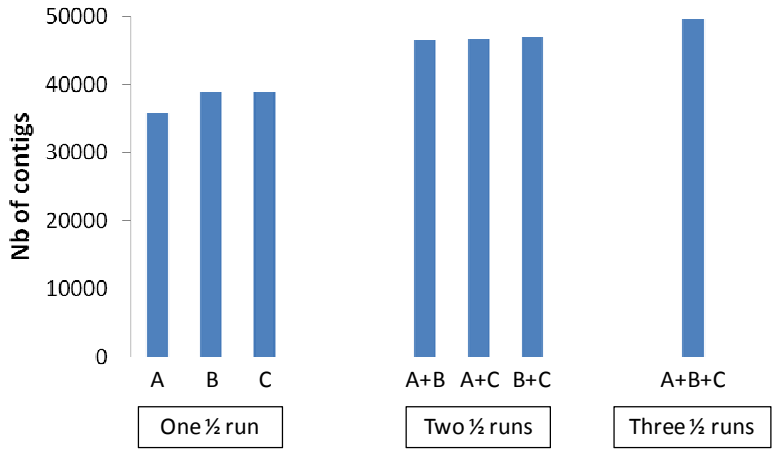
ADDITIONAL FILES

Additional file 1 – Results of the ANOVA for above-ground biomass

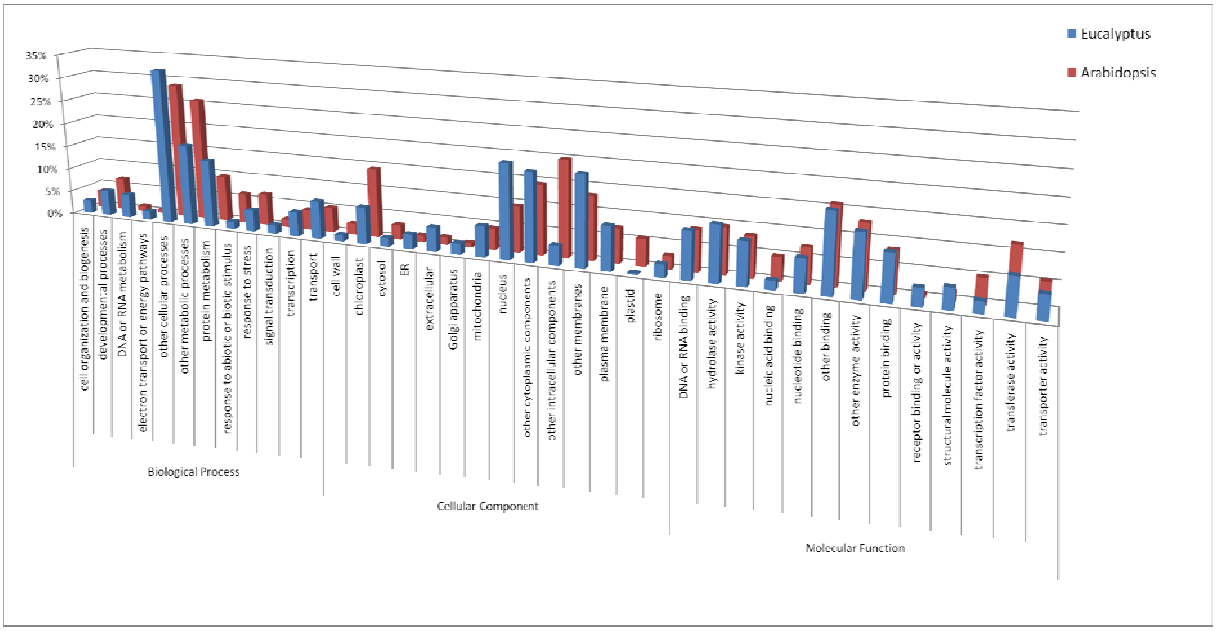
Source	DDL	Sum of squares	Mean of squares	F	Pr > F
Genotype	1	183.059	183.059	75.868	< 0.0001
Treatment	1	43.192	43.192	17.901	0.0001
Genotype x Treatment	1	0.934	0.934	0.387	0.537



Additional file 2 – Distribution of read length for the three half-runs



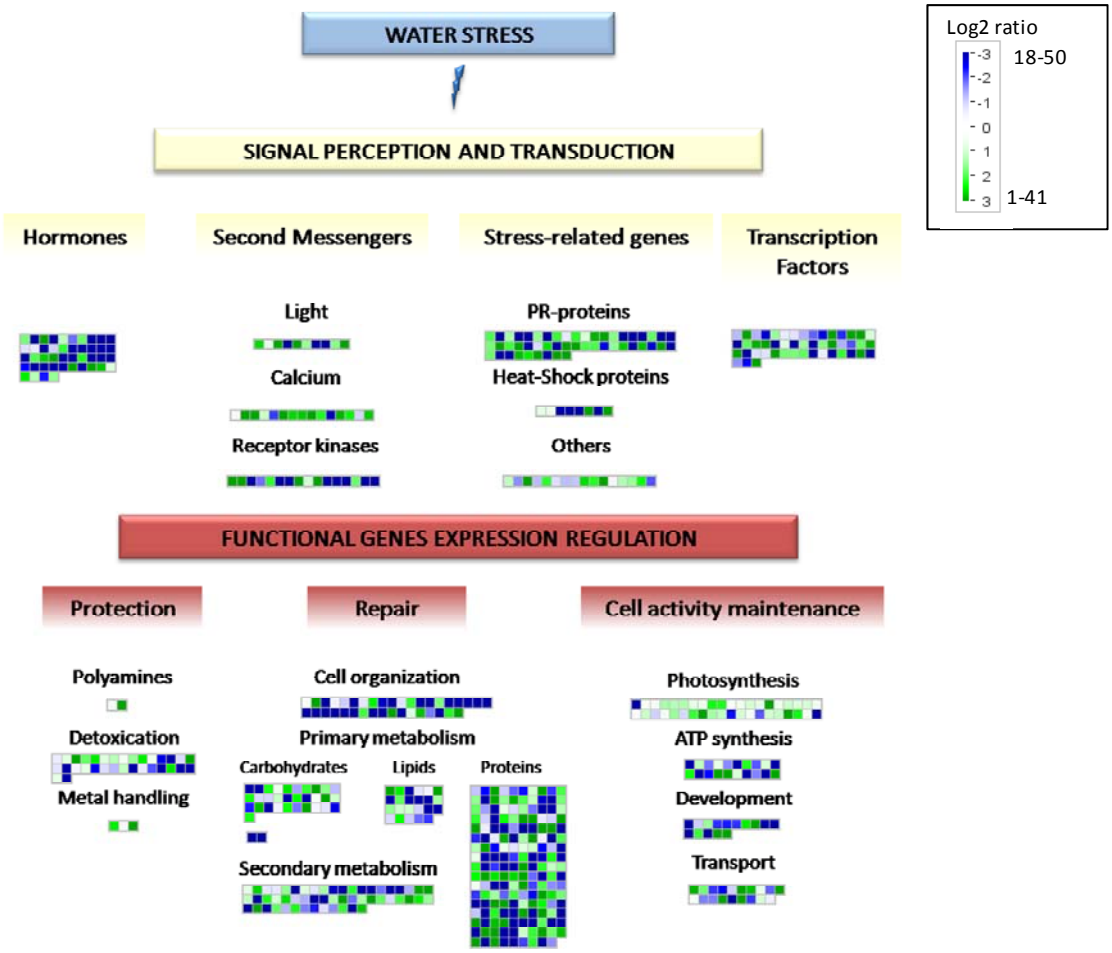
Additional file 3 – Increasing coverage with successive runs
Number of contigs represented in each half-run or combination of several half-runs. Performing a second half-run increased contig coverage by an average of 18%, and a third half run increased coverage by an average of 6%.



Additional file 4 – Comparison of the distribution of Gene Ontology (GO) categories between Eucalyptus spp unigene elements (UE) and Arabidopsis annotated unigenes
Proportion of each GO category (Biological Process, Cellular Component and Molecular Function) found in the *E. spp* sequencing set and in the annotated *Arabidopsis* genome.

Additional file 5 – Differentially expressed contigs displaying significant G, T and/or GxT effects.

Supplied in a separate Excel file



Additional file 6 – Distribution of "G" contigs between functional pathways

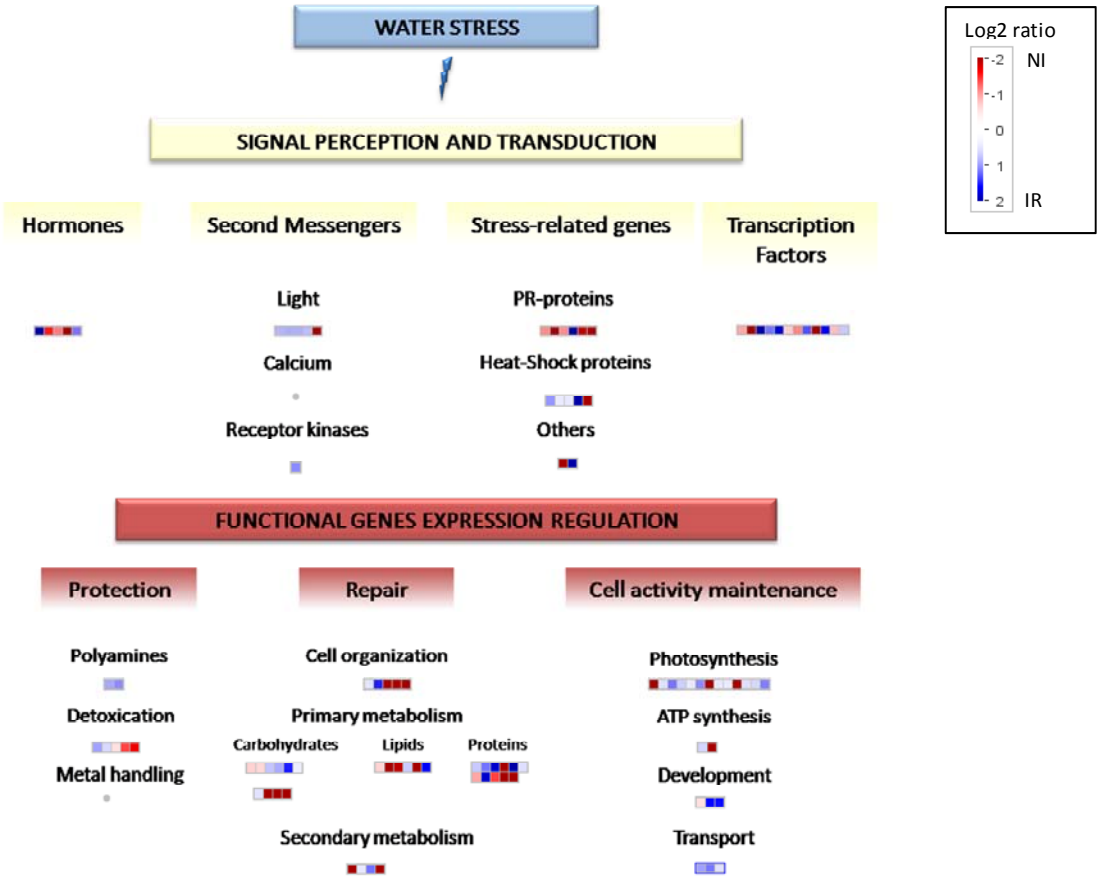
Each square represents the log2-transformed ratio of abundance between genotypes 1-41 and 18-50 for one contig. Contigs in green were overexpressed in genotype 1-41 and contigs in blue were overexpressed in genotype 18-50.

Additional file 7 – Significant categories for the Wilcoxon rank sum test, according to Mapman analysis for the pairwise comparison of differentially expressed contigs displaying genotype (G), treatment (T) and genotype x treatment (GxT) effects

	"G" contigs: (1-41 vs 18-50)		"T"contigs: (IR vs NI)		"GxT"contigs: 1-41 (IR vs NI)		"GxT"contigs: 18-50 (IR vs NI)		"GxT"contigs: IR (1-41 vs 18-50)		"GxT"contigs: NI (1-41 vs 18-50)	
Category	# contigs	mean FC	# contigs	mean FC	# contigs	mean FC	# contigs	mean FC	# contigs	mean FC	# contigs	mean FC
hormone metabolism	44	-2**	5	-0.77	7	-0.14	9	0.25	9	-1.42*	10	-0.51
hormone metabolism : auxin	8	-1.22	1	-1.53	2	-0.21	2	2.67*	2	-1.44	2	1.45
hormone metabolism : ethylene signal transduction	21	-3.22**	1	-6.36*	0	NA	3	-2.37	2	-5.24**	3	-5.39**
signalling : calcium	14	1.89**	0	NA	6	-0.48	4	0.57	6	0.87	6	2.56*
signalling : light	10	0.15	5	0.11	3	-0.47	2	2.88*	3	0.27	3	3.52*
signalling : receptor kinases	16	-1.07	1	1.09	1	1.25	3	-2.74*	1	-0.89	3	-4*
stress	73	-0.05	13	-0.68	17	-0.27	19	0.22	20	-0.81**	21	-0.21
abiotic stress : heat	10	-0.85	6	-0.26	9	-0.7*	10	1.48*	10	-1.34**	10	0.68
biotic stress : PR-proteins	29	0.51**	3	-1.63	3	-0.2	3	-2.76*	4	-0.4	5	-0.88
regulation of transcription : MYB-related transcription factor family	3	0.52	2	1.5*	0	NA	0	NA	0	NA	0	NA
regulation of transcription : unclassified	6	-2.15	2	-0.04	1	0.16	1	7.42*	1	1.25	1	8.51*
nitrilases	7	2.12*	1	0.97	0	NA	0	NA	0	NA	0	NA
photosystem	38	0.94*	10	0.73	13	0.1	13	0.99*	13	0.38	13	1.28
cell	37	-2.31**	5	-1.62	3	0.03	5	-2.7**	5	-1.37	6	-2.91**
major CHO metabolism : starch degradation	2	-3.67	3	-2.76**	1	0.56	2	-2.16*	1	-0.79	2	-2.67*
lipid metabolism	14	-0.59	1	-5.07	5	0.25	6	-1.47	5	1.19*	6	-0.49
protein degradation : serine protease	9	1.75**	0	NA	0	NA	1	-2.61	1	0.05	1	-4.43
protein degradation : ubiquitin E3	23	-2.04*	3	-1.99	5	-0.39	6	-0.4	5	0.02	6	-0.52
protein synthesis : ribosomal protein	47	-0.47	3	1.26*	2	1.93**	3	-1.6	3	-0.53	3	-3.72*
secondary metabolism : phenylpropanoids	24	0.11	2	0.84	4	-0.24	3	2.97**	4	1.7**	4	4.16**
development	14	-0.62	2	1.58*	6	-0.63	6	0.64	6	0.42	7	1.32

** Categories differentially expressed at an error rate threshold of 0.05

* Categories differentially expressed at an error rate threshold of 0.1



Additional file 8 – Distribution of "T" contigs between several functional pathways

Each square represents the log2-transformed ratio of abundance between irrigated (IR) and non-irrigated (NI) treatments for one contig. Contigs in blue were overexpressed for the IR treatment and contigs in red were overexpressed for the NI treatment.

Additional file 9 – "G" contigs related to secondary metabolism

Several contigs had the same hits on TIGR (Arabidopsis), and are considered to be different members of the same gene family on the basis of sequence alignment. Contigs associated with AT1G6180, AT3G26040 and AT5G48930 are all expressed more strongly in 1-41 than in 18-50, whereas contigs associated with other TIGR Arabidopsis hits had different expression patterns, depending on the contig considered.

Contig name	Hit on TIGR Arabidopsis	Gene Description	Log 2 transformed ratio between 1-41 and 18-50	Number of reads	
				1-41	18-50
F0F0A3T01ASD8T.1.eu.2	AT1G61680	terpene synthase/cyclase family protein	5.56	22	0
F0F0A3T01BJEVF.1.eu.2	AT1G61680	terpene synthase/cyclase family protein	3.80	13	1
F0F0A3T01CAOPZ.1.eu.2	AT3G26040	transferase family protein	5.27	18	0
F0F0A3T01A7GQ1.1.eu.2	AT3G26040	transferase family protein	1.42	112	45
F0F0A3T01AOJ15.1.eu.2	AT5G48930	transferase family protein (HCT)	5.01	15	0
F0F0A3T01A4QGQ.1.eu.2	AT5G48930	transferase family protein (HCT)	2.10	20	5
F0F0A3T01B6EJ3.1.eu.2	AT1G67980	caffeoyl-CoA 3-O-methyltransferase (CCoAOMT)	-5.94	0	33
F0F0A3T01A058P.1.eu.2	AT1G67980	caffeoyl-CoA 3-O-methyltransferase (CCoAOMT)	2.49	178	34
F0F0A3T01C80R9.1.eu.2	AT1G67980	caffeoyl-CoA 3-O-methyltransferase (CCoAOMT)	3.56	11	1
F0F0A3T01A0B6N.1.eu.2	AT1G67980	caffeoyl-CoA 3-O-methyltransferase (CCoAOMT)	1.56	107	39
F0F0A3T01A7NEB.1.eu.2	AT1G72680	cinnamyl-alcohol dehydrogenase, putative	1.91	28	8
F0F0A3T01AFFBH.1.eu.2	AT1G72680	cinnamyl-alcohol dehydrogenase, putative	-0.91	55	111
F0F0A3T01ADL4R.1.eu.2	AT3G25830	myrcene/(E)-beta-ocimene synthase (ATTPS-CIN)	-1.79	7	26
F0F0A3T01A3WRD.1.eu.2	AT3G25830	myrcene/(E)-beta-ocimene synthase (ATTPS-CIN)	2.19	81	19
F0F0A3T01AO3VN.1.eu.2	AT3G25830	myrcene/(E)-beta-ocimene synthase (ATTPS-CIN)	-1.36	28	77
F0F0A3T01A015O.1.eu.2	AT4G39230	isoflavone reductase, putative	-0.86	77	150
F0F0A3T01A0U1S.1.eu.2	AT4G39230	isoflavone reductase, putative	-1.62	17	56
F0F0A3T01A52VR.1.eu.2	AT4G39230	isoflavone reductase, putative	2.15	29	7
F0F0A3T01A2NPD.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	-6.45	0	47
F0F0A3T01A0VIZ.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	0.86	228	135
F0F0A3T01A0KXC.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	1.56	91	33
F0F0A3T01A1UWM.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	1.41	124	50
F0F0A3T01A9D1X.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	-4.71	0	14
F0F0A3T01AHD3X.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	-3.48	2	24
F0F0A3T01A81XL.1.eu.2	AT5G23960	terpene synthase/cyclase family protein	-5.07	0	18
F0F0A3T01A5HUB.1.eu.2	AT5G41040	transferase family protein	1.39	61	25
F0F0A3T01A1PTC.1.eu.2	AT5G41040	transferase family protein	-1.08	50	113
F0F0A3T01A02PQ.1.eu.2	AT5G41040	transferase family protein	0.92	661	374
F0F0A3T01A3WVC.1.eu.2	AT5G41040	transferase family protein	-7.20	0	79
F0F0A3T01A4UF8.1.eu.2	AT5G41040	transferase family protein	-7.93	0	131
F0F0A3T01AGUM8.1.eu.2	AT5G54160	O-methyltransferase 1 (ATOMT1)	-1.72	17	60
F0F0A3T01A0KSS.1.eu.2	AT5G54160	O-methyltransferase 1 (ATOMT1)	1.89	512	148
F0F0A3T01A1ZOT.1.eu.2	AT5G54160	O-methyltransferase 1 (ATOMT1)	-1.27	17	44
F0F0A3T01AGIR0.1.eu.2	AT5G54160	O-methyltransferase 1 (ATOMT1)	2.29	32	7
F0F0A3T01A02EC.1.eu.2	AT5G54160	O-methyltransferase 1 (ATOMT1)	-3.71	2	28

REFERENCES

1. Krutovskii KV, Neale DB: **Forest genomics for conserving adaptive genetic diversity**. *Forest Genetic Resources* 2001(29):6-8.

2. Hamrick JL: **Response of forest trees to global environmental changes**. *Forest Ecology and Management* 2004, **197**(1-3):323-335.

3. Jump AS, Penuelas J: Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters* 2005, 8(9):1010-1020.

4. **Eucalyptus universalis. Global cultivated forests map 2008.** [http://www.git-forestry.com/]

5. McGowen MH, Wiltshire RJE, Potts BM, Vaillancourt RE: **The origin of Eucalyptus vernicosa, a unique shrub eucalypt**. *Biological Journal of the Linnean Society* 2001, **74**(3):397-405.

6. Jones RC, Steane DA, Potts BM, Vaillancourt RE: **Microsatellite and morphological analysis of Eucalyptus globulus populations**. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 2002, **32**(1):59-66.

7. Holman JE, Hughes JM, Fensham RJ: **A morphological cline in Eucalyptus: a genetic perspective**. *Molecular Ecology* 2003, **12**(11):3013-3025.

8. Tripiana V, Bourgeois M, Verhaegen D, Vigneron P, Bouvet JM: **Combining microsatellites, growth, and adaptive traits for managing in situ genetic resources of Eucalyptus urophylla**. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 2007, **37**(4):773-785.

9. Payn KG, Dvorak WS, Janse BJH, Myburg AA: Microsatellite diversity and genetic structure of the commercially important tropical tree species Eucalyptus urophylla, endemic to seven islands in eastern Indonesia. *Tree Genetics & Genomes* 2008, 4(3):519-530.

10. Butcher PA, Williams ER: Variation in outcrossing rates and growth in Eucalyptus camaldulensis from the Petford Region, Queensland; Evidence of outbreeding depression. *Silvae Genetica* 2002, 51(1):6-12.

11. Shepherd M, Sexton TR, Thomas D, Henson M, Henry RJ: **Geographical and historical determinants of microsatellite variation in Eucalyptus pilularis**. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 2010, **40**(6):1051-1063.

12. McKinnon GE, Potts BM, Steane DA, Vaillancourt RE: Population and phylogenetic analysis of the cinnamoyl coA reductase gene in Eucalyptus globulus (Myrtaceae). *Australian Journal of Botany* 2005, 53(8):827-838.

13. Novaes E, Drost DR, Farmerie WG, Pappas GJ, Grattapaglia D, Sederoff RR, Kirst M: **High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome**. *Bmc Genomics* 2008, **9**.

14. Kulheim C, Yeoh SH, Maintz J, Foley WJ, Moran GF: Comparative SNP diversity among four Eucalyptus species for genes from secondary metabolite biosynthetic pathways. *Bmc Genomics* 2009, 10.

15. Sultan SE: Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science* 2000, 5(12):537-542.

16. Stape JL, Binkley D, Ryan MG: Production and carbon allocation in a clonal Eucalyptus plantation with water and nutrient manipulations. *Forest Ecology and Management* 2008, 255(3-4):920-930.

17. Campion JM, Nkosana M, Scholes MC: Biomass and N and P pools in above- and below-ground components of an irrigated and fertilised Eucalyptus grandis stand in South Africa. *Australian Forestry* 2006, 69(1):48-57.

18. Eyles A, Pinkard EA, Mohammed C: Shifts in biomass and resource allocation patterns following defoliation in Eucalyptus globulus growing with varying water and nutrient supplies. *Tree Physiology* 2009, 29(6):753-764.

19. Tatagiba SD, Pezzopane JEM, Reis EFd, Penchel RM: **Performance of six clones of eucalyptus in response to substrate water availability**. *Engenharia na Agricultura* 2009, **17**(3):179-189.

20. Shvaleva AL, Silva FCE, Breia E, Jouve L, Hausman JF, Almeida MH, Maroco JP, Rodrigues ML, Pereira JS, Chaves MM: **Metabolic responses to water deficit in two Eucalyptus globulus clones with contrasting drought sensitivity**. *Tree Physiology* 2006, **26**(2):239-248.

21. Pita P, Pardos JA: Growth, leaf morphology, water use and tissue water relations of Eucalyptus globulus clones in response to water deficit. *Tree Physiology* 2001, 21(9):599-607.

22. Merchant A, Peuke AD, Keitel C, Macfarlane C, Warren CR, Adams MA: Phloem sap and leaf delta C-13, carbohydrates, and amino acid concentrations in Eucalyptus globulus change systematically according to flooding and water deficit treatment. *Journal of Experimental Botany* 2010, 61(6):1785-1793.

23. Tsuchihira A, Hanba YT, Kato N, Doi T, Kawazu T, Maeshima M: Effect of overexpression of radish plasma membrane aquaporins on water-use efficiency, photosynthesis and growth of Eucalyptus trees. *Tree Physiology* 2010, 30(3):417-430.

24. Harb A, Pereira A: **Screening Arabidopsis Genotypes for Drought Stress Resistance**. *Plant Reverse Genetics: Methods and Protocols* 2011:191-198.

25. Deyholos MK: **Making the most of drought and salinity transcriptomics**. *Plant Cell and Environment* 2010, **33**(4):648-654.

26. Lefebvre V, Kiani SP, Durand-Tardif M: **A Focus on Natural Variation for Abiotic Constraints Response in the Model Species Arabidopsis thaliana**. *International Journal of Molecular Sciences* 2009, **10**(8):3547-3582.

27. Cohen D, Bogeat-Triboulot MB, Tisserant E, Balzergue S, Martin-Magniette ML, Lelandais G, Ningre N, Renou JP, Tamby JP, Thiec Dl *et al*: **Comparative transcriptomics of drought responses in Populus: a meta-analysis of genome-wide expression profiling in**

mature leaves and root apices across two genotypes. *BMC Genomics* 2010, **11**(630):(12 November 2010).

28. Berta M, Giovannelli A, Sebastiani F, Camussi A, Racchi ML: Transcriptome changes in the cambial region of poplar (Populus alba L.) in response to water deficit. *Plant Biology* 2010, 12(2):341-354.

29. Hamanishi ET, Raj S, Wilkins O, Thomas BR, Mansfield SD, Plant AL, Campbell MM: **Intraspecific variation in the Populus balsamifera drought transcriptome**. *Plant Cell and Environment* 2010, **33**(10):1742-1755.

30. Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, Marden JH: **Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing**. *Molecular Ecology* 2008, **17**(7):1636-1647.

31. Torres TT, Metta M, Ottenwalder B, Schlotterer C: **Gene expression profiling by massively parallel sequencing**. *Genome Research* 2008, **18**(1):172-177.

32. Barakat A, DiLoreto DS, Zhang Y, Smith C, Baier K, Powell WA, Wheeler N, Sederoff R, Carlson JE: Comparison of the transcriptomes of American chestnut (Castanea dentata) and Chinese chestnut (Castanea mollissima) in response to the chestnut blight infection. *BMC Plant Biology* 2009, 9(51):(9 May 2009).

33. Alagna F, D'Agostino N, Torchia L, Servili M, Rao R, Pietrella M, Giuliano G, Chiusano ML, Baldoni L, Perrotta G: **Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development**. *BMC Genomics* 2009, **10**(399):(26 August 2009).

34. Traas J, Vernoux T: **The shoot apical meristem: the dynamics of a stable structure**. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 2002, **357**(1422):737-747.

35. Lopez-Juez E, Dillon E, Magyar Z, Khan S, Hazeldine S, de Jager SM, Murray JAH, Beemster GTS, Bogre L, Shanahan H: **Distinct light-initiated gene expression and cell cycle programs in the shoot apex and cotyledons of Arabidopsis**. *Plant Cell* 2008, **20**(4):947-968.

36. Keller G, Marchal T, SanClemente H, Navarro M, Ladouce N, Wincker P, Couloux A, Teulieres C, Marque C: **Development and functional annotation of an 11,303-EST collection from Eucalyptus for studies of cold tolerance**. *Tree Genetics & Genomes* 2009, **5**(2):317-327.

37. Rengel D, San Clemente H, Servant F, Ladouce N, Paux E, Wincker P, Couloux A, Sivadon P, Grima-Pettenati J: **A new genomic resource dedicated to wood formation in Eucalyptus**. *BMC Plant Biology* 2009, **9**(36):(27 March 2009).

38. Faria DA, Mamani EMC, Pappas MR, Pappas GJ, Grattapaglia D: A Selected Set of EST-Derived Microsatellites, Polymorphic and Transferable across 6 Species of Eucalyptus. *Journal of Heredity* 2010, 101(4):512-520.

39. Mizrachi E, Hefer CA, Ranik M, Joubert F, Myburg AA: De novo assembled expressed gene catalog of a fast-growing Eucalyptus tree produced by Illumina mRNA-Seq. *BMC Genomics* 2010, 11(681):(1 December 2010).

40. Reid KE, Olsson N, Schlosser J, Peng F, Lund ST: An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biology* 2006, 6(Nov.):6 (Nov.) 11pp. http://www.biomedcentral.com/1471-2229/1476/1427.

41. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen ZT *et al*: **Genome sequencing in microfabricated high-density picolitre reactors**. *Nature* 2005, **437**(7057):376-380.

42. Ueno S, Provost Gl, Leger V, Klopp C, Noirot C, Frigerio JM, Salin F, Salse J, Abrouk M, Murat F *et al*: **Bioinformatic analysis of ESTs collected by Sanger and pyrosequencing methods for a keystone forest tree species: oak**. *BMC Genomics* 2010, **11**(650):(23 November 2010).

43. Pertea G, Huang XQ, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B *et al*: **TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets**. *Bioinformatics* 2003, **19**(5):651-652.

44. Wang LK, Feng ZX, Wang X, Wang XW, Zhang XG: DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 2010, 26(1):136-138.

45. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing**. *Journal of the Royal Statistical Society Series B-Methodological* 1995, **57**(1):289-300.

46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic Local Alignment Search Tool**. *Journal of Molecular Biology* 1990, **215**(3):403-410.

47. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A: **High-throughput functional annotation and data mining with the Blast2GO suite**. *Nucleic Acids Research* 2008, **36**(10):3420-3435.

48. Usadel B, Poree F, Nagel A, Lohse M, Czedik-Eysenberg A, Stitt M: **A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize**. *Plant Cell and Environment* 2009, **32**(9):1211-1229.

49. Vega-Arreguin JC, Ibarra-Laclette E, Jimenez-Moraila B, Martinez O, Vielle-Calzada JP, Herrera-Estrella L, Herrera-Estrella A: **Deep sampling of the Palomero maize transcriptome by a high throughput strategy of pyrosequencing**. *BMC Genomics* 2009, **10**(299):(6 July 2009).

50. Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF *et al*: **Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences**. *Botanical Journal of the Linnean Society* 2000, **133**(4):381-461.

51. Lynch M, Walsh JB: **Genetics and Analysis of Quantitative Traits**. Sunderland, MA; 1998.

52. Ekblom R, Galindo J: Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 2010, aop(current).

53. Metzker ML: APPLICATIONS OF NEXT-GENERATION SEQUENCING Sequencing technologies - the next generation. *Nature Reviews Genetics* 2010, 11(1):31-46.

54. Zerbino DR, Birney E: Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 2008, 18(5):821-829.

55. Morozova O, Marra MA: Applications of next-generation sequencing technologies in functional genomics. *Genomics* 2008, 92(5):255-264.

56. Zhou X, Su Z, Sammons RD, Peng YH, Tranel PJ, Stewart CN, Yuan JS: **Novel software package for cross-platform transcriptome analysis (CPTRA)**. *Bmc Bioinformatics* 2009, **10**.

57. Wall PK, Leebens-Mack J, Chanderbali AS, Barakat A, Wolcott E, Liang HY, Landherr L, Tomsho LP, Hu Y, Carlson JE *et al*: **Comparison of next generation sequencing technologies for transcriptome characterization**. *BMC Genomics* 2009, **10**(347):(1 August 2009).

58. Nowrousian M: Next-Generation Sequencing Techniques for Eukaryotic Microorganisms: Sequencing-Based Solutions to Biological Problems. *Eukaryotic Cell* 2010, 9(9):1300-1310.

59. Hegarty MJ, Barker GL, Brennan AC, Edwards KJ, Abbott RJ, Hiscock SJ: **Changes to gene expression associated with hybrid speciation in plants: further insights from transcriptomic studies in Senecio**. *Philosophical Transactions of the Royal Society B-Biological Sciences* 2008, **363**(1506):3055-3069.

60. Pavey SA, Collin H, Nosil P, Rogers SM: **The role of gene expression in ecological speciation**. *Year in Evolutionary Biology* 2010, **1206**:110-129.

61. Potts BM, Dungey HS: Interspecific hybridization of <i>Eucalyptus</i>: key issues for breeders and geneticists. *New Forests* 2004, 27(2):115-138.

62. Shinozaki K, Yamaguchi-Shinozaki K: **Gene networks involved in drought stress response and tolerance**. *Journal of Experimental Botany* 2007, **58**(2):221-227.

63. Bray EA: **Plant responses to water deficit**. *Trends in Plant Science* 1997, **2**(2):48-54.

64. Beaudoin N, Serizet C, Gosti F, Giraudat J: **Interactions between abscisic acid and ethylene signaling cascades**. *Plant Cell* 2000, **12**(7):1103-1115.

65. Sharp RE, LeNoble ME: **ABA, ethylene and the control of shoot and root growth under water stress**. *Journal of Experimental Botany* 2002, **53**(366):33-37.

66. Voisin AS, Reidy B, Parent B, Rolland G, Redondo E, Gerentes D, Tardieu F, Muller B: Are ABA, ethylene or their interaction involved in the response of leaf growth to soil water deficit? An analysis using naturally occurring variation or genetic transformation of ABA production in maize. *Plant Cell and Environment* 2006, 29(9):1829-1840.

67. Hong-Bo S, Li-Ye C, Ming-An S: Calcium as a versatile plant signal transducer under soil water stress. *Bioessays* 2008, 30(7):634-641.

68. Schroeder JI, Kwak JM, Allen GJ: Guard cell abscisic acid signalling and engineering drought hardiness in plants. *Nature* 2001, 410(6826):327-330.

69. Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, Havaux M: **Early light-induced proteins protect arabidopsis from photooxidative stress**. *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(8):4921-4926.

70. Martindale JL, Holbrook NJ: **Cellular response to oxidative stress: Signaling for suicide and survival**. *Journal of Cellular Physiology* 2002, **192**(1):1-15.

71. Kitajima S, Sato F: Plant pathogenesis-related proteins: Molecular mechanisms of gene expression and protein function. *Journal of Biochemistry* 1999, 125(1):1-8.

72. Lee BR, Jung WJ, Lee BH, Avice JC, Ourry A, Kim TH: Kinetics of drought-induced pathogenesis-related proteins and its physiological significance in white clover leaves. *Physiologia Plantarum* 2008, 132(3):329-337.

73. Przymusinski R, Rucinska R, Gwozdz EA: Increased accumulation of pathogenesis-related proteins in response of lupine roots to various abiotic stresses. *Environmental and Experimental Botany* 2004, 52(1):53-61.

74. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L: **MYB transcription factors in Arabidopsis**. *Trends in Plant Science* 2010, **15**(10):573-581.

75. Yamaguchi K, Takahashi Y, Berberich T, Imai A, Takahashi T, Michael AJ, Kusano T: **A protective role for the polyamine spermine against drought stress in Arabidopsis**. *Biochemical and Biophysical Research Communications* 2007, **352**(2):486-490.

76. Warren CR, Bleby T, Adams MA: Changes in gas exchange versus leaf solutes as a means to cope with summer drought in Eucalyptus marginata. *Oecologia* 2007, 154:1-10.

77. Hajheidari M, Eivazi A, Buchanan BB, Wong JH, Majidi I, Salekdeh GH: **Proteomics uncovers a role for redox in drought tolerance in wheat**. *Journal of Proteome Research* 2007, **6**(4):1451-1460.

78. Alexandersson E, Fraysse L, Sjovall-Larsen S, Gustavsson S, Fellert M, Karlsson M, Johanson U, Kjellbom P: **Whole gene family expression and drought stress regulation of aquaporins**. *Plant Molecular Biology* 2005, **59**(3):469-484.

79. Landry CR, Oh J, Hartl DL, Cavalieri D: Genome-wide scan reveals that genetic variation for transcriptional plasticity in yeast is biased towards multi-copy and dispensable genes. *Gene* 2006, 366(2):343-351.

80. Hodgins-Davis A, Townsend JP: **Evolving gene expression: from G to E to G x E**. *Trends in Ecology & Evolution* 2009, **24**(12):649-658.

81. Chinnusamy V, Zhu J-K: **Epigenetic regulation of stress responses in plants**. *Current Opinion in Plant Biology* 2009, **12**(2):133-139.

82. Mazzucotelli E, Mastrangelo AM, Crosatti C, Guerra D, Stanca AM, Cattivelli L: Abiotic stress response in plants: When post-transcriptional and post-translational regulations control transcription. *Plant Science* 2008, 174(4):420-431.

83. Bray EA: **Molecular Responses to Water-Deficit**. *Plant Physiology* 1993, **103**(4):1035-1040.

84. Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K: **Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future**. *Current Opinion in Biotechnology* 2006, **17**(2):113-122.

Proteomic plasticity of two *Eucalyptus* genotypes under contrasted water regimes in the field

Running title: Eucalyptus proteomes under water deficit

Frank Bedon[1,6], Emilie Villar[1,2,3], Delphine Vincent[1,7], Jean-William Dupuy[4], Anne-Marie Lomenech[4], André Mabialangoma[3], Philippe Chaumeil[1], Aurélien Barré[5], Christophe Plomion[1] and Jean-Marc Gion[1,2]

[1]INRA, UMR1202 BIOGECO, 69 route d'Arcachon, F-33612, France

[2]CIRAD, UPR39, Génétique Forestière, Campus de Baillarguet TA 10C, F–34398 Montpellier Cédex 5, France

[3]CRDPI, BP1291, Pointe Noire, République du Congo

[4]Plateforme Protéome du Centre Génomique Fonctionnelle de Bordeaux, Université Bordeaux Segalen, F-33076 Bordeaux, France

[5]Centre de Bioinformatique de Bordeaux, Plateforme Génomique Fonctionnelle de Bordeaux, Université de Bordeaux 2, F-33076 Bordeaux, France

[6]*Present address*: CSIRO Plant Industry, GPO Box 1600, Canberra ACT 2601, ACT, Australia

[7]*Present address*: Research School of Biology, RN Robertson Building, The Australian National University, Canberra 0200, ACT, Australia

Correspondence: F. Bedon Fax: +61 2 6246 5000; e-mail: frank.bedon@csiro.au

1 Abstract

Water deficit affects tree growth and limits wood production. In an attempt to identify the molecular triggers of adaptation mechanisms to water deficit in *Eucalyptus,* we investigated protein expression patterns of two ecophysiologically contrasted *Eucalyptus* genotypes. They were grown in the field in either natural conditions or irrigated for seven weeks during the dry season in the Republic of Congo. At the phenotypic level, genotype (G), treatment (T) and/or GxT interaction effects were observed for above and below ground biomass-related traits. At the molecular level, changes in protein abundance were recorded in leaves (acidic pH 4-7, and basic pH 7-11, proteomes) and stems (acidic proteome) using two-dimensional gel electrophoresis. One-third of the detected protein spots displayed significant G, T and/or GxT effects, and 158 of them were identified by tandem mass spectrometry (LC-MS/MS) analysis. Thus several proteins whose molecular plasticity was genetically controlled (*i.e.* GxT effect) were revealed, highlighting adaptive mechanisms to water deficit specific to each genotype, namely cell wall modification, cell detoxification and osmoregulation. Transcript abundances corresponding to GxT proteins were also investigated by quantitative RT-PCR. These proteins represent relevant targets to improve drought resistance in this ecologically and economically important forest tree genus.

Keywords: Eucalyptus, proteome, two-dimensional gel electrophoresis, water deficit, adaptation, molecular plasticity

2 Introduction

Water deficit is a major abiotic factor which affects tree growth and wood formation (Martinez-Meier et al. 2008; Arend and Fromm 2007). Water deficit may alter plantation quality and viability in regions subject to natural cycles of dry and rainy seasons, or at temperate latitudes where climate change is predicted to result in reduced summer rainfalls and warmer winters (IPCC 2007). To adapt to such rapid changes, often within the time scale of one or only a few generations, trees need to rely predominantly on their level of plasticity. This has been well investigated at the phenotypic level (reviewed by Matesanz, Gianoli & Valladares 2010), however changes at the molecular level are far less well documented in these long-lived organisms.

During their evolution higher plants have developed a range of complex resistance mechanisms to adapt to soil moisture and vapour pressure deficit as they are sessile and incapable of escaping from stressful environments (Turner 1986). Several physiological, biochemical and molecular responses to drought stress have been identified (reviewed by Chaves, Maroco & Pereira 2003; Shao *et al.* 2008). For example, some plants maintain water potential and limit water loss due to transpiration by closing stomata, decreasing leaf area, or shedding older leaves. Such modifications affect carbon and nitrogen metabolisms and subsequently growth of the whole plant as the photosynthesis rate decreases. Plants can also maximise water uptake by modifying the biomass allocation pattern from aerial to root organs for a better soil prospection and uptake of water and nutrients (*e.g.* Li and Wang 2003). In tree stems, water deficit can affect secondary xylem cavitation and induces changes in phloem sap composition; xylem vessel size and cell wall thickness are indeed critical traits for resistance to drought, while sap is involved in biomass partitioning (Awad *et al.* 2010; Merchant *et al.* 2010).

A proteomic approach is adequate for discovering the key molecular components of abiotic stresses (reviewed by Timperio, Egidi & Zolla 2008). In perennials, proteomic investigations related to water deficit have mainly considered the model tree poplar (Plomion *et al.* 2006; Bogeat-Triboulot *et al.* 2007; He *et al.* 2008, Xiao *et al.* 2009; Bonhomme *et al.* 2009a, 2009b; Yang, Wang & Miao 2010). These studies have focused on the leaf proteome, as it represents the main active interface between the atmospheric environment and the plant. Therefore, most of the identified proteins with accumulation patterns altered by water deficit were associated to photosynthesis, photorespiration, cell homeostasis and oxidative stress.

Often two or more genotypes contrasting for their leaf physiological parameters were exposed to different soil-moisture treatments, thus highlighting the molecular plasticity of small sets of genotype-specific proteins (Xiao *et al.* 2009; Bonhomme *et al.* 2009a, 2009b; Yang *et al.* 2010).

Eucalyptus plantations are estimated to cover approximately 20 million ha over many countries in several parts of the world: South America, South Africa, Asia, Australia and South-Western Europe (Iglesias & Wistermann 2008). The *Eucalyptus* genus includes about 900 species native to Australia where a wide range of environmental conditions exists from moist temperate to hot arid zones (Boland *et al.* 2006; Specht 1970). Some *Eucalyptus* species, considered as the fastest growing trees in tropical and sub-tropical areas, are the most widely used in cultivated forests thus representing an important source of hardwood for wood, paper and charcoal industries (Kellison 2001). Several mechanisms for drought avoidance have been identified in these evergreen species (Whitehead & Beadle 2004). In previous studies it was shown that water deficit leads to the development of different physiological (*e.g.* relative water content, osmotic potential) and morphological (*e.g.* leaf area, growth) adjustments. Such modifications allow trees to maintain growth, gas exchange and water-use during dry seasons. It was also pointed out that these mechanisms varied between clones (Pita and Pardos 2001), provenances (Li *et al.* 2000; Tuomela, Koskela & Gibson 2001; Li 1998; Tuomela 1997), or species (Schulze *et al.* 2006; White, Beadle & Worledge 1996; White, Turner & Galbraith 2000; Merchant *et al.* 2007; Ngugi *et al.* 2003; Wang, Bachelard & Banks 1988).

In this study we investigate the phenotypic and molecular response to water deficit of two *Eucalyptus* genotypes (1-41 and 18-50) used for reforestation in the coastal plain of the Republic of Congo. These genotypes have been shown to differ in terms of productivity and water-use efficiency at adult stage (Safou-Matondo *et al.* 2005; Nouvellon Y., unpublished results). The objectives of our work are three-fold: (i) evaluate their phenotypic response in terms of above and below ground biomass production, (ii) assess the effect of the tested factors (genotypes G, treatment T) on leaf and stem proteomes, and (iii) identify specifically those proteins showing GxT interactions to reveal genotype dependent strategies to cope with water deficit in field conditions; the expression pattern of some of these at the transcript level was confirmed by RT-PCR analyses.

3 Materials and Methods

3.1 Location and environmental conditions

The study was conducted in the Atlantic coastal plain of the Republic of Congo, (4°20'S, 11°38'E, 50 m a.s.l.). The climate is tropical humid with high mean annual air humidity and temperature (85% and 24°C, respectively) with low seasonal variations. Two seasons are distinguished: the dry season from May to October which is marked by cloudiness, no rainfall and low global radiations, and the rainy season characterized by high precipitations, higher temperatures and global radiations.

Soils are characterized by their great depth, chemical poverty and texture homogeneity in the landscape, they are classified as ferralic arenosols according to the FAO classification with high sand (80-90%), low clay (8-10%) and silt (2-2.5%) contents (Trouvé *et al.* 1994). These soils have low water retention, small cationic exchange capacity and low organic matter content.

Environmental conditions (rainfall and air temperatures) were recorded every 30 minutes by a weather station (Davis Instruments Corp., Hayward, CA, USA). Cumulated rainfall values were produced and air temperature means calculated per week. Soil moisture was measured once a week by three different Time Domain Reflectometer engine (TDR; Trase system, Soil moisture, Santa Barbara, CA, USA) per plots at 15 and 50 cm depths (*i.e.* six measurements per treatment).

3.2 Plant material and experimental design

Two *Eucalyptus* genotypes (1-41 and 18-50) were used: i/ 1-41 (taxon ID: 764271) is a natural hybrid of *E. alba* Reinw. ex Blume x unknown father which appeared naturally among the first afforestation and were the most planted clone around Pointe Noire (Delwaulle 1988). This genotype represents a reference for breeding and silvicultural trials (Laclau, Bouillet & Ranger 2000; Safou-Matondo *et al.* 2005; Saint-Andre *et al.* 2005); ii/ 18-50 (taxon ID: 765255) is an artificial hybrid of *E. urophylla* S.T. Blake *x E. grandis* W. Hill ex Maiden, which is currently replacing clone 1-41 as it is more productive in industrial plantations (25.1 m^3 ha^{-1} $year^{-1}$ for 1-41 *versus* 28.5 m^3 ha^{-1} $year^{-1}$ for 18-50, Safou-Matondo *et al.* 2005).

Cuttings were produced from the two genotypes using standard procedures and then transferred into the field.

The experimental design was set up on June 1st 2007 near the village of Lamba (4°20'S, 11°38'E). Cuttings were planted at a spacing of 2.5 m x 3.2 m to constitute one block of four plots of 64 trees each (*i.e.* one plot per genotype, 1-41 *vs*. 18-50, and per treatment, irrigated IR *vs*. not irrigated NIR). Following planting, all cuttings were irrigated during the first two weeks (1L per tree and per day) to avoid transplantation stresses. In the mean time, they received starter fertilization (200g of ammonitrate CAN® per tree) and insecticide treatment (pyrethrenoids) to prevent cricket attack. Herbaceous vegetation was controlled manually. During the dry season, watering was applied for seven weeks (starting on July 31st 2007) to half of the cuttings (irrigated treatment). Irrigation was performed by aspersion to compensate for potential evapotranspiration losses, estimated at 25mm per week (Bouillet JP, personal communication). On September 15th 2007 trees were sampled for phenotypic measurements and molecular analysis.

3.3 Measurements of phenotypic parameters

Total height was measured on 24 cuttings located in the middle of each plot buffered by 40 surrounding trees. Then, the 10 closest cuttings to the median were selected, from which leaves and stems were sampled and dried at 65°C for seven days to estimate above-ground biomass. Dried below-ground biomass was measured on 6 cuttings whose stumps and roots were excavated from the soil at 1m deep. Total leaf number was measured on 9 cuttings. Leaves were collected and scanned flat against a white background. Images were then imported and processed using Scion Image software (Scion Corporation, Frederick, MD, USA) to measure leaf surface area.

Stomatal density was determined based on 6 leaves from 4 cuttings per condition (2 genotypes x 2 treatments). Using transparent nail varnish and tape, a print of each side of leaves (adaxial and abaxial) was performed (Ceulemans, Van Praet & Jiang 1995). These prints were observed using an Olympus BM2 RFCA light microscope (Olympus America Inc., Melville, New York, USA) and 5 microscopic fields per print of leaves (*i.e.* n=30 per leaves side and per condition) at a 20X magnification (representing 0.568 mm^2) were analyzed.

3.4 Protein extraction and assay

Total proteins were extracted using 500 mg of fresh leaf tissues (*i.e.* full expanded leaves without neither the central mibrid nor the petiole) or stems (*i.e.* young lignified terminal stems) following the TCA/acetone procedure described by Damerval *et al.* (1986), and modified by Gion *et al.* (2005), except that the "TCT" resolubilization buffer contained 8% CHAPS instead of 4% (*i.e.* 7 M urea, 2 M thiourea, 0.4% v/v Triton X-100, 8% w/v CHAPS, 10 mM DTT and 1% v/v IPG buffer). Leaves and young stems harvested from 4 different plants per treatment (IR and NIR) and per genotype (1-41 and 18-50) were used for proteins extraction (*i.e.* 16 independent extractions per organ). Protein content assay were performed in triplicate for each extract using a modified Bradford assay, as described by Ramagli & Rodriguez (1985). Averaged concentrations were then used to pool equal protein amounts from each of the four biological replicates in order to highlight treatment and genotype effects instead of biological variations due to the sampling in non controlled environments (Karp & Lilley 2009). Pooled protein extracts were then used for the subsequent two-dimensional electrophoresis (2-DE) steps. Four technical replicates were produced.

3.5 2-DE and gel staining

2-DE was used to analyse total soluble proteins from leaf and young stem samples following the procedure of Bahrman *et al.* (1997). For the isoelectric focusing (IEF), 24 cm immobilized pH gradients (IPG) strips (Immobiline DryStrip, GE Healthcare, Uppsala, Sweden) were used with a linear 4-7 pH gradient (4-7L) for both leaves and young stems, and a non linear 7-11 pH gradient (7-11NL) for leaves only. Basic proteins were not investigated in stems due to insufficient protein contents. For the acidic strips (4-7L), 600 µg of proteins from leaves and 800 µg from young stems were resuspended into 470 µL of "TCT 4-7" solution (TCT with 1% 4-7 IPG buffer). Acidic strips were passively rehydrated with 470 µL of protein samples for one hour at room temperature prior to the IEF run which was performed using the IPGphor system (Amersham), according to Gion *et al.* (2005). For the basic strips (7-11NL), 600 µg of proteins from leaves were resuspended into 150 µL of "TCT 7-11" solution (TCT with 1% 7-10 IPG buffer). Basic strips were passively rehydrated in an Immobiline DryStrip reswelling tray (GE Healthcare) by adding 470 µL of Destreak Rehydration Solution (GE Healthcare) with additional IPG buffer (1% pH 7-11NL and 0.5% pH 3-10NL) for one hour at

room temperature, then covered with mineral oil and left to further rehydrate overnight for twelve hours. Cup loading and IEF run for basic IPG strip were carried out as described in Vincent *et al.* (2009). SDS-PAGE and Colloidal Coomassie Blue gel staining were performed as described in Gion *et al.* (2005).

3.6 Image analysis, spot detection and statistical analysis

Image acquisition of the stained gels was done as described in Gion *et al.* (2005). Image analysis was performed in two steps: *i*) gel image alignment and spot matching between gels using Progenesis SameSpot software (Nonlinear Dynamics Ltd, Durham, NC, USA), and *ii*) spots detection and analysis using Progenesis PG240 software (Nonlinear Dynamics). Background subtraction (lowest on boundary) was applied to all images. Spots normalization was applied following the equation of the spot volume normalization procedure of Progenesis SameSpots software: (Volume of each spot / Volume of all the spots)*10^6.

Normalized volumes were used to perform two-way analysis of variance (ANOVA) following the linear model:

$$Y_{ijk}=\mu+G_i+T_j+(GxT)_{ij}+\varepsilon_{ijk}\ ,$$

where Y_{ijk} denoted protein abundance of Genotype *i* (i= 18-50 or 1-41) and Treatment *j* (j=IR or NIR), for the kth technological replicate. The interaction term $(GxT)_{ij}$ accounted for the interaction between the main factors. A false discovery rate (FDR) procedure was applied on *p*-values to minimize the number of false positives (Benjamini & Hochberg 1995). Effects were declared significant at a *p*-adjusted value<0.05. This analysis was performed using R cran version 2.6.2 using aov and p.adjust functions (R Development Core Team, 2008). Normalized volumes were standardised prior to hierarchical clustering which was performed using Expander software (Ulitsky *et al.* 2010) and the following parameters ((1-Pearson correlation)/2; average linkage).

3.7 Protein identification by LC-MS/MS

A total of 158 spots presenting at least one significant effect (G, T and/or GxT) were manually excised and processed as described in Vincent *et al*. (2009). Briefly, distained spots were trypsin-digested and peptide mixtures were analysed using online nanoHPLC (LC

Packings, Amsterdam, Netherlands) coupled to a nanospray LCQ Deca XP Plus ion trap mass spectrometer (Thermo-Finnigan, San Jose, CA, USA).

Peptides were identified with SEQUEST through the Bioworks 3.3.1 interface (Thermo-Finnigan, Torrence, CA, USA) using the 1,315,726 indexed entries of the constructed *Eucalyptus* database including expressed sequence tags (ESTs) from Novaes *et al.* (2008) and from public databases (downloaded from NCBI). ESTs were translated in all six reading frames using the transeq program of the sequence analysis package EMBOSS. Positive matches were blasted against the Swiss-Prot database for identification and then against the non redundant (nr, *viridiplantae*) database of NCBI using tblastx when unsuccessful results were obtained (for 9 proteins, see Table S1). All the spectra generated in this experiment and the peptide sequences identified have been submitted to the proteomics identification database PRIDE (Martens *et al.* 2005; Barsnes *et al.* 2009), with accession numbers 12916-13073 inclusive (see Table S1).

4 Results

4.1 Phenotypic response to water deficit

In the irrigated (IR) treatment, water supply consisted of 25 mm per week for seven weeks and resulted in about 12-14% of soil moisture at 15 and 50 cm depths at the sampling point (Fig. 1). In the non irrigated (NIR) treatment, soil moisture reached 5% and 8% at 15 and 50 cm depths, respectively. Since about 90% of the root biomass was located between 0 and 50 cm deep regardless of the genotypes and treatments, one can conclude that water availability was reduced by at least half under the NIR condition.

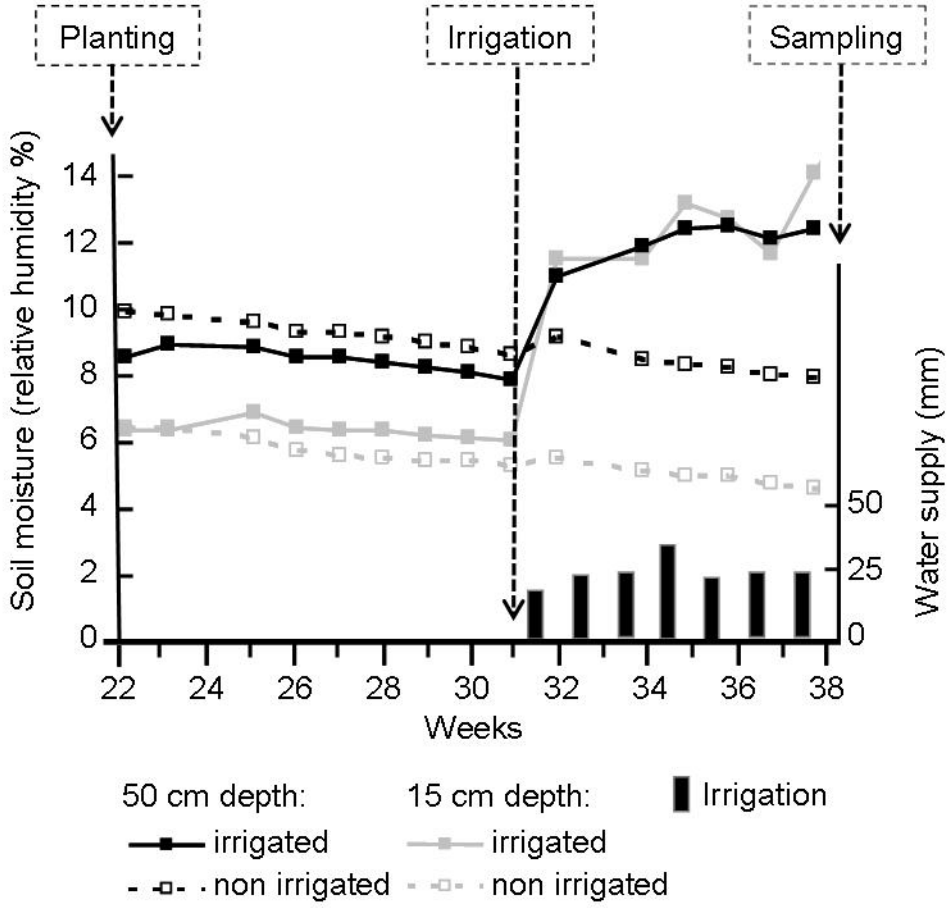


Figure 1. Weekly measurements of water supply (irrigation) and soil moisture at 15 and 50 cm depth for both irrigated (IR) and non irrigated (NIR) treatments during the dry season.

We used two-way ANOVA (*P*adjust < 0.05) to highlight the effects of genotype "G" (1-41 *vs.* 18-50), treatment "T" (IR *vs.* NIR) and their interaction "GxT" (Table 1) on the traits measured.

Table 1. Phenotypic variables of four month-old eucalyptus genotype 1-41 and 18-50 under irrigated (IR) and non irrigated (NIR) treatments.

	Genotype 1-41		Genotype 18-50		ANOVA
	IR	NIR	IR	NIR	effects
General variables					
Ht, shoot height (m) (n=24)	0,63 ± 0,11 a	0,49 ± 0,07 b	0,58 ± 0,15 a	0,47 ± 0,08 b	T***
Ldw, leaf dry weight (g) (n=10)	23,27± 6,86 a	13,36 ± 4,85 b	21,80 ± 5,14 a	11,94 ± 1,38 b	T***
Sdw, stem dry weight (g) (n=10)	9,41 ± 3,04 a	4,30 ± 1,47 b	11,02 ± 2,79 a	5,64 ± 0,82 b	G*/T***
Rdw, root dry weight (g) (n=6)	5,83 ± 1,31a	3,08 ± 0,91 b	6,80 ± 1,15 a	4,00 ± 0,54 b	G*/T***
Rs, ratio [root / shoot] (n=6)	0,17 ± 0,02 b	0,20 ± 0,03 ab	0,20 ± 0,04 ab	0,24 ± 0,03 a	G*/T*
Leaves variables					
Sab, abaxial stomate density (mm^{-2}) (n=4)	113,78 ± 31,94	128,71 ± 23,27	110,10 ± 5,6	135,74 ± 8,82	n.s.
Sad, adaxial stomate density (mm^{-2}) (n=4)	143,39 ± 25,87 a	162,79 ± 25,87 a	6,86 ± 1,21 b	10,23 ± 1,60 b	G***
Tla, total leaf area (dm^2) (n=9)	33,5 ± 9,08 a	12,64 ± 2,63 b	30,4 ± 8,18 a	14,38 ± 2,47 b	T***
Tln, total leaf number (n=9)	117,78 ± 33,20 b	69 ± 18,94 c	208,09 ± 49,08 a	138,5 ± 31,56 b	G***/T***
Mla, mean leaf area (dm^2) (n=9)	0,29 ± 0,03 a	0,19 ± 0,03 b	0,14 ± 0,01 c	0,11 ± 0,01 d	G ***/ T***/GxT***

Means ± Standard Deviation. Numbers of biological replicates (n) are indicated into brackets in the first column. Shoots from Rs included stems and leaves. For each variable, a two-way ANOVA followed by Bonferoni test was performed and values not sharing common letters are significantly different. In the last column ANOVA effects are indicated for genotype (G), treatment (T) and interaction (GxT), as well as their associated *p*-value<0.05 (*), 0.01 (**), or 0.001 (***), ns indicate no significant difference, and combined effects are symbolized by slash (/).

Firstly, the imposed irrigation treatment presented a strong effect on all growth variables, with increased shoot height (Ht), root (Rdw) and shoot (Ldw and Sdw) dry weight, total leaf area (Tla) and total leaf number (Tln), both of which resulting in a higher SLA (specific leaf area Tla/ Ldw). Secondly, while both genotypes had similar Tla, they presented opposite strategies in terms of leaf number (Tln), leaf area (Mla) and leaf dry weight (Ldw), with genotype 18-50 producing more leaves but shorter and thinner than those of genotype 1-41. Finally, a strong GxT interaction was observed for Mla suggesting that both clones exhibited a contrasted behaviour between the two treatments.

4.2 Analysis of protein abundance variation in leaves and stems

The extent of molecular plasticity of both genotypes was assessed in leaf and stem based on 1,133 and 715 reproducible protein spots revealed by CBB-stained two-dimensional gel

electrophoresis (2-DE), respectively. Leaf proteins were separated across two pH gradients, 4-7 (acidic leaf proteome symbolised by aL, 901 spots, Fig. S1. A-B) and 7-11 (basic leaf proteome symbolised by bL, 232 spots, Fig. S1. C-D). Stem proteins (stem proteome symbolised by S, 715 spots, Fig. S1. E-F) were analysed using an acidic 4-7 pH gradient only. Two-way ANOVA (*P*adjust < 0.05) with false discovery rate (FDR) correction for multiple testing revealed that a total of 593 spots (aL: 294 spots, 32.6%; bL: 59 spots, 25.4%; S: 240 spots, 33.5%) displayed significant G, T and/or GxT effects (Fig. 2).

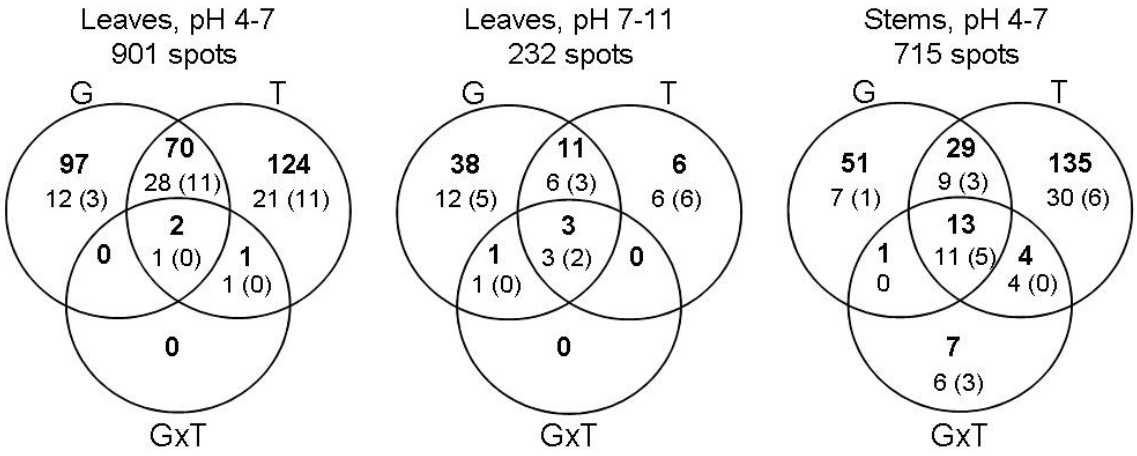


Figure 2. Venn diagrams showing the distribution of significant spots following two-way ANOVA.
Total reproducible spots identified from 2-D gels of each experiment are indicated (acidic and basic proteins from leaves, and acidic proteins from stems). In bold is the total number of significant protein spots following a two-way ANOVA (*P*adjust < 0.05) with genotype (G), treatment (T) and/or genotype by treatment (GxT) effects; below is the number of excised/sequenced protein spots, and into brackets the number of protein spots with single protein identification after LC-MS/MS analysis.

Overall, taking into account all the 2-D gels, 265 (14.3%), 186 (10.1%) and 7 (0.4%) spots displayed pure T, G and GxT effect, respectively, showing that beside genotypic differences the treatment had an important effect on protein accumulation. In addition, a total of 110 spots (5.9%) presented combined G and T effects. These few proteins with differential expression between genotypes and involved in the stress response are of little value to breeders. Finally, a rather small number of proteins presented GxT effect, either in combination with other main effects (32 spots, 1.7%) or alone (7 spots, 0.4%), highlighting genotype specific molecular strategies in the different conditions; these proteins are potential breeding targets. These proportions varied according to the tissue type and the pH gradient used (Fig. 2), with a higher level of molecular plasticity revealed in stems relative to leaves (135 acidic spots, 18.9%, *vs*. 124 acidic spots, 13.8%, with T effect only), and with a higher percentage of spots

displaying genotypic variation for basic proteins in leaves (38 spots, 16.4%, for pH 7-11 *vs*. 97 spots, 10.8%, for pH 4-7).

The distribution of the fold-change (FC) ratio (NIR/IR) for the 593 spots with at least one significant effect indicates that more proteins of genotype 1-41 ranked between FC -1.5 to 1.25 compared to genotype 18-50 proteins mostly distributed outside this FC range (Fig. S2). The difference in protein abundance levels between both genotypes was more important in leaves than in stems (Fig. S3). These results suggest that 18-50 presented a higher degree of plasticity in terms of protein abundance. It should also be noted that most of the spots were found to be down-regulated in the NIR treatment regardless of genotypes and organs. When considering FC outside the arbitrary cut-offs of -1.75 and 1.75 (representing 191 of the 593 significant spots, *i.e.* 32.7%), the proportion of down-regulated spots reached about 70-80% (see inset in Fig. S2). Such down-regulation in protein abundance has been previously reported in poplar leaves submitted to drought stress, indicating that water deficit negatively affects a range of biochemical processes (Bonhomme *et al.* 2009a; Plomion *et al.* 2006). In Figure 3, we plotted the FC ratio of these 191 spots for both genotypes according to the type of ANOVA effect. This chart reveals that: i/ 60.7% (116/191) of these highly responding proteins were significantly affected by water deficit alone and 23.5% (45/191) displayed opposite accumulation profile under the dry condition in each genotype, ii/ 71.7%(137/191) of these proteins were down-regulated (FC<-1.75) when water availability was limited, and iii/ the amplitude of the water deficit-response was higher for genotype 18-50 proteins than that of genotype 1-41, especially in leaves.

4.3 Identification of differentially accumulated proteins

Subsequent to the statistical analysis of spot abundance, 158 of the most abundant spots leading to significant G, T and/or GxT effects (P*adjust*<0.05) were manually excised from 2-D gels of leaves (63 acidic spots and 28 basic spots), and stems (67 acidic spots), and further analysed by nLC-ESI-MS/MS (Fig. 2, Table S1). No conclusive identification was obtainable for 5 spots (3%) based on our criteria, probably due to low quality spectra and/or incomplete database coverage. A total of 58 spots (40%) contained a single protein. Their detailed analysis and function are given in Table 2 (Supplied at the end of the manuscript). In the following sections, we only considered these proteins for biological interpretations of spot volume variation, thereby allowing for a true quantitative approach (Mechin *et al.* 2004; Zhu *et al.* 2005; Yang *et al.* 2007; Vincent *et al.* 2007). The hundred other spots presented multiple

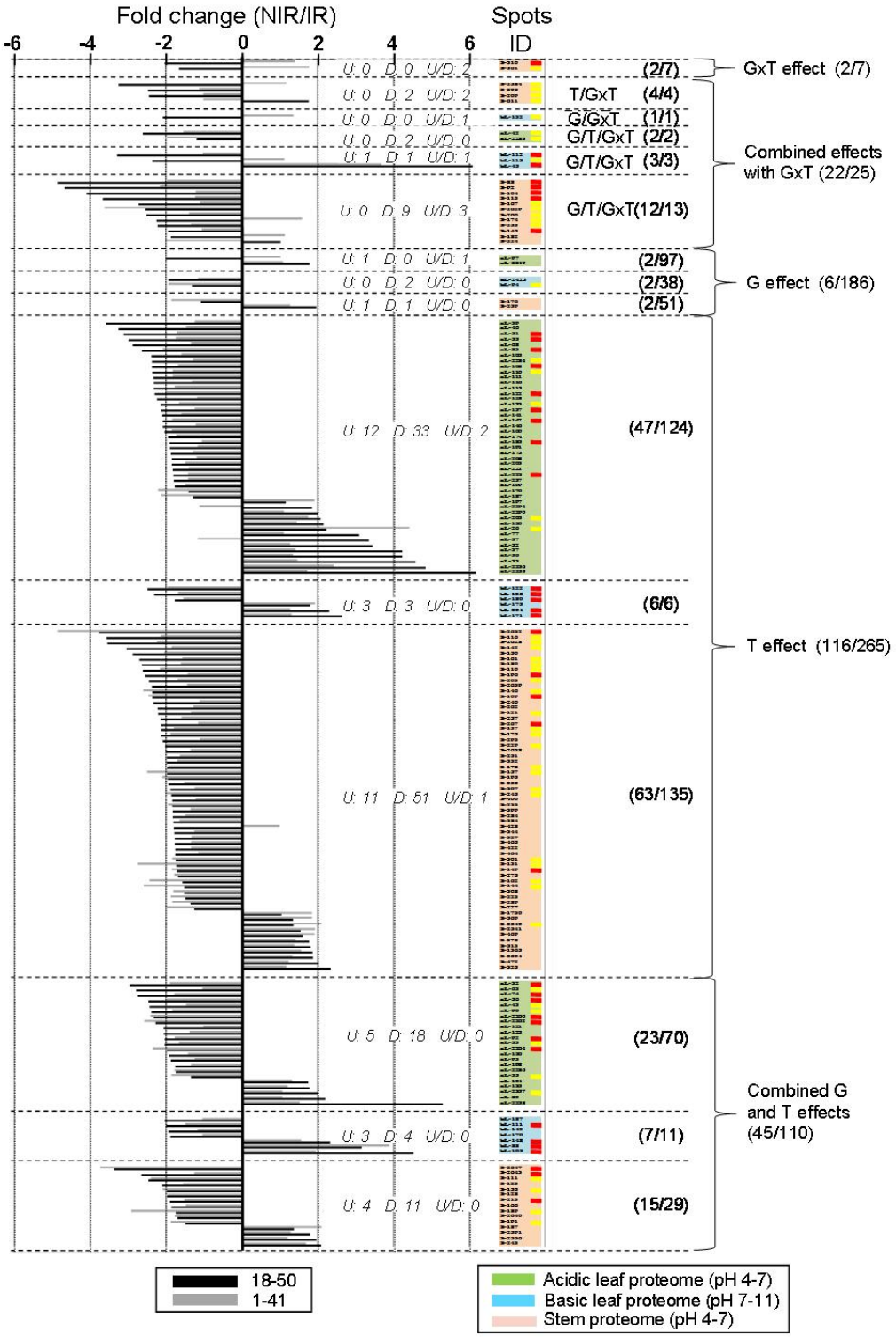


Figure 3. Fold-change ratios between non irrigated (NIR) and irrigated (IR) treatments, and associated ANOVA effects for 191 significant spots (-1.75>FC>1.75).

Fold changes (-1.75>FC>1.75, observed at least in one genotype) based on spots normalized volumes for each genotypes (18-50 and 1-41) from each 2-D gels experiment (Stem: S, pH 4-7; acidic leaves proteins: aL, pH 4-7; and basic leaves proteins: bL, pH 7-11) were used to construct the graph. Rectangle colors on the right of the spot ID indicate one protein (red) or multiple proteins (yellow) identification. In italic are presented the numbers of up- (U), down- (D), and up- and down-regulated (U/D) spots for each 2-D experiment and each ANOVA effect. Into brackets are indicated the number of corresponding spots *versus* the total number of spots with ANOVA effect without FC cut-off. ANOVA effects (*P*adjust < 0.05) are indicated for any groups of protein spots from each 2-D gel experiment (G, genotype; T, treatment; GxT, interaction).

protein identifications (ranging from 2 to 9 proteins) and required further experiments, such as spectral counting to determine the proportion of each protein per spot (Lundgren *et al.* 2010). Of these 58 proteins, 44 showed a decrease in abundance in the NIR treatment in both genotypes, while 6 increased and 8 presented opposite expression patterns between both genotypes (Table 2).

These 58 proteins were distributed into 9 groups based on functional categories using the UniProtKB database (http://www.uniprot.org/uniprot/). Several categories were represented for each class of effect (Table 2). Carbohydrate metabolism was the most represented with 16 proteins, followed by energy metabolism (10 proteins among which 6 correspond to ATP synthase subunits), amino acid metabolism (8), protein fate (8), stress/defence (7), nucleotide metabolism (3 adenosine kinases, ADK), structure (1 actin and 4 tubulins), and secondary metabolism (caffeic acid 3-O-methyltransferase). Several spots corresponding to the same protein were found with similar experimental p*I* or MW indicating potential isoforms (*e.g.* adenosine kinase 2 in spot #S-145, #S-109 and #S-310). Additionally, an ATP synthase subunit beta (UniprotKB Entry Q49KZ1) was identified in both acidic proteome of stem (spot #S-113) and leaf (spot #aL-51) with similar p*I* and MW suggesting a constitutive expression across organs.

Multivariate analysis using hierarchical clustering was carried out to group these 58 proteins according to their expression profiles. Six clusters, with 2 to 30 proteins, were obtained (Clusters I to VI, Fig. 4). These clusters were illustrated by reaction norm diagrams consisting of a graphical representation highlighting the contribution of each condition (IR and NIR) to the observed protein abundance for each genotype (18-50 and 1-41). Cluster I comprised 3 proteins (spots #S-694 Quinone reductase-like, #S-586 protein disulfide isomerase, #S-310 adenosine kinase 2) showing a clear signature of GxT interaction, the two genotypes presenting contrasted behaviour in IR *vs*. NIR treatments. The 9 proteins of cluster II were highly abundant in genotype 18-50 under IR treatment only. This trend was particularly strong for 5 of them therefore presenting also significant GxT interaction (spots #S-145, Adenosine kinase 2; #S-88, Actin-7; #S-104, Caffeic acid 3-O-methyltransferase; #S-113, ATP synthase subunit beta cp and #bL-112, Serine glyoxylate aminotransferase). These first two clusters highlighted the genotype-specific molecular strategies in the different conditions.

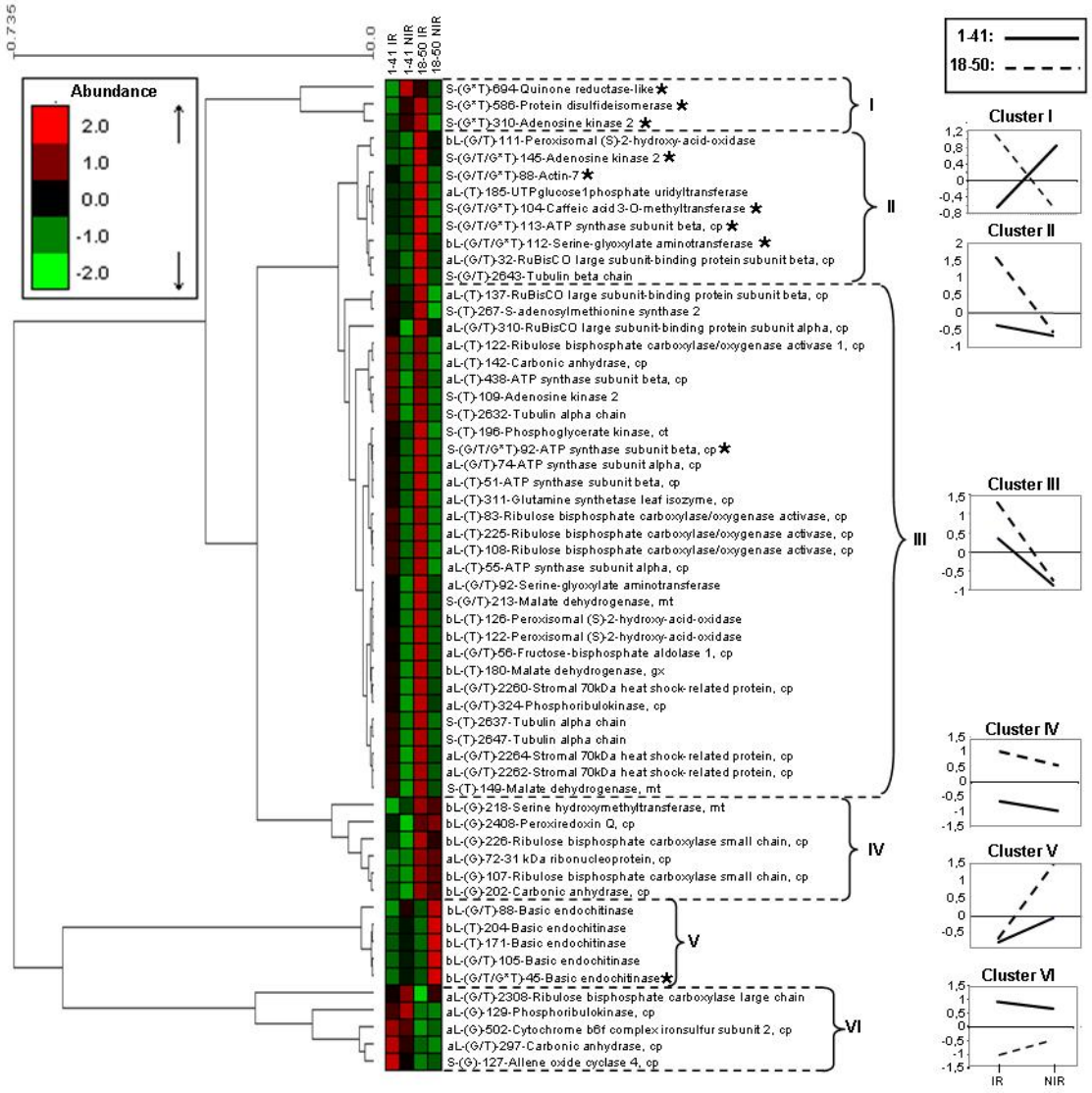


Figure 4. Hierarchical clustering of protein abundance and associated reaction norms of 58 significant protein spots identified with a single protein.

Each line of the hierarchical cluster corresponds to a protein spot identification with its origin, *i.e.* acidic proteins from leaves (aL, 4-7 pH range) and basic proteins from leaves (bL, 7-11 pH range) and stem (S, 4-7 pH range), its ANOVA effect (G, genotype; T, treatment, and GxT interaction also highlighted by an asterisk) and the spot number. Each column represents the mean of four 2-D gel replicates for genotypes 1-41 and 18-50 in either irrigated (IR) or non irrigated (NIR) condition. Protein decrease and increase under water deficit are indicated in green and red, respectively. Graphs of the reaction norms were generated with the means of each standardized normalized spot volumes for each defined cluster. Abbreviations for mitochondrial, chloroplastic, glyoxysomal and cytosolic are respectively mt, cp, gx and ct.

The largest cluster (cluster III with 30 proteins) presented reaction norms similar to those of cluster II, but with higher differences in protein accumulation between both treatments, especially for genotype 1-41. Only one spot (#S-92, ATP synthase subunit beta mt) in this cluster presented a significant GxT effect.The 6 and 5 proteins of clusters IV and VI, respectively, displayed mainly G effect, *i.e.* parallel reaction norms, with distinct expression levels between genotypes. Finally, cluster V gathered proteins with strong T or G/T effects but with opposite reaction norms compared to clusters II and III, with higher protein accumulation under NIR treatment and to a greater extent for genotype 18-50. Interestingly, this cluster comprised five endochitinase isoforms, one (#aL-45) presenting a significant GxT effect.

5 Discussion

5.1 Phenotypic effect of water deficit

In both *Eucalyptus* genotypes water deficit significantly affected all the measured growth variables (*i.e.* height, biomass, leaf number and area) (Table 1). Such reduction in biomass due to water shortage has been previously reported in forest tree species including the evergreen *Eucalyptus* and the deciduous poplar (*e.g.* Merchant *et al.* 2007; Li & Wang 2003; Liang *et al.* 2006; Monclus *et al.* 2007). Water deficit reduces gas exchange and carbon assimilation by leaves, inhibits cell division and expansion, subsequently resulting in slower growth rates (Farquhar & Sharkey 1982; Lecoeur *et al.* 1995). An increase in root/shoot ratio (Rs) is usually observed under severe drought stress to maximise water uptake associated with an expansion of the root system development (*e.g.* Li & Wang 2003; Guo, Zhang & Huang 2010). In this study our results show that the NIR treatment did not lead to an increased root dry biomass. NIR treatment distinctly affected the aerial growth of both genotypes. Because 1-41 had significantly less leaves of smaller size than 18-50 in both treatments, 1-41 presented a higher ability, *i.e.* plasticity, to reduce leaf size than 18-50 under water deficit. Leaf shedding and generation of small leaves has been already reported as an advantage to cope with drought by reducing transpiration rate (Chaves *et al.* 2003). Thus, in our study both *Eucalyptus* genotypes might be distinguished by two mechanisms to cope with the dry season: leaf shedding for 1-41 and smaller leaf size for 18-50. Small leaf sizes have been described in common bean (*Phaseolus vulgaris L.*) and in Pima cotton (*Gossypium barbadense L.*) cultivars bred to better cope with dry conditions (Mencuccini & Comstock 1999; Cornish *et*

al. 1991). In *Eucalyptus*, Pita *et al.* (2003) identified the most drought-resistant *Eucalyptus globulus* clone as the one displaying the smallest leaves under both well-watered and drought-treated conditions. In the present study, such observations would designate the genotype 18-50 as the most able to deal with water shortage.

About 20 times less adaxial stomata were observed in genotype 18-50 relative to 1-41, while similar numbers of abaxial stomata were found. *E. urophylla* and potential hybrids were previously described to lack stomata almost entirely on the adaxial surface (Pryor, Williams & Gunn 1995). In the present study, we showed that these stomata density was not under the influence of water supply.

Although phenotypic data did not allow us to conclude to which genotype was more tolerant to water shortage, the evidence that both are differentially affected by the irrigation regime suggests a genetic control of the water deficit response at the biochemical and molecular levels.

5.2 Identification of water-deficit responsive proteins involved in cell wall modification, cell detoxification and osmoregulation

Leaf and stem proteome analyses aimed at identifying water deficit-responsive proteins whose patterns contrasted between the two *Eucalyptus* genotypes (*i.e.* showing GxT interaction). Among the 58 identified spots (out of 158 in total, 37%) containing single protein identification, ten proteins presented pure (3) or combined (7) interaction effects (Table 2). Eight of these 10 identified proteins were observed in stems and therefore highlighted the wood forming tissue as a powerful target to decipher the molecular mechanisms underlying tree response to water shortage. In the following sections, we discuss the role and function of this set of 10 proteins, thereby representing potential markers of the water-deficit response. Transcript levels of most of the corresponding genes were also assessed from the same plant material but revealed little correlation with protein abundance patterns (see Supplementary Data S1 and Fig. S4).

5.2.1 Alteration of cell wall components

In plant tissues, water content controls cell turgor pressure by either relaxing or tightening the wall through several physiological, biochemical and genetic mechanisms (Bartels & Sunkar 2005). As previously reported in plants elsewhere (*e.g.* Vincent *et al.* 2007), an actin (spot #S-88) was found to be down-regulated during the dry season, particularly in genotype 18-50. Actins are structural proteins essential during cellulose deposition within secondary cell wall

(Wigthman & Turner 2007). Alteration of the actin protein abundance observed in genotype 18-50 could change the cell morphology and cyto-architecture, and modify the plant ability to cope with the dry condition by maintaining or not a viable cell turgor pressure.

A basic endochitinase (EC 3.2.1.14, spot#bL-45) was the only protein with a GxT interaction effect up-regulated during the dry season in both genotypes and at high levels, particularly in genotype 18-50 (Table 2). The same trend was also observed at the transcript level (see Supplementary Data S1 and Fig. S4). Biotic and abiotic stresses, including drought, have been already reported to induce the accumulation of chitinase in plants (Hong & Hwang, 2002; Torabi *et al.* 2009). Chitinase up-regulation might also be a consequence of the secondary cell wall modification during dry conditions; indeed Hernandez-Blanco *et al.* (2007) have shown that alteration of the cellulose synthesis activated the defence pathways and the generation of an antimicrobial-enriched environment hostile to pathogens. Increased activities of chitinases are assumed to modify cell wall properties potentially important for plant survival under a variety of environmental stresses, diseases and pest attack (Pinheiro, Chaves & Ricardo 2001). Thus, the higher protein level from genotype 18-50 could help dealing with water deficit by maintaining a higher level of cell wall integrity; however further experiments are required to assess the chitinases role in *Eucalyptus* during dry conditions. Our study is the first to report the differential accumulation of chitinase in response to water shortage, possibly because this protein resides within the seldom prospected basic pH gradient (experimental p*I* of 10).

Additionally, a caffeic acid/5-hydroxyferulic 3-O-methyltransferase protein (COMT, EC 2.1.1.68, spot #S-104), an enzyme of the phenylpropanoid pathway involved in lignin biosynthesis, was found to be down-regulated in stem of both genotypes during the dry season (4 times more in genotype 18-50). Such reduction in COMT abundance, together with a reduction in lignin content of methylated monolignols, was observed in water-deficit stressed maize leaves (Vincent *et al.* 2005). Cell wall modifications induced by lower turgor pressure during water-deficit could participate in the drought tolerance mechanism by maintaining essential tissues growth (Wu & Cosgrove 2000; Vincent *et al.* 2005). In this study, the higher COMT down-regulation in NIR condition observed for genotype 18-50 would lead to less lignification and enable a better control of the cell turgor pressure, thus allowing for growth resumption when environmental conditions improve.

5.2.2 Cell detoxification and osmoregulation

Water deficit, especially under high light intensity, disrupts photosynthesis and increases photorespiration, altering the normal homeostasis of cells and causing an increased production of reactive oxygen species (ROS) (reviewed by Miller *et al.* 2010). ROS are toxic by-products of stress metabolism and also represent important signal transduction molecules during abiotic stress responses. A link between the drought-induced oxidative stress and the Ca^{2+} signalling response involving the endoplasmic reticulum (ER) homeostasis control through an unknown mechanism of ER-Ca^{2+} release control by ER-resident proteins was recently hypothesized (Valente *et al.* 2008). In the present study we identified a protein disulfide-isomerase (EC 5.3.4.1, PDI, in spot #S-146) that was down-regulated for genotype 18-50 during the dry season, while it was up-regulated in genotype 1-41. PDIs are ubiquitous enzymes catalyzing the disulfide bond formation in proteins, a step in the oxidative folding pathway located in ER of higher organisms (Gruber *et al.* 2006). In response to stress, the unfolded or misfolded accumulating proteins will generate an ER stress due to perturbation of its homeostasis (reviewed by Urade 2007). ER-resident proteins like PDI help to maintain the ER homeostasis and therefore might have a role in drought stress tolerance. The fact that this protein was up-regulated only in genotype 1-41 in the dry condition suggests a higher efficiency to control the ER-homeostasis and might improve the drought-stress tolerance by limiting protein misfolding.

A serine glyoxylate aminotransferase (EC 2.6.1.45, SGAT, spot #bL-112) was down-regulated during the dry season, especially in genotype 18-50. SGATs are key enzymes of photorespiration and catalyse the reaction of glutamate and glyoxylate to produce 2-oxoglutarate and glycine. Photorespiration is likely to increase under water-deficit stress as the intercellular CO_2 concentration and stomatal conductance are reduced. Such a mechanism would prevent photoinhibition in leaves through energy dissipation, and the resulting glycine would be used for the production of glutathione involved in the anti-oxidative system in plants (Wingler *et al.* 2000). Additionally two isoforms of adenosine kinase 2 (EC 2.7.1.20, Adenosine 5'-phosphotransferase 2, ADK) were identified. In the dry season, both isoforms were down-regulated at similar levels in genotype 18-50, but in genotype 1-41, one isoform was up-regulated (spot #S-310) and the second was less down-regulated (spot #S-145) compared to genotype 18-50. This enzyme catalyses ATP formation and is thought to maintain methylation activities during salt stress (Weretilnyk *et al.* 2001). Under drought stress, methylated compounds like glycine betaine are accumulated and function as osmolytes

to maintain cell turgor and to stabilize protein conformation and cell structure (Bartels & Sunkar, 2005). ADK was reported to be up-regulated in several plants organs under salt stress (Weretilnyk *et al.* 2001; Vincent *et al.* 2007) as well as in *Eucalyptus* roots after 36 hours of osmotic stress (Bedon *et al.* 2011). Osmotic adjustment has been considered as a crucial process in plant adaptation to drought stress (Chaves *et al.* 2003). Therefore, accumulation of SGAT and ADK for glycine production and transformation is likely to be less affected by water-deficit in genotype 1-41 compared to genotype 18-50 and might be seen as a way to improve water-deficit stress tolerance.

Two proteins corresponding to ATP synthase subunit beta from chloroplast (spot #S-113) and mitochondria (spot #S-92) were down-accumulated during the dry season. This is consistent with previous observations indicating that water deficit alters ATP production, as a consequence of ROS accumulation in response to enhanced energy transduction from the low CO_2 assimilation and photorespiration (reviewed by Lawlor & Tezara, 2009). Recently, Kohzuma *et al.* (2009) concluded that plant acclimation to long-term environmental stress involves ATP synthase as a key control point for regulating electron transfer and proton motive force. The higher down-regulation observed in genotype 18-50 compared to 1-41 suggests that both genotypes differ in their ability to detoxify ROS and maintain ATP synthase activity. Quinone reductase-like proteins (spot #S-694), which catalyses the reduction of several electron acceptors, like electrophilic quinines, have been hypothesised to play a role in the protection of plant cells from oxidative damage (Sparla *et al.* 1999). Up-regulation of these enzymes should help detoxifying toxic compounds produced under water-deficit. Induction of quinine reductase-like and ATP synthase proteins was recently observed in roots of rice and tomato grown under phosphorous deficiency and aluminium stress, respectively (Torabi *et al.* 2009; Zhou, Sauve & Thannhauser 2009). Taken together, these results indicate that the genotype 1-41 may tolerate water-deficit better than 18-50, as several enzymes that postpone cellular damages and reduce osmotic potential losses were up-regulated (PDI, spot #S-586; ADK, spot #S-310; quinone reductase, spot #S-694) or less down-regulated (ADK, spot #S-145; SGAT, spot #S-112) compared to 18-50 during the dry season.

In conclusion, two *Eucalyptus* genotypes with contrasted phenotypic variables were used to asses the genetic control of their molecular plasticity under water-deficit. Both genotypes are used in plantation in the Republic of Congo but the genotype 18-50 is currently preferred as it provides higher wood yield (Safou-Matondo *et al.* 2005). In our study, the two-month old

Eucalyptus genotypes cuttings submitted to seven weeks of dry season exhibited different foliar behaviours. Proteome analyses of leaves and stems using two-dimensional gel electrophoresis and tandem mass spectrometry identified water deficit-responsive proteins displaying genotype-specific expression. Comparison of these proteins abundance levels between irrigated and non-irrigated treatments highlighted potential drought-tolerant mechanisms involved in cell wall modification, ROS detoxification and osmoregulation during the dry season. Our data suggest the genotype 1-41 is potentially less susceptible to drought stress; although further experiments at the phenotypic, physiological and post-genomic levels are required for validation (Villar E. et al. unpublished results). Moreover, we reported several isoforms of a basic endochitinase that were up-regulated under water shortage at higher level in genotype 18-50, associating drought susceptibility with cellular damage. The set of proteins we have identified represents potential markers of the water-deficit adaptation and future targets for molecular breeding.

Acknowledgements

The authors gratefully acknowledge Dr. Matias Kirst for providing us with nucleotide sequences (454 Roche ESTs) prior publication. The authors thank Dr. Stephane Claverol (Pôle de Protéomique, Université de Bordeaux 2) for his expertise in proteomic analysis. We thank Liam Cassidy for critical suggestions on improving the manuscript. This work was supported by INRA and ANR under an ERA-NET project (Eucagen). Part of the experiments (q-PCR) presented in the present publication were performed at the Genotyping and Sequencing facility of Bordeaux (grants from the Conseil Regional d'Aquitaine n°20030304002FA and 20040305003FA and from the European Union, FEDER n°2003227).

REFERENCES

Arend M. & Fromm J. (2007) Seasonal change in the drought response of wood cell development in poplar. *Tree Physiology* **27**, 985-992.

Awad H., Barigah T., Badel E., Cochard H. & Herbette S. (2010) Poplar vulnerability to xylem cavitation acclimates to drier soil conditions. *Physiologia Plantarum* **139**, 280-8.

Bahrman N., Plomion C., Petit R.J. & Kremer A. (1997) Contribution to maritime pine genetics using proteins revealed by two- dimensional electrophoresis. *Annals of Forest Science* **54**, 225-236.

Barsnes H., Vizcaino J.A., Eidhammer I. & Martens L. (2009) PRIDE Converter: making proteomics data-sharing easy. *Nature Biotechnology* **27**, 598-599.

Bartels D. & Sunkar R. (2005) Drought and Salt Tolerance in Plants. *Critical Reviews in Plant Sciences* **24**, 23-58.

Bedon F., Majada J., Feito I., Chaumeil P., Dupuy J.-W., Lomenech A.M., Barre A., Gion J.-M. & Plomion C. (2011) Interaction between environmental factors affects the accumulation of root proteins in hydroponically grown *Eucalyptus globulus* (Labill.). *Plant Physiology and Biochemistry* **49**, 69-76.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of Royal Statistical Society B* **57**, 289-300.

Bogeat-Triboulot M.B., Brosché M., Renaut J., et al. (2007) Gradual soil water depletion results in reversible changes of gene expression, protein profiles, ecophysiology, and growth performance in *Populus euphratica*, a poplar growing in arid regions. *Plant Physiology* **143**, 876-892.

Boland D.J., Brooker M.I.H., Chippendale G.M., Hall N., Hyland B. P.M., Johnson R.D., Kleinig D.A., McDonald M.W. & Turner J.D. (2006) *Forest trees of Australia,* Melbourne: CSIRO.

Bonhomme L., Monclus R., Vincent D., Carpin S., Claverol S., Lomenech A.M., Labas V., Plomion C., Brignolas F. & Morabito D. (2009a) Genetic variation and drought response in two *Populus x euramericana* genotypes through 2-DE proteomic analysis of leaves from field and glasshouse cultivated plants. *Phytochemistry* **70**, 988-1002.

Bonhomme L., Monclus R., Vincent D., Carpin S., Lomenech A.M., Plomion C., Brignolas F. & Morabito D. (2009b) Leaf proteome analysis of eight *Populus x euramericana* genotypes: genetic variation in drought response and in water-use efficiency involves photosynthesis-related proteins. *Proteomics* **17**, 4121-4142.

Ceulemans R., Van Praet L. & Jiang X. N. (1995) Effects of CO2 enrichment, leaf position and clone on stomatal index and epidermal cell density in poplar (*Populus*). *New Phytologist* **131**, 99-107.

Chaves M.M., Maroco J.P. & Pereira J.S. (2003) Understanding plant response to drought - from genes to the whole plant. *Functional Plant Biology* **30**, 239-264.

Cornish K., Radin J.W., Turcotte E.L., Lu Z. & Zieger E. (1991) Enhanced photosynthesis and stomatal conductance of Pima cotton (*Gossypium barbadense* L.) bred for increased yield. *Plant Physiology* **97**, 484-489.

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

Delwaulle J.C. (1988) Plantations clonales au Congo. Point des recherches sur le choix des clones dix ans apres les premieres plantations. In: *Proceedings of the IUFRO Meeting on Genetic Improvement of Tropical Trees*, Pattaya, Thailand, p. 4.

Farquhar G.D. & Sharkey T.D. (1982) Stomatal conductance and photosynthesis. *Annual Review of Plant Physiology* **33**, 317-345.

Gion J.M., Lalanne C., Le Provost G., et al. (2005) The proteome of maritime pine wood forming tissue. *Proteomics* **14**, 3731-3751.

Gruber C.W., Cemazar M., Heras B., Martin J.L. & Craik D.J. (2006) Protein disulfide isomerase: the structure of oxidative folding. *Trends in Biochemical Sciences* **31**, 455-464.

Guo X.-Y., Zhang X.-S. & Huang Z.-Y. (2010) Drought tolerance in three hybrid poplar clones submitted to different watering regimes. *Journal of Plant Ecology* **3**, 79-87.

He C., Zhang J., Duan A., Zheng S., Sun H. & Fu L. (2008) Proteins responding to drought and high-temperature stress in *Populus x euramericana* cv. '74/76'. *Trees* **22**, 803-813.

Hernandez-Blanco C., Feng D.X., Hu J., et al., (2007) Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. *The Plant Cell* **19**, 890-903,

Hong J.K. & Hwang B.K. (2002) Induction by pathogen, salt and drought of a basic class II chitinase mRNA and its in situ localization in pepper (*Capsicum annuum*). *Physiologia Plantarum* **114**, 549-558.

Iglesias G.T. & Wistermann D. (2008) *Eucalyptus universalis*. Global cultivated *Eucalyptus* forest map. Eucalyptologics: Information on *Eucalyptus* cultivation worlwide. GIT Forestry Consulting. www.git-forestry.com.

IPCC, Summary for Policymakers. In: S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, H.L. Miller, (Eds.), Climate Change 2007: The physical science basis, Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, Cambridge University Press, New York, 2007.

Karp N.A. & Lilley K.S. (2009) Investigating sample pooling strategies for DIGE experiments to address biological variability. *Proteomics* **9**, 388-397.

Kellison R.C. (2001). Present and future uses of eucalyptus wood in the world. Proceedings of the IUFRO International Symposium Developing the eucalypt of the future. Valdivia, Chile: INFOR Instituto Forestal (published on CD).

Kohzuma K. Cruz J.A., Akashi K., Hoshiyasui S., Munekage Y.N., Yokota Y. & Kramer D.M. (2009). The long-term responses of the photosynthetic proton circuit to drought. *Plant, Cell and Environment* **32**, 209-219.

Laclau J.P., Bouillet J.P. & Ranger J. (2000) Dynamics of biomass and nutrient accumulation in a clonal plantation of *Eucalyptus* in Congo. *Forest Ecology and Management* **128**, 181-196.

Lawlor D.W & Tezara W. (2009) Causes of decreased photosynthetic rate and metabolic capacity in water-deficient leaf cells: a critical evaluation of mechanisms and integration of processes. *Annals of Botany* **103**, 561-579.

Lecoeur J., Wery J., Turc O. & Tardieu F. (1995) Expansion of pea leaves subjected to short water-deficit: cell number and cell-size are sensitive to stress at different periods of leaf development. *Journal of Experimental Botany* **46**, 1093-1101.

Li C. & Wang K. (2003) Differences in drought responses of three contrasting *Eucalyptus microtheca* F. Muell. Populations. *Forest Ecology and Management* **179**, 377-385.

Li C. (1998) Some aspects of leaf water relations in four provenances of Eucalyptus microtheca seedlings. *Forest Ecology and Management* **111**, 303-308.

Li C., Berninger F., Koskela J. & Sonninen E. (2000) Drought responses of *Eucalyptus microtheca* provenances depend on seasonality of rainfall in their place of origin. *Plant Biology* **27**, 231-238.

Liang Z.-S., Yang J.-W., Shao H.-B. & Han R.-L. (2006) Investigation on water consumption characteristics and water use efficiency of poplar under soil water deficits on the Loess Plateau. *Colloids and Surfaces B: Biointerfaces* **53**, 23-28.

Lundgren D.H., Hwang S.I., Wu L. & Han D.K. (2010) Role of spectral counting in quantitative proteomics. *Expert Review of Proteomics* **1**, 39-53.

Martens L., Hermjakob H., Jones P., Adamski M., Taylor C., States D., Gevaert K., Vandekerckhove J. & Apweiler R. (2005) PRIDE: the proteomics identifications database. *Proteomics* **5**, 3537-3545.

Martinez-Meier A., Sanchez L., Pastorino M., Gallo L. & Rozenberg, P. (2008) What is hot in tree rings? The wood density of surviving Douglas-firs to the 2003 drought and heat wave. *Forest Ecology and Management* **256**, 837-843.

Matesanz S., Gianoli E. & Valladares F. (2010) Global change and the evolution of phenotypic plasticity in plants. *Annals of the New York Academy of Sciences* **1206**, 35-55.

Mechin V., Balliau T., Chateau-Joubert S., Davanture M., Langella O., Negroni L., Prioul J.-L., Thevenot C., Zivy M. & Damerval C. (2004) A two-dimensional proteome map of maize endosperm. *Phytochemistry* **65**, 1609-1618.

Mencuccini M. & Comstock J. (1999) Variability in hydraulic architecture and gaz exchange of common bean (*Phaesolus vulgaris*) cultivars under well-watered conditions: interactions with leaf size. Australian Journal of Plant Physiology **26**, 115-124.

Merchant A., Callister A., Arndt S., Tausz M. & Adams M. (2007) Contrasting physiological responses of six *Eucalyptus* species to water deficit. *Annals of Botany* **100**, 1507-1515.

Merchant A., Peuke A.D., Keitel C., Macfarlane C., Warren C.R. & Adams M.A. (2010) Phloem sap and leaf delta13C, carbohydrates, and amino acid concentrations in *Eucalyptus*

globulus change systematically according to flooding and water deficit treatment. *Journal of Experimental Botany* **61**, 1785-1793.

Miller G., Suzuki N., Ciftici-Yilmaz S. & Mittler R. (2010) Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant, Cell and Environment* **33**, 453-467.

Monclus R., Dreyer E., Villar M., Delmotte F.M., Delay D., Petit J.M., Barbaroux C., Le Thiec D., Bréchet C. & Brignolas F. (2007) Impact of drought on productivity and water use efficiency in 29 genotypes of *Populus deltoides x Populus nigra*. *New Phytologist* **169**, 765-777.

Ngugi M.R., Doley D., Hunt M.A., Dart P. & Ryan P. (2003) Leaf water relations of *Eucalyptus cloeziana* and *Eucalyptus argophloia* in response to water deficit. *Tree Physiology* **23**, 335-343.

Novaes E., Drost D.R., Farmerie W.G., Pappas G.J. Jr., Grattapaglia D., Sederoff R.R. & Kirst M. (2008) High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* **9**, 312.

Pinheiro C., Chaves M.M. & Ricardo C.P.P. (2001) Alterations in carbon and nitrogen metabolism induced by water deficit in the stems and leaves of *Lupinus albus* L. *Journal of Experimental Botany* **52**, 1063-1070.

Pita P. & Pardos J.A. (2001) Growth, leaf morphology, water and tissue water relations of *Eucalyptus globulus* clones in response to water deficit. *Tree Physiology* **21**, 599-607.

Pita P., Gascó A. & Pardos J.A. (2003) Xylem cavitation, leaf growth and leaf water potential in *Eucalyptus globulus* clones under well-watered and drought conditions. *Functional Plant Biology* **30**, 891 - 899.

Plomion C., Lalanne C., Claverol S., et al. (2006) Mapping the proteome of poplar and application to the discovery of drought-stress responsive proteins. *Proteomics* **6**, 6509-6527.

Pryor L. D., Williams E. R. & Gunn B. V. (1995) A Morphometric Analysis of *Eucalyptus urophylla* and Related Taxa with Descriptions of Two New Species. *Australian Systematic Botany* **8**, 57-70.

R Development Core Team. (2008) R: A language and environment for statistical computing. *R Foundation for Statistical Computing*; Vienna, AU

Ramagli L.S. & Rodriguez L.V. (1985) Quantitation of micrograms amounts of proteins in two-dimensional polyacrylamide gel electrophoresis. *Electrophoresis* **6**, 559-563.

Safou-Matondo R., Deleporte P., Laclau J.P. & Bouillet J.P. (2005) Hybrid and clonal variability of nutrient content and nutrient use efficiency in Eucalyptus stands in Congo. *Forest Ecology and Management* **210**, 193-204.

Saint-Andre L., M'Bou A.T., Mabiala A., Mouvondy W., Jourdan C., Roupsard O., Deleporte P., Hamel O. & Nouvellon Y. (2005) Age-related equations for above- and below-ground biomass of a *Eucalyptus* hybrid in Congo. *Forest Ecology and Management* **205**, 199-214.

Schulze E.-D., Turner N.C., Nicolle D. & Schumacher J. (2006) Leaf and wood carbon isotope ratios, specific leaf areas and wood growth of *Eucalyptus* species across a rainfall gradient in Australia. *Tree Physiology* **26**, 479-492.

Shao H.B., Chu L.Y., Jalleel C.A. & Zhao C.X. (2008) Water-deficit stress-induced anatomical changes in higher plants. *Comptes Rendus Biologies* **331**, 215-225.

Sparla F., Tedeschi G., Pupillo P. & Trost P. (1999) Cloning and heterologous expression of NAD(P)H:quinone reductase of *Arabidopsis thaliana*, a functional homologue of animal DT-diaphorase. *FEBS Letters* **463**, 382-386.

Specht R.L. (1970) Vegetation. In: Leeper, G.W. (Ed.), The Australian Environment. *CSIRO and Melbourne University Press*, Carlton, Victoria 1970, pp. 44-67.

Timperio A.M., Egidi M.G. & Zolla L. (2008) Proteomics applied on plant abiotic stresses: role of heat shock proteins (HSP). *Journal of Proteomics* **71**, 391-411.

Torabi S., Wissuwa M., Heidari M., Naghavi M.-R., Gilany K., Hajirezaei M.-R., Omidi M., Yazdi-Samadi B., Ismail A.M. & Salekdeh G.H. (2009) A comparative proteome approach to decipher the mechanism of rice adaptation to phosphorous deficiency. *Proteomics* **9**, 159-170.

Trouvé C., Mariotti A., Schwartz D. & Guillet B. (1994) Soil organic carbon dynamics under *Eucalyptus* and *Pinus* planted on savannas in the Congo. *Soil Biology and Biochemistry* **26**, 287-295.

Tuomela K. (1997) Leaf water relations in six provenances of *Eucalyptus microtheca*: a greenhouse experiment. *Forest Ecology and Management* **92**, 1-10.

Tuomela K., Koskela J. & Gibson A. (2001) Relationship between growth, specific leaf area and water use in six populations of *Eucalyptus microtheca* seedlings from two climates grown in controlled conditions. *Australian Forestry* **64**, 75-79.

Turner N.C. (1986) Adaptation to water deficits: a changing perspective. *Australian Journal of Plant Physiology* **13**, 175-190.

Ulitsky I., Maron-Katz A., Shavit S., Sagir D., Linhart C., Elkon R., Tanay A., Sharan R., Shiloh Y. & Shamir R. (2010) Expander: from expression microarrays to networks and functions. *Nature Protocols* **5**, 303–322.

Urade R. (2007) Cellular response to unfolded proteins in the endoplasmic reticulum of plants. *FEBS Journal* **274**, 1152-1171.

Valente M.A.S., Faria J.A.Q.A., Soares-Ramos J.R.L., et al. (2008) The ER luminal binding protein (BiP) mediates an increase in drought tolerance in soybean and delays drought-induced leaf senescence in soybean and tobacco. *Journal of Experimental Botany* **60**, 533-546.

Vincent D., Balesdent M.-H., Gibon J., et al.. (2009) Hunting Down Fungal Secretomes using Liquid-Phase IEF Prior to High Resolution 2-DE. *Electrophoresis* **30**, 4118-4136.

Vincent D., Ergül A., Bohlman M.C. et al. (2007) Proteomic analysis reveals differences between *Vitis vinifera* L. cv. Chardonnay and cv. Cabernet Sauvignon and their responses to water deficit and salinity. *Journal of Experimental Botany* **58**, 1873-1892.

Vincent D., Lapierre C., Pollet B., Cornic .G, Negroni L. & Zivy M. (2005) Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation. *Plant Physiology* **137**, 949-960.

Wang D., Bachelard E.P. & Banks C.G. (1988) Growth and water relations of two subspecies of *Eucalyptus globulus*. *Tree Physiology* **4**, 129-138.

Weretilnyk E.A., Alexander K.J., Drebenstedt M., Snider J.D., Summers P.S. & Moffatt B.A. (2001) Maintaining methylation activities during salt stress. The involvement of adenosine kinase. *Plant Physiology* **125**, 856-865.

Westbrook J., Yan J., Wait R., Welson S. & Dunn M. 2001. Zooming-in on the proteome: very narrow-range inmobilised pH gradients reveal more protein species and isoforms. *Electrophoresis* **22**, 2865-2871.

White D. A., Beadle C. L. & Worledge D. (1996) Leaf water relations of *Eucalyptus globulus ssp. Globulus* and *E. nitens*: seasonal, drought and species effects. *Tree Physiology* **16**, 469-476.

White D.A., Turner N. C. & Galbraith J. H. (2000) Leaf water relations and stomatal behaviour of four allopatric *Eucalyptus* species planted in Mediterranean southwestern Australia. *Tree Physiology* **20**, 1157-1165.

Whitehead D. & Beadle C. L. (2004) Physiological regulation of productivity and water use in *Eucalyptus*: a review. *Forest Ecology and Management* **193**, 113-140.

Wightman R. & Turner S.R. (2008) The roles of the cytoskeleton during cellulose deposition at the secondary cell wall. *Plant Journal* **54**, 794-805.

Wingler A., Lea P.J., Quick W.P. & Leegood R.C. (2000) Photorespiration: metabolic pathways and their role in stress protection. *Philosophical Transactions of the Royal Society B: Biological Sciences* **355**, 1517-1529.

Wu Y. & Cosgrove D.J. (2000) Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. *Journal of Experimental Botany* **51**, 1543-1553.

Xiao X., Yang F., Zhang S., Korpelainen H. & Li C. (2009) Physiological and proteomic responses of two contrasting *Populus cathayana* populations to drought stress. *Physiologia Plantarum* **136**, 150-168.

Yang F., Wang Y. & Miao .L.F. (2010) Comparative physiological and proteomic responses to drought stress in two poplar species originating from different altitudes. *Physiologia Plantarum* **139**, 388-400.

Yang Y., Thannhauser T.W., Li L. & Zhang S. (2007) Development of an integrated approach for evaluation of 2-D gel image analysis: impact of multiple proteins in single spots on comparative proteomics in conventional 2-D gel/MALDI workflow. *Electrophoresis* **12**, 2080-2094.

Zhou S., Sauve R. & Thannhauser T.W. (2009) Proteome changes induced by aluminium stress in tomato roots. *Journal of Experimental Botany* **60**, 1849-1857.

Zhu W., Venable J., Giometti C.S., Khare T., Tollaksen S., Ahrendt A.J. & Yates III J.R. (2005) Large-scale LC-MS/MS for silver- and Coomassie blue-stained polyacrylamide gels. *Electrophoresis* **26**, 4495-4507.

Table 2. Protein identity of the 58 spots harbouring a single protein identification and with significant genotype (G), treatment (T) and interaction (GxT) effects.

									MW (kDa)		pI		ANOVA (p-value)j)			FC (NIR/IR)k)	
Spot n°[a)]	Organ Type[b)]	Assignment	Cat.[c)]	Sequence reference[d)]	# Pept.[e)]	% Cov.[f)]	UniprotKB Entry[g)]	Species	Th.[h)]	Exp.[i)]	Th.[h)]	Exp.[i)]	Genotype (G)	Treatment (T)	GxT	1-41	18-50
GxT effects																	
310	S	*Adenosine kinase 2	NM	CT983393	4	21.8	Q9LZG0	*A. thaliana*	37.8	59.5	5.1	5.1	0.67476	0.19668	0.01584	1.41	**-2.03**
586	S	*Protein disulfideisomerase	PF	KIRST.672.C1	9	22.8	Q43116	*R. communis*	55.6	75.2	4.9	4.7	0.20355	0.47385	0.01461	1.39	-1.55
694	S	*Quinone reductase-like	EM	CT983353	2	24.3	-	*V. vinifera*	-	58.6	-	6.7	0.61339	0.1168	0.01361	1.71	-1.25
Combined effects with GxT																	
113	S	ATP synthase subunit beta cp	EM	KIRST.23124.C1	2	20.7	Q49KZ1	*E. globulus*	53.7	70	5.3	5.3	0.0288	0.00056	0.01461	-1.22	**-3.66**
92	S	ATP synthase subunit beta mt	EM	ES596525	6	35.1	P19023	*Z. mays*	59.1	70.6	6	5.4	0.00247	0	0.00134	**-2.16**	**-4.67**
145	S	*Adenosine kinase 2	NM	ES592020	3	16.5	Q9LZG0	*A. thaliana*	37.8	59.2	5.1	5.2	0.00061	0.00212	0.01895	-1.05	**-1.95**
88	S	*Actin-7	S	KIRST.617.C1	6	14.7	P53492	*A. thaliana*	41.7	64.5	5.3	5.5	0.00072	0	0.00105	**-1.96**	**-4.85**
104	S	*Caffeic acid 3-O-methvltransferase	SM	KIRST.4436.C1	6	36.7	P46484	*E. gunnii*	40	61.2	5.5	5.6	0.00375	0.00002	0.00059	-1.21	**-4.08**
112	bL	*Serine glyoxylate aminotransferase	AM	CU400252	7	44.9	Q56YA5	*A. thaliana*	44.2	63.4	7.7	8.7	0.00001	0.00005	0.00006	-1.02	**-3.28**
45	bL	*Basic endochitinase	SD	KIRST.38305.C1	2	50.8	P51613	*V. vinifera*	33.4	44.1	8.4	10.1	0.00185	0.00021	0.03499	**3.69**	**6.09**
T effects																	
311	aL	*Glutamine synthetase leaf isozyme cp	AM	KIRST.2424.C5	5	15.6	P15102	*P. vulgaris*	47.2	55.8	6.8	5.5	0.13348	0.00028	0.1619	-1.25	-1.66
267	S	S-adenosyl methionine synthetase 2	AM	ES596251	2	13.6	Q96552	*C. roseus*	43	54.2	5.5	4.8	0.98338	0.01271	0.1934	-1.15	**-2.13**
126	bL	*Peroxisomal (S)-2-hydroxy-acid oxidase	AM	KIRST.16228.C1	7	22.8	P05414	*S. oleracea*	40.3	60.4	9.2	10	0.1254	0.02144	0.65324	-1.68	**-2.31**
122	bL	*Peroxisomal (S)-2-hydroxy-acid oxidase	AM	KIRST.16228.C1	9	23.4	P05414	*S. oleracea*	40.3	60.5	9.2	9.9	0.27124	0.039	0.77317	**-1.95**	**-2.48**
438	aL	ATP synthase subunit beta cp	EM	KIRST.737.C1	2	6.4	Q2L917	*G. hirsutum*	53.6	67.3	5.2	5.4	0.4332	0.00005	0.74357	-1.49	-1.38
51	aL	ATP synthase subunit beta cp	EM	KIRST.23124.C1	5	45.3	Q49KZ1	*E. globulus*	53.7	68.3	5.3	5.6	0.11539	0.00012	0.11701	-1.72	**-3.11**
55	aL	ATP synthase subunit alpha cp	EM	KIRST.28079.C1	2	33.3	Q49L13	*E. globulus*	55.5	70.6	5.1	5.4	0.60107	0.00167	0.32322	-1.74	**-2.98**
149	S	Malate dehydrogenase mt	CM	ES590359	4	20.8	P17783	*C. lanatus*	36.2	58.6	8.9	6	0.19234	0.00207	0.91367	**-1.84**	-1.72
180	bL	*Malate dehydrogenase gx	CM	KIRST.7544.C1	7	19	Q2QLQ9	*O. sativa*	37.4	54.4	8.1	9	0.26735	0.03277	0.80743	-1.54	**-1.77**
196	S	Phosphoglycerate kinase cytosolic	CM	CT983242	4	18.1	Q42962	*N. tabacum*	42.4	51.8	5.7	5	0.55062	0.0076	0.30888	-1.42	**-2.55**
122	aL	Ribulose bisphosphate carboxylase/oxygenase activase 1 cp	CM	KIRST.1744.C2	6	9.1	Q7X9A0	*L. tridentata*	52.1	62.7	6.4	5	0.90452	0.00001	0.65025	**-1.99**	**-2.29**
108	aL	Ribulose bisphosphate carboxylase/oxygenase activase 1 cp	CM	KIRST.1744.C2	6	16.7	Q7X9A0	*L. tridentata*	52.1	64.5	6.4	5	0.18899	0.00001	0.1001	-1.67	**-2.37**

Table 2 continued

									MW (kDa)		pI		ANOVA (p-value)j)			FC (NIR/IR)k)	
Spot n°[a)]	Organ Type[b)]	Assignment	Cat.[c)]	Sequence reference[d)]	# Pept.[e)]	% Cov.[f)]	UniprotKB Entry[g)]	Species	Th.[h)]	Exp.[i)]	Th.[h)]	Exp.[i)]	Genotype (G)	Treatment (T)	GxT	1-41	18-50
83	aL	Ribulose bisphosphate carboxylase/oxygenase activase cp	CM	ES589623	5	20.3	Q40281	*M. domestica*	48.1	64.2	8.2	5.1	0.07788	0	0.18118	**-2.08**	**-2.64**
225	aL	Ribulose bisphosphate carboxylase/oxygenase activase cp	CM	ES589623	3	13.9	Q40281	*M. domestica*	48.1	59.6	8.2	5.3	0.07739	0.00001	0.09586	-1.41	**-1.81**
185	aL	UTP glucose 1 phosphate uridylyltransferase	CM	KIRST.13533.C1	5	18.8	Q5F1U8	*S. tuberosum*	51.9	69.2	5.7	6.3	0.07285	0.00954	0.08195	-1.05	**-1.91**
142	aL	Carbonic anhydrase cp	EM	ES589377	11	66.1	P17067	*P. sativum*	35.4	41.2	7	6.8	0.86113	0.00013	0.69547	**-1.83**	**-2.08**
109	S	*Adenosine kinase 2	NM	ES592020	3	14.6	Q9LZG0	*A. thaliana*	37.8	59.9	5.1	5.4	0.17166	0	0.78936	**-2.45**	**-2.36**
137	aL	RuBisCO large subunit binding protein subunit beta cp	PF	KIRST.3840.C2	10	21.2	Q37267	*P. sativum*	63	71.1	5.8	5.2	0.87448	0.00969	0.25673	-1.23	**-2.11**
2632	S	Tubulin alpha-3/alpha-5 chain	S	KIRST.44.C7	5	7.6	P20363	*A. thaliana*	49.6	54.9	4.9	5.7	0.06367	0	0.65629	**-4.85**	**-3.75**
2637	S	Tubulin alpha-3/alpha-5 chain	S	KIRST.44.C7	4	7.1	P20363	*A. thaliana*	49.6	55.4	4.9	5.8	0.13544	0.0001	0.43477	**-3.35**	**-3.82**
171	bL	*Basic endochitinase	SD	CU400566	3	15.7	Q9FRV1	*S. cereale*	33.6	30.5	8.3	10.1	0.11644	0.02144	0.2801	1.32	**2.64**
204	bL	*Basic endochitinase	SD	KIRST.38305.C1	2	33.9	P51613	*V. vinifera*	33.4	27.3	8.4	10	0.1254	0.02144	0.31438	1.27	**2.31**
Combined T and G effects																	
92	aL	*Serine glyoxylate aminotransferase	AM	ES596682	6	27.3	O81248	*A. thaliana*	44.2	44.5	7.7	6.6	0.02272	0.00105	0.39232	**-1.76**	**-2.03**
111	bL	*Peroxisomal (S)-2-hydroxy-acid oxidase	AM	KIRST.16228.C1	3	8.4	P05414	*S. oleracea*	40.3	51.5	9.2	10.1	0.00012	0.00164	0.08037	-1.49	**-1.99**
74	aL	ATP synthase subunit alpha cp	EM	KIRST.28079.C1	4	60	Q49L13	*E. globulus*	55.5	70.7	5.1	5.3	0.0268	0.00001	0.06423	**-1.78**	**-2.75**
213	S	Malate dehydrogenase mt	CM	KIRST.6923.C1	4	16.8	P17783	*C. lanatus*	36.2	54.7	8.9	6	0.04177	0.00122	0.32337	-1.51	**-1.89**
324	aL	*Phosphoribulokinase cp	CM	KIRST.15477.C1	11	31.4	P27774	*M. crystallinum*	44.1	57.6	6	5.8	0.0056	0.00014	0.553	-1.39	-1.45
56	aL	Fructose-bisphosphate aldolase 1 cp	CM	ES596411	4	22.6	Q9SJU4	*A. thaliana*	42.9	56.1	6.2	6.7	0.03722	0.0008	0.3215	**-1.97**	**-2.45**
2308	aL	Ribulose bisphosphate carboxylase large chain	CM	KIRST.1789.C3	2	7.5	Q49KZ0	*E. globulus*	52.7	26.1	6	5	0.0034	0.00084	0.26119	1.11	1.33
297	aL	Carbonic anhydrase cp	EM	ES588641	4	27.1	P17067	*P. sativum*	35.4	26.6	7	5.9	0.00012	0.04368	0.3215	-1.24	-1.07
2264	aL	*Stromal 70 kDa heat shock-related protein cp	PF	KIRST.137.C1	8	22.7	Q02028	*P. sativum*	75.5	77.1	5.2	4.9	0.01411	0.00006	0.82554	**-2.34**	**-1.98**
2260	aL	*Stromal 70 kDa heat shock-related protein cp	PF	KIRST.137.C1	7	22.7	Q02028	*P. sativum*	75.5	77	5.2	5	0.00143	0.00001	0.29364	**-2.4**	**-2.31**
2262	aL	*Stromal 70 kDa heat shock-related protein cp	PF	KIRST.137.C1	8	22.7	Q02028	*P. sativum*	75.5	77	5.2	4.9	0.02165	0.00002	0.68079	**-2.57**	**-2.27**

Table 2 continued

									MW (kDa)		pI		ANOVA (p-value)j)			FC (NIR/IR)k)	
Spot n°[a)]	Organ Type[b)]	Assignment	Cat.[c)]	Sequence reference[d)]	# Pept.[e)]	% Cov.[f)]	UniprotKB Entry[g)]	Species	Th.[h)]	Exp.[i)]	Th.[h)]	Exp.[i)]	Genotype (G)	Treatment (T)	GxT	1-41	18-50
310	aL	RuBisCO large subunit binding protein subunit alpha cp	PF	KIRST.742.C2	10	18.4	P08926	*P. sativum*	62	72.6	5.1	5	0.00188	0.0004	0.88556	-1.41	-1.34
32	aL	RuBisCO large subunit binding protein subunit beta cp	PF	KIRST.3840.C2	22	36.7	P08927	*P. sativum*	63	72.1	5.8	5.4	0.00072	0.00013	0.06111	**-1.89**	**-2.95**
2643	S	Tubulin beta8 chain	S	EW688270	7	29.8	Q9FF94	*A. thaliana*	50.6	52	4.7	5.6	0.04314	0.00918	0.11028	-1.25	**-2.65**
2647	S	Tubulin alpha-3/alpha-5 chain	S	KIRST.44.C7	6	15.6	P20363	*A. thaliana*	49.6	56.4	4.9	5.7	0.01638	0	0.32272	**-3.72**	**-3.37**
88	bL	*Basic endochitinase	SD	KIRST.38305.C1	2	33.9	P51613	*V. vinifera*	33.4	43.6	8.4	10.2	0.02306	0.00067	0.70028	**3.9**	**3.16**
105	bL	*Basic endochitinase	SD	KIRST.38305.C1	2	33.9	P51613	*V. vinifera*	33.4	44.5	8.4	10	0.007	0.00083	0.05454	**1.96**	**4.53**
G effects																	
218	bL	Serine hydroxymethyltransferase mt	AM	KIRST.681.C1	3	14.1	P50433	*S. tuberosum*	57.1	38.6	8.4	9.8	0.0088	0.93938	0.65324	1.32	-1.13
129	aL	*Phosphoribulokinase cp	CM	KIRST.15477.C1	2	4.2	P27774	*M. crystallinum*	44.1	46.3	6	5.7	0.00004	0.21298	0.28157	1.23	-1.03
107	bL	Ribulose bisphosphate carboxylase small chain cp	CM	CU395366	4	22.8	P24007	*P. pyrifolia*	20.5	42.8	9	9.5	0.00008	0.24309	0.96193	-1.45	-1.22
226	bL	Ribulose bisphosphate carboxylase small chain cp	CM	CU395366	5	23.3	P24007	*P. pyrifolia*	20.5	10.6	9	9.5	0.0057	0.11802	0.97837	-1.24	-1.19
202	bL	Carbonic anhydrase cp	CM	ES589377	2	9.5	P17067	*P. sativum*	35.4	9.7	7	8.4	0.00469	0.48644	0.97837	-1.23	-1.13
72	aL	31 kDa ribonucleoprotein cp	PF	ES588930	2	11.5	Q04836	*A. thaliana*	35.8	53.9	4.6	4.4	0.00001	0.34033	0.58022	-1.03	-1.15
127	S	Allene oxide cyclase 4 cp	SD	CT982271	4	29	Q9FX69	*A. thaliana*	27.8	34.8	9.1	5.8	0.00086	0.09036	0.05737	-1.68	1.27
2408	bL	Peroxiredoxin Q cp	SD	KIRST.3093.C1	9	19.6	Q6QPJ6	*P. jackii*	23.4	19.2	9.6	10	0.04543	0.72579	0.77317	-1.25	1.03
502	aL	Cytochrome b6f complex ironsulfur subunit 2 cp	EM	CT982835	3	20.1	Q02585	*N. tabacum*	24.1	24.2	8.1	5.2	0.00052	0.87816	0.38411	-1.07	1.07

SUPPORTING INFORMATION

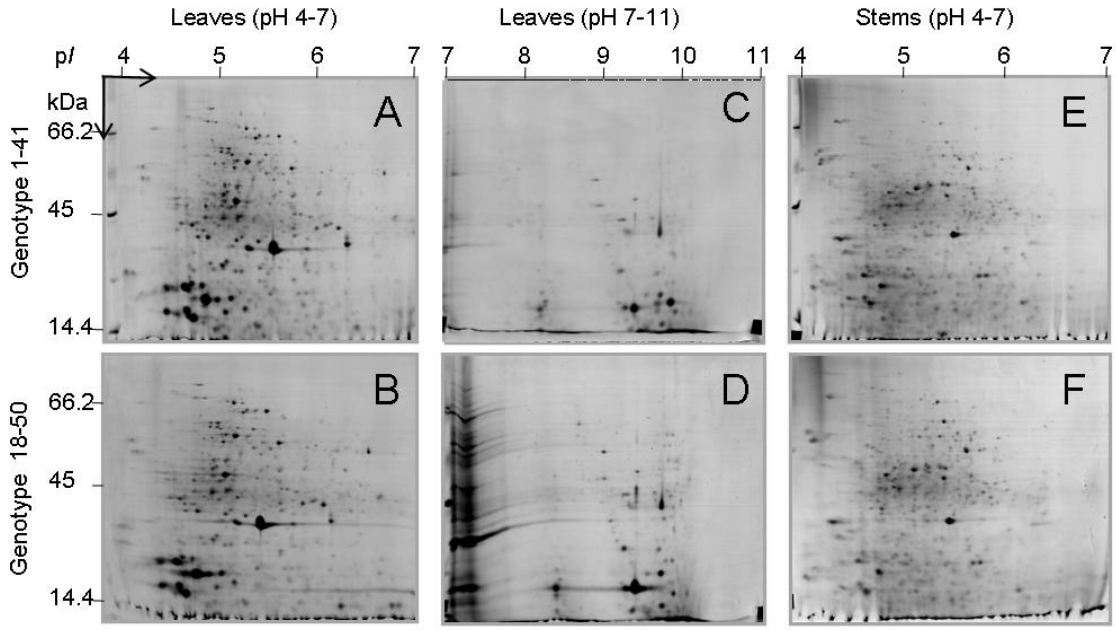


Figure S1. Representative 2-D gels of the soluble proteins from the irrigated Eucalyptus leaves and stems along acidic and basic pH gradients.
The relative molecular weight (MW) is given on the left hand side for the three 2-D gels. The pH gradient is shown above each gel. 2-D gels from genotypes 1-41 and 18-50 in irrigated treatment are presented for leaves proteins separated along acidic, A) and B), and basic pH ranges, C) and D), and for stems proteins along an acidic pH range, E) and F), respectively.

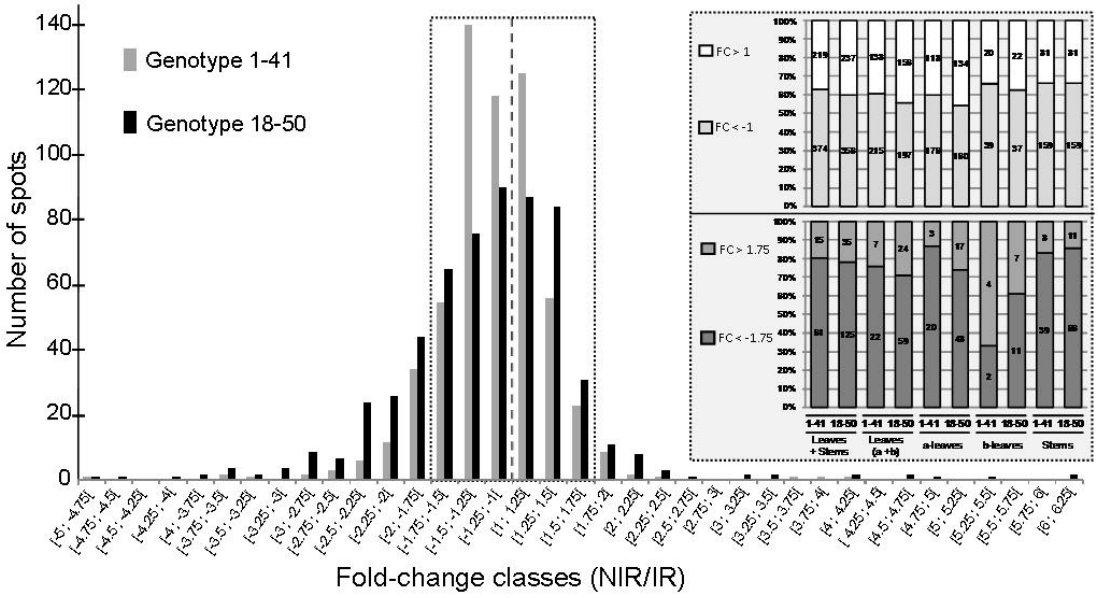


Figure S2. Fold-change distribution of each genotype (1-41 and 18-50) between non irrigated (NIR) and irrigated (IR) treatments based on the normalized volumes of the 593 significant spots from leaves and stems.

Significant spots (*P*adjust < 0.05) from leaves (4-7 and 7-11 pH gradients) and stems (4-7 pH gradient) were used to construct this graph. Negative or positive fold-changes (FC) indicate that the spot volumes respectively decrease or increase in the NIR treatment compared to the IR treatment. FCs were generated from the means of four technical replicates for IR and NIR treatments in leaves and stems. In the inset, upper and below graphs indicate the percentage of differentially regulated protein spots according to the FC (NIR/IR), based on FC occurring at least in one genotype. The corresponding number of protein spots is reported for each genotype in leaves (a-leaves, acidic leaves proteome along 4-7 pH gradient; b-leaves, basic leaves proteome along 7-11 pH gradient) and stems (4-7 pH gradient).

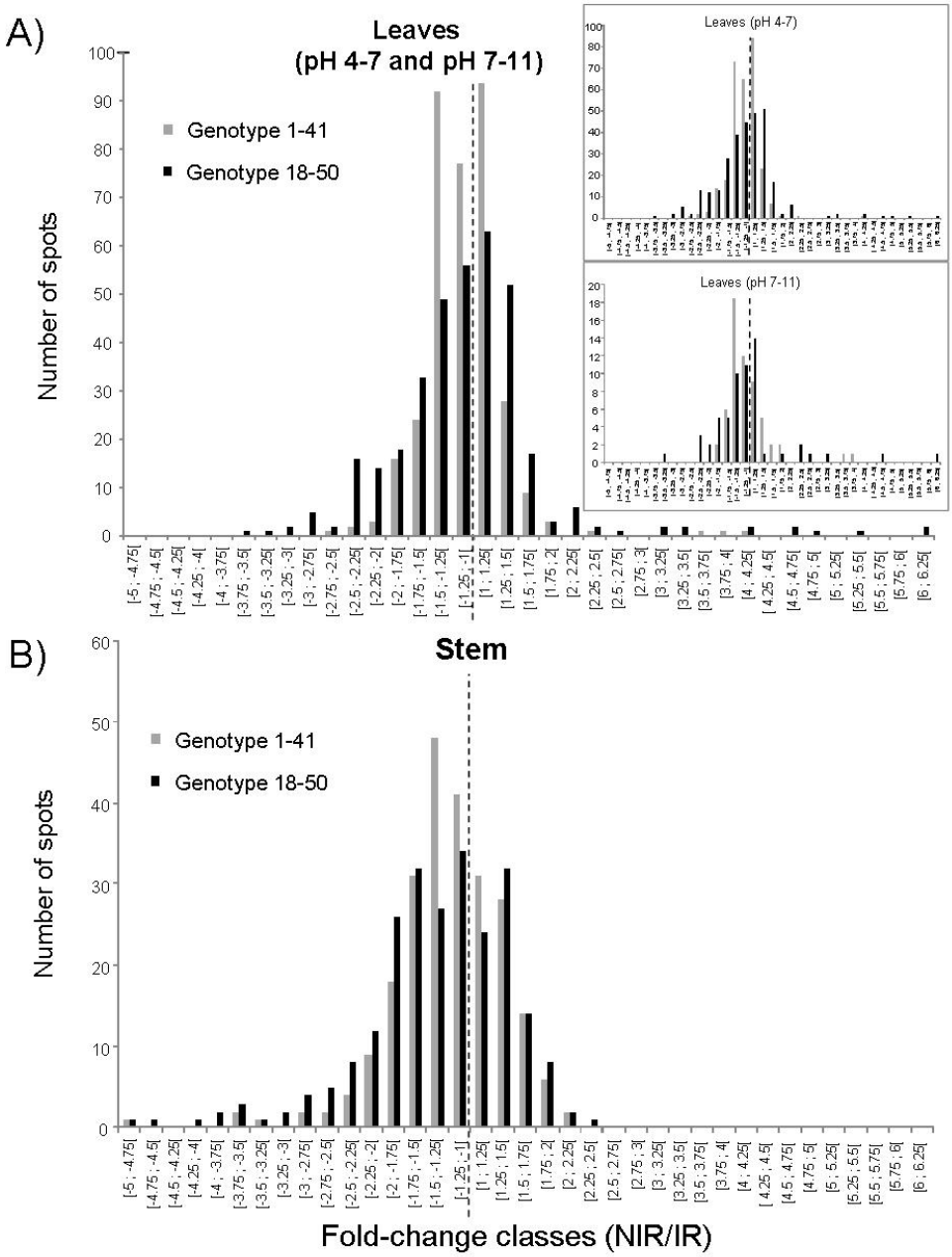


Figure S3. Fold-change distribution of each genotype (1-41 and 18-50) between non irrigated (NIR) and irrigated (IR) treatments based on significant protein spots normalized volumes from leaves (A) and stems (B).
Significant spots (*P*adjust < 0.05) were used to construct this chart, A) 353 spots from leaves including 294 spots separating along 4-7 pH gradient (top inset), and 59 spots separating along 7-11 pH gradient (bottom inset); B) 240 spots from stems (4-7 pH gradient). Negative or positive fold-changes (FC) indicate that the spot volumes respectively decrease or increase in the NIR treatment compared to the IR treatment. FCs were generated from the means of four technical replicates.

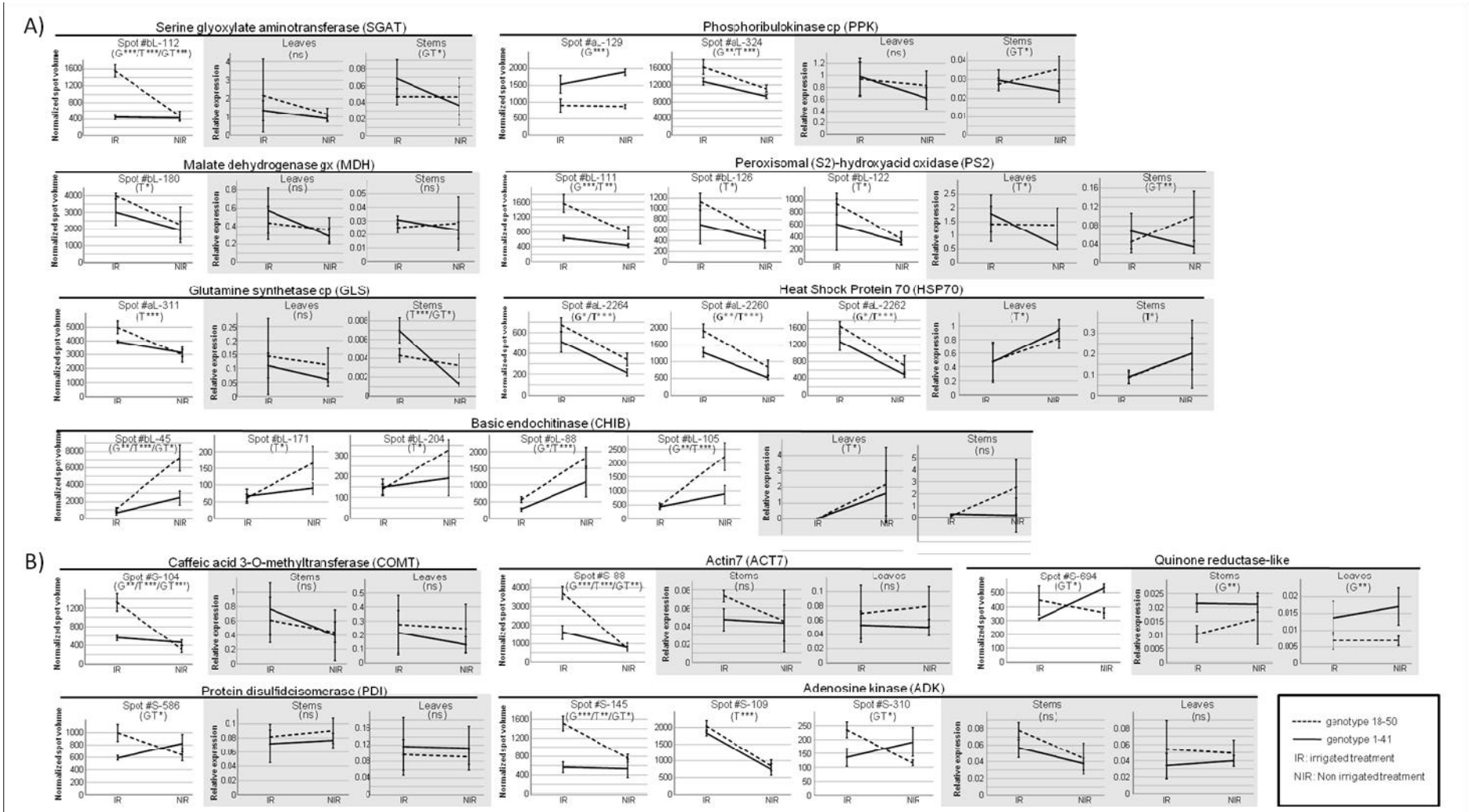


Figure S4. Reaction norm diagrams of transcript and protein abundance in leaves and stems of 12 significant protein spots identified from leaves A) or stems B).
Charts with white background display reaction norms for protein abundance (y axis: normalized spot volume) in leaves A) or in stems B). "Spot aL" and "Spot bL" correspond respectively to acidic and basic protein spots from leaves, while "Spot S" corresponds to acidic proteins from stems. Charts with grey background show transcript abundance, in leaves and/or in stems, of the corresponding encoding gene (y axis: relative expression to EF1α or G3PDH housekeeping genes, see Method section). The result of the two-way ANOVA is indicated into brackets, i.e. genotype (G) and/or treatment (T) and/or interaction effect (GxT), and the associated *p*-value<0.05 (*), 0.01 (**), or 0.001 (***), ns indicates no significant difference, '/' indicate combined ANOVA effects. Irrigated (IR) and not irrigated (NIR) treatments are indicated in the x-axis. No ANOVA effect was found for the housekeeping genes. Standard deviations are presented for spots (4 technical replicates) and for gene transcripts (4 biological replicates).

Table S1. Detailed results of the 158 spots harboring significant two-way ANOVA effects and analysed by LC-MS/MS (Supplied as a separated Excel file).

Colored values in the "two-way ANOVA" section indicate different levels of significance, in blue $0.05 \geq p\text{-value} \geq 0.01$, in pink $0.01 \geq p\text{-value} \geq 0.001$ and in red $0.001 \geq p\text{-value}$. Identical assignment and sequences are indicated in bold. Blue letters show slight amino acid variations between proteins of similar assignment. See caption in *table 1* and material and methods section. PRIDE accession numbers are indicated. Proteins assignments with an asterisk were identified after blast against the non redundant (nr, *viridiplantae*) database of NCBI using tblastx.

Supplementary data for qRT-PCR (S1).

1 Methods

1.1 Eucalyptus DNA sequences amplification

Eucalyptus DNA sequences corresponding to twelve proteins identified by LC-MS/MS were PCR amplified and sequenced for each genotype (1-41 and 18-50). Primers for sequencing were defined based on nucleotide sequences available from Novaes *et al.* (2008). Then, to avoid undesirable nucleotide polymorphisms that would bias the RT-qPCR results (Boyle, Dallaire & MacKay 2009) we designed new PCR primer pairs specific to the studied genotype sequences. Primer pairs and the isolated DNA sequences are listed in Tables A and B, respectively. Amplifications were performed using a Tetrad 2 PTC-0240 Thermo Cycler (MJ Research, Whaltam, MA, USA) in 20 µL final reaction volume composed of 2 µL of 10X PCR reaction Buffer (Invitrogen Carlsbad, CA, USA), 0.8 µL of dNTPs at 5 mM), 1 µL of each primer solution (10 µM), 20 ng of cDNA from leaves or stems, 0.8 units of native Taq polymerase (Invitrogen) and H_2O to complete the final reaction volume. The cycling conditions were: an initial denaturation of 5 min at 94°C, 35 cycles of 30 s at 94°C, 30 s at 58°C, 1 min at 72°C and a final extension of 10 min at 72°C.

Nucleotide sequences of *Elongation factor 1-alpha* and *Glyceraldehyde-3-Phosphate Dehydrogenase* genes were isolated by 3' RACE System for rapid Amplification of cDNA Ends kit (Invitrogen), following the manufacturer's protocol. Each RACE product was cloned independently using the TOPO-TA cloning kit for sequencing (Invitrogen) according to the manufacturer's protocol. PCR and cloned RACE products were sequenced using Big Dye Terminator V 1.1 cycle sequencing kit (Applied Biosystems, Carlsbad, CA, USA) according to the manufacturer's protocol and electrophoresis run were performed using an ABI 3730 DNA Analyzer (Applied Biosystems).

1.2 RNA isolation, first strand synthesis and quantitative RT-PCR

The same samples used for the proteomic study were used for RNA isolation and subsequent analyses. Liquid nitrogen frozen samples were grounded to fine powder using a mill (Retsch GmbH, Haan, Germany) at 30 vibrations per second for 1 min. RNAs were extracted from

each sample following the procedure of Chang *et al.* (1993) in 1.5 ml microtubes. RNA concentration and quality were assessed with the Nanodrop 8000 spectrophotometer (NanoDrop technologies, Wilmington, DE, USA). The extracted RNA samples were treated with RQ1 RNase free DNase (Promega, Madison, WI, USA) according to manufacturer's protocol. First strand cDNA synthesis was performed with 1µg of DNAse treated RNAs by using the ImpromII Reverse Transcription System (Promega) according to the manufacturer's instructions. The resulting cDNA was diluted 1:50 with sterile RNAse-free water and stored at -20ºC. Gene transcript abundance was analysed via real-time PCR on 384-well plates and monitored using Light Cycler 480 (Roche Applied Science, Indianapolis, IN, USA). All reactions were performed in triplicates. Primer pairs were carefully designed using Primer3 software (Rozen & Skaletsky 2000) to anneal on both genotypes and near the 3'-end of each transcript when possible (usually in 3'UTR) in order to ensure primer specificity. Primer pairs used for qRT-PCR are available in Table A. The PCR reaction mixture (10 µL) consisted of 5 µL of 2X ABsolute QPCR SYBR Green Mix (ABgene, Epsom, UK), 0.5 µL of each primer (1.4 µM), 3µL of cDNA sample and 1 µL of RNase-free water. An Hamilton StarLet liquid handling workstation (Hamilton Robotics Reno, NV, USA) was used to assemble cDNA and master mix. The cycling conditions were 95ºC for 15 min., then 40 cycles at 95ºC for 10 sec., 60ºC for 30 sec. and 72°C for 20 sec. Fluorescence reading was performed at each end of the cycle. The readings of melting curves were carried out every 0.2ºC from 60ºC to 97ºC. Standard curves were established for each amplified fragment and the fluorescent signal was computed using LinRegPCR (Ramakers *et al*. 2003; Ruijter *et al*. 2009). Starting concentrations were calculated using the following formula: $N_o = F /(E\text{^}Ct)$, where N_o is the starting concentration for each sample, F is the fluorescence threshold at Ct, E is the PCR efficiency, and Ct is the threshold cycle. Expression levels for transcripts were normalized using one of the two housekeeping genes: *elongation factor 1-alpha (EF1-α)* and *Glyceraldehyde-3-phosphate dehydrogenase* (*G3PDH*), both of which are widely used (Nicot *et al.* 2005; Barsalobres-Cavallar *et al.* 2008). The target gene relative abundance (A_{tg}) is expressed as the ratio between the target gene starting concentration (N_{tg}) and the housekeeping gene starting concentration (N_{hkg}) : $A_{tg} = N_{tg}/N_{hkg}$.

2 Results

Twelve genes encoding proteins with significant ANOVA effects (mainly GxT or T) were selected and their corresponding cDNAs sequenced. Transcript abundance was assessed by

quantitative RT-PCR in leaves and stems. mRNAs were extracted from the same leaf and stem tissues as used for proteomics, but were not pooled. This could explain the higher standard deviation obtained for transcriptomic data, and therefore the fact that less significant effects were detected. Transcript data are presented using reaction norms along with those of proteomic data for comparative purpose (Fig. S4). Significant T effects were obtained for 3 out of 7 leaf proteins: *basic endochitinase* and *heat shock protein 70* were over-expressed in NIR, while *peroxisomal (S2)-hydroxyacid oxidase* was over-expressed in IR. Interestingly, expression of these transcripts showed more significant differences in stems: 6 gene products presenting T or GxT effects. Conversely, out of the 5 stem proteins, only one gene transcript encoding a q*uinone reductase-like protein* (spot #S-694) showed a significant G effect in leaves and stem. In general, there were few similar reaction norms between proteome and transcriptome data. Only one *basic endochitinase* gene indicated a transcript pattern in leaves analogous to that found for its five potential protein isoforms in the same organ. *Heat shock protein 70* transcript level showed a significant T effect in leaves and stems but reaction norms were inversed to those found at the protein level. Lack of correlations between transcript and protein abundances has been observed in several studies (Bogeat-Triboulot *et al*. 2007; Gion *et al.* 2005), which highlights the importance of post-transcriptional regulations in cellular stress adaptation (Gygi *et al.* 1999).

REFERENCES

Barsalobres-Cavallari C.F., Severino F.E., Maluf M.P. & Maia I.G. (2009) Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. *BMC Molecular Biology* **10**, 1.

Boyle B., Dallaire N. & MacKay J. (2009) Evaluation of the impact of single nucleotide polymorphisms and primer mismatches on quantitative PCR. *BMC Biotechnology* **9**, 75.

Chang S., Puryear J. & Cairney J. (1993) A simple and efficient method for isolating RNA from Pine Trees. *Plant Molecular Biology Reporter* **11**, 113-116.

Gygi S.P., Rochon Y., Franza B.R. & Aebersold R. (1999) Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology* **19**, 1720-1730.

Nicot N., Hausman J., Hoffmann L., & Evers D. (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* **56**, 2907-2914.

Novaes E., Drost D.R., Farmerie W.G., Pappas G.J. Jr., Grattapaglia D., Sederoff R.R. & Kirst M. (2008) High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* **9**, 312.

Ramakers C., Ruijter J.M., Deprez R.H. & Moorman A.F. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**, 62.

Rozen S. & Skaletsky H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Molecular Biology* **132**, 365-86.

Ruijter J.M., Ramakers C., Hoogaars W.M.H., Karlen Y., Bakker O., van den Hoff M.J.B. & Moorman A.F.M. (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**, e45.

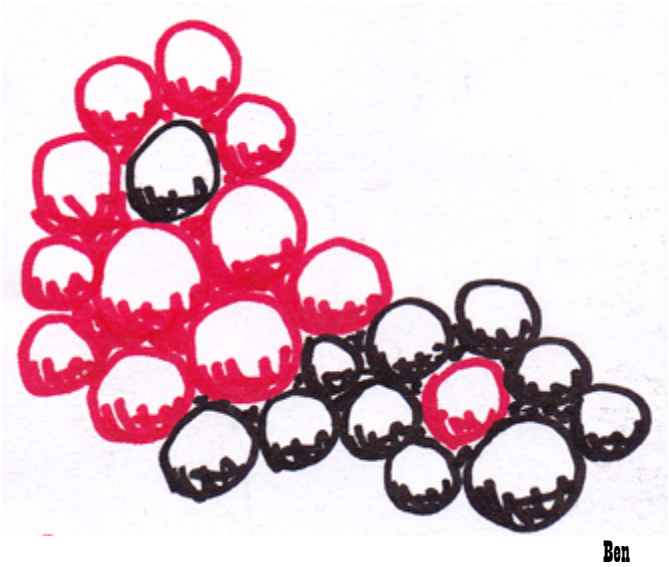
Table A. Primers pairs for PCR amplification and qRT-PCR.

		PCR amplification and sequencing				Quantitative RT-PCR		
Gene name	Identified as unique protein in spot#	Reference sequence	Forward primer (5'-3')	Reverse primer (5'-3')	Amplicon Length (nt)	Forward primer (5'-3')	Reverse primer (5'-3')	Amplicon Length (nt)
Caffeic acid 3-O-methyltransferase	S-104	KIRST.4436.C1	CTTGGTCAAGAACGAGGA	AATGCCTCGAACTCCTTC	672	GTGGATATGCCATGACTGGA	AATGCCTCGAACTCCTTC	219
Actin-7	S-88	KIRST.617.C1	GCATCCAGTCCTTCTAACTG	CAACGAAGGCTGGAAGAG	499	GCATCCAGTCCTTCTAACTG	AGCGAGATCAAGACGGAGAA	244
Adenosine kinase 2	S-310, S-145 and S-109	CT983393 and ES592020	GGATGAAGAAGAATGCAAAA	CCAGCTTTAACACAGTCCTC	673	GGATGAAGAAGAATGCAAAA	GTGCCCAGTTTTCTGGTTGT	177
Stromal 70 kDa heat shock-related protein	aL-2264, aL-2260 and aL-2262	KIRST.137.C1	GCAAGACATCACAATCACG	CAATCCTTGGCTCACTTG	486	GCAAGACATCACAATCACG	TCCACCTTCTCCTTCACTGG	210
Serine glyoxylate aminotransferase	bL-112 and aL-92	CU400252 and ES596682	ATGAATGGGGAGTTGATGT	ATCAAGGGGGTGGTATTTT	628	ATGAATGGGGAGTTGATGT	TGAATGGAAGGGGTGTATGG	199
Basic endochitinase	bL-45, bL-171, bL-204, bL-88 and bL-105	KIRST.38305.C1 and CU400566	CAAGGGCTTCTACACTTACG	CTCAACGAACTGGTTAGGTC	565	CAAGGGCTTCTACACTTACG	AGCAGTATCCCCACGCATAG	182
Glutamine synthetase	aL-311	KIRST.2424.C5	CCATGACAGCAAAAACATGG	TCTCGGCCGAATGACTTATC	609	CCGAGGTGGCAACAATATCT	TCTTGCTCAATCCCATACCA	147
Malate dehydrogenase gx	bL-180	KIRST.7544.C1	CGCATCGCAAGAATCTCC	GCAACAAATGCACATTCGAC	869	ATTTCTGGGACAGCCACAAC	CGATAGCTTTTGGGCATGAT	173
Protein disulfideisomerase	S-586	KIRST.672.C1	GAACCCATACCCCAAGATAACA	TCAGACACGACGATTCCTAATG	506	GCAAATGATGTCCCAAGTGA	CTGCGACTTTCTCCTCTGCT	178
Phosphoribulokinase cp	aL-324 and aL-129	KIRST.15477.C1	TTTGGACATCAGCAATGAGG	GCAGATCCTTCCCTTCTCAA	744	ATGTGGAAAGCCATTTGAGC	AGCCCAACTATGGTCTGGAA	127
Peroxisomal (S)-2-hydroxy-acid oxidase	bL-126, bL-122 and bL-111	KIRST.16228.C1	TGCTTCATATGTTGCTGGTCA	AGTGAAAATGCAGGCAGGAG	595	CCGGCGTAAGAAAGGTACTG	ACATCCCATTCGGTCACAAT	118
Quinone reductase-like	S-694	CT983353	TGTTGAAGGAGTTGAAGCACA	CCTTCTCCATCTCAAACATGC	386	TTGGTATGATGGCTGCTCAA	AGTCAACGCTGTGGTCTCCT	143
*Elongation factor 1-alpha	-	KIRST.18.C7	GTCATCATCATGAACCACCC	GGCCACGCGTCGACTAGTAC	738	TTGAAGAACGGTGATGCTGG	AATCCTGCAATCGACAGCAG	223
*Glyceraldehyde-3-Phosphate Dehydrogenase	-	KIRST.745.C11	ATGAAGGACTGGAGAGGTGG	GGCCACGCGTCGACTAGTAC	720	TGTCAGTTGTGGATCTCACC	AATTCCAGCCTTGGCATCGA	185

*: these sequences were obtained by Rapid Amplification of cDNA Ends (RACE), see Method section.

Table B. DNA sequences obtained by PCR amplification in eucalyptus (Supplied as a separated excel file)

Chapitre VI - Régulation de l’expression des gènes de la voie de biosynthèse des lignines



Régulation de l'expression des gènes de la voie de biosynthèse des lignines

1 Introduction

Dans le chapitre IV sur l'analyse de la plasticité foliaire, au niveau expressionnel, différentes voies métaboliques semblaient discriminer les deux génotypes testés dans leur réponse à la contrainte hydrique. L'une d'elles est une voie métabolique majeure chez les arbres forestiers : la voie de biosynthèse des lignines. De plus, cette voie de biosynthèse pourrait être en relation avec certaines variations phénotypiques observées au cours de cette étude, telles que l'accumulation de phénols, et la diminution du diamètre des vaisseaux du xylème en traitement non-irrigué. Dans cette dernière partie, nous avons donc choisi de nous focaliser plus particulièrement sur l'analyse de l'expression de gènes de la voie de biosynthèse des lignines en étudiant l'expression de plusieurs gènes dans différents organes.

La lignine étant un composant majeur de la paroi secondaire des cellules du xylème, ce biopolymère suscite de vifs intérêts chez les industriels du bois. En effet, dans le secteur de la pâte à papier, les lignines sont des composés indésirables. Intimement liées à la cellulose, elles doivent être éliminées par la mise en place de procédés chimiques polluants coûteux et à faible rendement (Baucher *et al.* 2003). Au contraire, dans le secteur de la production de charbon de bois, de fortes teneurs en lignines sont souvent recherchées car pour une unité de masse égale, la quantité d'énergie produite par combustion de lignines est plus importante que celle produite par combustion de cellulose (Savidge, 2000). Les polymères de lignine sont déposés dans la paroi secondaire de certaines cellules végétales durant la xylogénèse leur conférant hydrophobicité et résistance mécanique. En plus de ce dépôt induit par des mécanismes développementaux, la synthèse des lignines est également induite par des stress biotiques et abiotiques tels que des stress osmotiques, des attaques par des pathogènes et des blessures (Moura *et al.*, 2010). Par exemple, dans le cadre de la contrainte hydrique, Vincent *et al.* (2005) ont montré chez le maïs que la lignification des parois cellulaires de différents tissus avait pour conséquence de diminuer la conductance hydraulique des tissus. Des mesures d'expression de gène ont également

permis de montrer la régulation de cette voie de biosynthèse en cas de stress osmotique, ou suite à l'application d'ABA (Kim *et al.,* 2005 ; Chaumeil, 2006 ; Hu *et al.,* 2009).

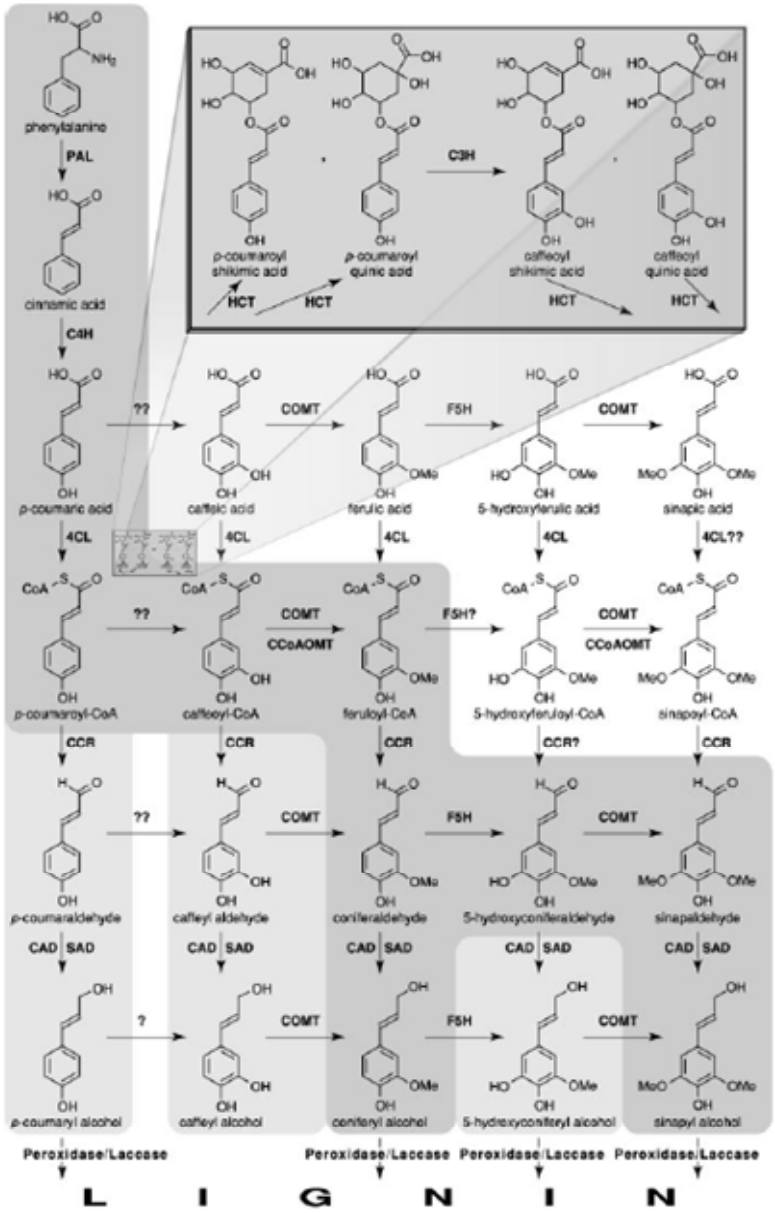


Figure 1 - Voie de biosynthèse des phénylpropanoïdes et des monolignols (d'après Boerjan *et al.*, 2003). La voie la plus foncée est celle qui semble être la plus probable et la plus favorisée chez les angiospermes, tandis que la voie en gris clair peut aussi être « empruntée » chez certaines espèces de plantes. Les parties en blanc ne semblent pas jouer un rôle dans la biosynthèse des lignines.

De part les intérêts économiques qu'elle représente, cette voie de biosynthèse a donc été l'objet de nombreuses études chez les espèces forestières de plantation mais aussi chez les plantes modèles. Les travaux de séquençage du génome menés sur *Arabidospsis thaliana*, *Populus trichocarpa* ou *Oryza sativa* ont notamment permis de montrer que la plupart des gènes de la voie de biosynthèse des lignines appartiennent à des familles multigéniques La voie de biosynthèse des lignines peut être appréhendée comme la succession de deux voies de biosynthèse (figure 1). Tout d'abord, la voie générale des phénylpropanoïdes, qui fait intervenir les enzymes suivantes : la phenylalanine amonia-lyase (PAL), la cinnamate-4-hydroxylase (C4H), la p-coumarate-3-hydroxylase (C3H) et la 4-coumarate :CoA ligase (4CL). A partir de la phénylalanine, cette cascade enzymatique va produire du p-coumaroyl CoA, qui est un intermédiaire métabolique pour de nombreuses voies de biosynthèse, telles que celle des monomères de lignine, des flavonoïdes et d'autres phénols libres ou estérifiés à la paroi. Le p-coumaroyl CoA est notamment le substrat de deux enzymes, l'hydroxycinnamoyl transférase (HCT) et la chalcone synthase (CHS) qui conduisent à la voie de biosynthèse des flavonoïdes. Ensuite, la voie spécifique des monolignols, qui aboutit à la synthèse de trois sous-unités hydroxyphényl (H), guaiacyl (G) et syringyl (S) fait intervenir plusieurs enzymes, telles que la cinnamoyl CoA reductase (CCR) et la cinnamyl alcool déshydrogénase (CAD). D'autres gènes interviennent dans ces deux voies : la ferrulate-5-hydroxylase (F5H), la caffeoyl-CoA O-methyltransferase (CCoAOMT) et la caffeic acid O-methyltransferase (COMT).

avec, pour certains d'entre eux, l'existence de nombreux paralogues plus ou moins bien conservés (Xu *et al.*, 2009). Des études ont pu montrer que l'expression des différents paralogues pouvait varier en fonction des organes (Shi *et al.*, 2010), de stimuli exogènes, biotique ou abiotiques, ou du développement (Lauvergeat *et al.*, 2001). Par exemple, chez CCR, Lauvergeat *et al.* (2001) ont montré que deux membres de la famille CCR étaient exprimées différemment chez Arabidopsis thaliana en fonction du dévleoppement et de l'état infectieux de la plante. Chez l'eucalyptus, des séquences pour chacun des gènes de cette voie de biosynthèse (excepté *C4H* et *HCT*) ont été publiées, et des gènes de régulation de cette voie ont également été mis en évidence (Goicoechea *et al.* 2005; Legay *et al.*, 2007; Foucart *et al.*, 2009). L'effet de la variabilité de ces gènes sur les teneurs en lignines a aussi été montré chez l'eucalyptus dans différentes espèces par des analyses QTL (Gion, 2000 ; Thumma *et al.*, 2010) ou encore par test d'association en population non apparentée (Thumma *et al.*, 2005; Mandrou, 2010).

De plus, au cours du chapitre 4, nous avons identifié une accumulation de phénols dans les feuilles au cours de la contrainte hydrique, présentant un effet de l'interaction GxT. Au niveau anatomique, il a été observé que le diamètre des vaisseaux du xylème variait en fonction du traitement. Il est possible que les gènes de la voie de biosynthèse des lignines interviennent dans de tels processus. De plus, l'analyse en haut-débit du transcriptome d'apex a montré que cette voie de biosynthèse tendait à être sur-exprimé en condition IR, particulièrement pour un génotype.

Etant donné les connaissances déjà acquises sur cette voie de biosynthèse, son implication à la fois dans les processus de développement, mais aussi dans la réponse à des contraintes abiotiques, et l'observation de variation de caractères phénotypiques potentiellement liés, nous avons décidé d'étudier plus particulièrement cette voie de biosynthèse, dans le cadre du dispositif expérimental de cette thèse. Nous avons tout d'abord exploité les banques de séquences d'ADNc d'apex pour rechercher les différents membres de 10 familles multigéniques et identifier leur variation d'abondance. Puis l'analyse de la plasticité moléculaire a été réalisées pour 9 familles multigéniques, dont 2 étaient représentées par 2 membres, dans différents tissus (feuilles, tiges et xylème), au cours d'une cinétique de 5 points de prélèvement.

2 Matériel et méthode

2.1 Cadre expérimental

L'expérimentation et les prélèvements réalisés sont détaillés dans le chapitre 2 (§ Moyens mis en œuvre) et le chapitre 3. Les génotypes utilisés sont le 1-41 (moins productif) et le 18-50 (plus productif), soumis à deux traitements : irrigués (IR) ou non-irrigués (NIR) au cours des deux saisons sèches étudiées.

Cette partie présente les résultats des mesures d'abondance des transcrits des gènes liés à la voie de biosynthèse des phényl-propanoïdes réalisées *in silico* et par RT-qPCR. Les mesures d'abondance *in silico* ont été effectuées à partir du séquençage d'ADNc issus de parties apicales prélevées à la fin de la seconde saison sèche, dont le protocole est détaillé dans le chapitre 4. Les mesures d'expression par RT-qPCR ont été effectuées sur différents tissus prélevés tout au long du suivi de l'expérimentation, au cours de 5 points de prélèvements (P2 à P6; figure 2): P2 effectués en Septembre 2007 (saison sèche 1), P3 en

Décembre 2007 (saison des pluies 1), P4 en Avril 2008 (saison des pluies 1), P5 en Septembre 2008 (saison sèche 2) et P6 en Janvier 2009 (saison des pluies 2). Le point de prélèvement P1 a été abandonné des analyses car trop peu de matériel végétal a pu être prélevé à ce stade de développement des plants (1 mois après plantation).

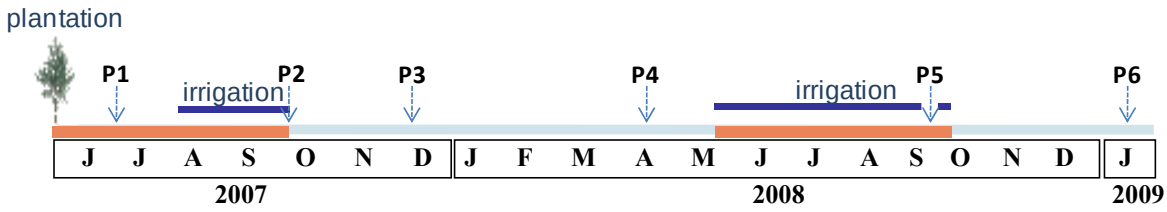


Figure 2 - Déroulement de l'expérimentation de terrain au cours des 20 mois de suivi. La 1ère année, le traitement irrigué (barres bleu foncé) n'a été appliqué qu'une partie de la saison sèche (barres rouge), tandis qu'elle a été apporté tout au long de la saison sèche en 2ème année. Les points de prélèvement (P) sont indiqués par des flèches.

2.2 Analyse d'abondance *in silico*

Tout d'abord, les homologies avec les 63 séquences protéiques codant pour des enzymes de la voie de biosynthèse des lignines, issues de l'article de Xu *et al.*, 2009[1] ont été cherchées dans la banque de séquences spécifiques au matériel végétal de l'étude, issues du séquençage 454 d'apex, par tblastn (BLAST de séquences protéiques sur des séquences nucléotidiques). Les seuils des e-value étaient les suivants : 10^{-30} pour les gènes *4CL*, *C3H*, *C4H*, *CAD*, *CCoAOMT*, *F5H*, *HCT* et *PAL* ; 10^{-24} pour *CCR* et 10^{-8} pour *COMT*, de façon à pouvoir comparer les résultats obtenus à ceux de l'étude de Xu *et al.* (2009). Les domaines protéiques correspondant à chaque gène ont ensuite été recherchés en utilisant le logiciel InterproScan (Hunter *et al.*, 2009) parmi les séquences homologues aux gènes de la voie de biosynthèse des lignines identifiés précédemment. Par la suite, seuls les contigs (contenant plus de deux lectures) qui ont une séquence homologue aux gènes de la voie de biosynthèse des lignines chez *Arabidopsis* et qui portent le domaine conservé ont donc été considérés.

Pour chaque gène, les différentes séquences de contigs traduites en séquences protéiques ont été alignées avec l'algorithme ClustalW implémenté dans le logiciel BioEdit (URL:

[1] Cet article présente la recherche et la comparaison de gènes homologues à 10 gènes de la voie de biosynthèse des lignines parmi les génomes de 14 espèces de plantes et une espèce fongique. Il expose l'expansion possible de ces différentes familles au cours de l'évolution des plantes.

http://www.mbio.ncsu.edu/BioEdit/) afin d'identifier i) les variants issus d'épissage alternatif, ii) les gènes homologues, iii) les variants allèliques s'il y a lieu. Sur la base des deux membres connus *CCoAOMT1* (Accession ACY66930.1) et *CCoAOMT2* (Accession ACY66929.1) chez *Eucalyptus camaldulensis*, qui comportent 8 substitutions d'acide aminés sur une longueur de 217 acides aminés, nous avons défini le seuil d'un changement tous les 20 acides aminés pour définir différents membres d'une même famille de gènes. En dessous de ce seuil, si les alignements protéiques présentaient des motifs d'insertion ou de déletion de plus de 10 acides aminés, les 2 contigs ont été considéré soit comme des variants d'épissage alternatif, soit comme différents allèles mal assemblés, après vérification par alignement des séquences nucléotidiques.

De la même façon que pour l'analyse des séquences 454, la fréquence de lecture a servi à estimer l'abondance des transcrits. En bref, le nombre de lectures par sous-librairies (12 au total) a été normalisé par le nombre de séquences total issues de la librairie, et par la longueur du contig. Ensuite, deux tests statistiques ont été appliqués. Un test non paramétrique implémenté dans le logiciel DEG-Seq (Wang *et al.*, 2009), le taux de faux positifs étant corrigé, a permis de mettre en évidence les contigs différentiellement exprimés entre conditions deux à deux, et d'estimer le ratio d'abondance entre les deux conditions à comparer. Ensuite, une analyse de variance à deux facteurs a permis d'évaluer les effets du génotype (G), du traitement (T) et de leur interaction (GxT).

2.3 Analyse d'expression par RT-qPCR

2.3.1 Extraction des ARN

Pour chaque point de prélèvement, des fragments de tige, des apex caulinaires (feuilles et tiges juvéniles, primordia foliaires, méristèmes), des feuilles, et du xylème en différenciation ont été prélevés et immédiatement plongés dans de l'azote liquide (figure 3). Ces prélèvements ont ensuite été acheminés congelés à l'INRA de Pierroton pour les analyses moléculaires à l'aide de cryoconservateurs (Voyageur +, Air Liquide, Puteaux, France). Les tissus ont été broyés à l'aide d'un broyeur à bille (Retsch GmbH, Haan, Germany) et leurs ARN totaux ont été extraits en suivant le protocole développé par Reid *et al.* (2006). Ils ont ensuite été traités à la RQ1 RNase free DNase (Promega, Madison, WI, USA), dosés au Nanodrop (NaNoDrop technologies, Wilmington, DE, USA), et leur qualité a été vérifiée sur gel d'agarose. Des ADN complémentaires (ADNc) ont été synthétisés par transcription inverse à partir d'1 µg d'ARN total, en utilisant le kit

ImpromII Reverse Transcription System (Promega, Madison, WI, USA). Les ADNc ont ensuite été dilués à 1/50ème avec de l'eau stérile pour l'analyse quantitative.

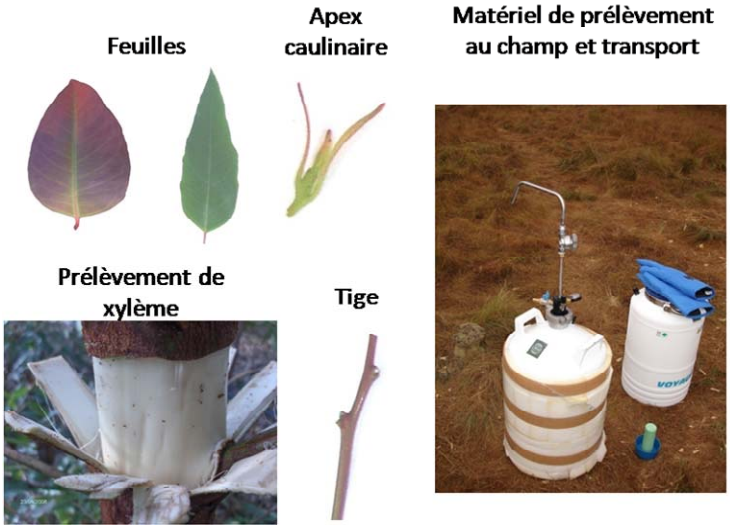


Figure 3 - Illustration des organes et tissus prélevés pour les analyses effectuées par RT-qPCR, ainsi que du matériel de congélation et de transport des échantillons du champ au laboratoire.

2.3.2 Sélection de gènes et construction d'amorces oligonucléotidiques

Cette partie a été réalisée avant l'obtention des séquences par pyroséquençage. Nous ne disposions alors pas de séquences propres au matériel végétal étudié. Etant donné l'importance d'une bonne hybridation des amorces pour la quantification par RT-qPCR (figure 4), il a été décidé de séquencer les gènes d'intérêt et les gènes de ménage (EF1α, G3PDH) afin de dessiner des amorces spécifiques.

Pour cela, les extrémités 3' des ADNc obtenus ont été amplifiées par RACE 3' PCR (Invitrogen, Carlsbad, CA, USA) à l'aide d'amorces définies à partir d'EST ou de séquences d'eucalyptus disponibles dans les bases de données publiques. Les produits PCR ont ensuite été clonés avec le kit TOPO-TA Cloning kit (Invitrogen, Carlsbad, CA, USA) et séquencés sur séquenceur ABI 3730. Ainsi les séquences des extrémités 3' des ADNc ont pu être alignées et comparées grâce au logiciel BioEdit.

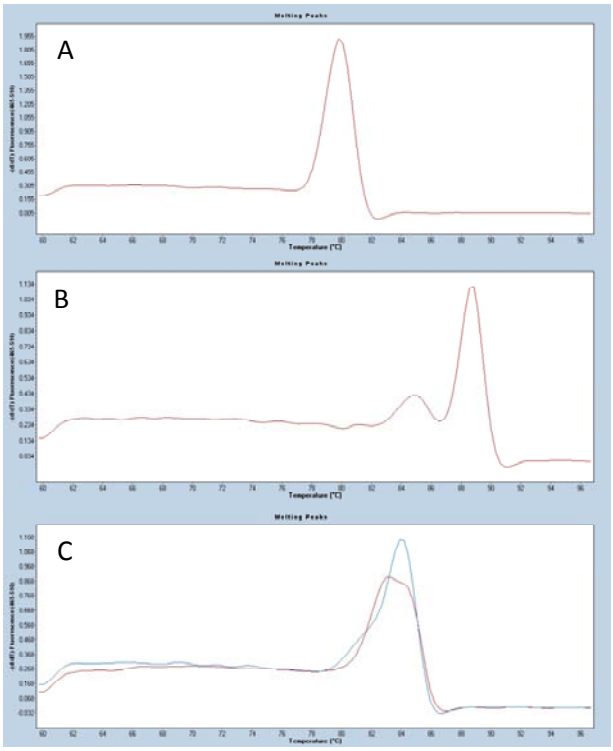


Figure 4: Les courbes de fusion de l'amplicon permettent de déterminer la qualité de l'amplification. La courbe A ne présente qu'un pic, les amorces utilisées permettent donc d'amplifier correctement le fragment choisi. Par contre, la courbe B présente deux pics, le premier suggérant une hybridation entre les amorces qui biaisera la quantification par estimation du cycle d'amplification seuil. Enfin, dans le cas de la courbe C, deux fragments semblent être amplifiés en même temps, les amorces n'étant pas assez spécifiques, ce qui rend impossible la quantification de transcrit.

Pour *EF1α*, *G3PDH*, *CCoAOMT* et *PAL* plusieurs membres ont pu être identifiés. Pour *CCoAOMT* et *PAL*, deux paires d'amorces ont été construites de façon à amplifier spécifiquement ces deux membres séparément. Pour *EF1-α* et *G3PDH*, un seul membre a été choisi, et les amorces ont été définies dans la région 3'UTR non codante qui permet de discriminer les différents membres d'une même famille de gènes. Les paires d'amorces ont été conçues en utilisant le logiciel Primer3 (Rozen & Skaletsky 2000) et sont décrites dans le tableau 1.

Tableau 1 - Gènes, séquences de référence chez *Eucalyptus* et amorces utilisées pour l'amplification par RT-qPCR.

Enzyme	Sequence de référence	Amorce forward	Amorce reverse
4CL	AB368720	CGATGAAGGATGAGGTTGC	AACACGACCTGCTTCGAGAT
C3H	AJ438348	ACGACCTCAGCGAAGATACC	CTCCTCTTGGGCCTTCTGTT
COMT2	AF168777	ACCCAGAACCCCGAGGCAC	TGACCAAGAACTTGCACACC
PAL1	AF167487	AAGCCAAAGCAGGACAGGTA	GCGAACATGAGCTTCCCTAT
PAL2		GGACTTCGCCTCAGTGGTTA	TTGTCCATCGAGACACCAAC
F5H	AJ249093	ACGATCTGCTGGCCTTCTAC	CCATGGCCCACTCGATAG
CCoAOMT1	AF046122	GTATGTGAGGTACTACAGGG	CTCTCTTTTTCGTGGGTTCT
CCoAOMT2	AF168780	GCAAGATGAGAAGAACCATG	AGCACGAAGTCCCGGTAGTA
CAD2	X65631	TATCCCTGTGGTTCACCCTC	GCTCCCTATGAAACTCCCAG
CCR	X79566	GACCAAGTGCTCTGATGAGG	CTCAGGGTTGGTATCTCCAG
EF1α	Non publiées	GTAAGGAGCTTGAGAAGGAG	ACACTCTTGATGACACCGAC
G3PDH	Non publiées	TGTCAGTTGTGGATCTCACC	AATTCCAGCCTTGGCATCGA

2.3.3 Quantification de l'expression

Le mélange réactionnel utilisé pour la PCR contenait 5 µL de 2X ABsolute QPCR SYBR Green Mix (ABgene, Epsom, UK), 0.5 µL de chaque amorce (1.4 µM), 3µL d'ADNc et 1 µL d'eau stérile par puit. Les analyses ont été effectuées en plaque 384, et l'assemblage entre mélange réactionnel de PCR et ADNc a été effectué grâce à une station de travail Hamilton (Hamilton Robotics Reno, NV, USA). L'analyse d'expression a été réalisée à l'aide d'un thermocycleur à lecture de fluorescence Light Cycler 480 (Roche Applied Science, Indianapolis, IN, USA). Sur chaque plaque, les ADNc étaient répliqués 4 fois, pour mesurer l'expression de 3 gènes cible, et celle d'un gène de ménage, qui a permis de normaliser les expressions mesurées d'une plaque à l'autre. Les conditions d'amplification étaient les suivantes : dénaturation initiale à 95°C pendant 15 min, ensuite 40 cycles comprenant une dénaturation à 95°C pendant 10s, une hybridation à 60°C pendant 30s, et une élongation à 72°C pendant 20s. A la fin de chaque cycle, la fluorescence était lue et enregistrée. Enfin, une courbe de fusion a été réalisée de 60°C à 97°C par lecture de la fluorescence des amplicons tout les 0.2°C, pour vérifier la présence d'un seul produit de PCR. Les réactions ont été effectuées en duplicat, et en triplicat si les résultats des deux 1ères mesures n'étaient pas cohérents.

Les courbes d'amplification ont été traitées avec le logiciel LinRegPCR (Ramakers *et al.* 2003; Ruijter *et al.* 2009) afin de soustraire les lignes de base, calculer les efficacités individuelles (Eff_{ind}), et déterminer le cycle seuil (Ct) et le seuil de fluorescence à Ct (F). Ensuite, les concentrations à l'origine ont été calculées à partir de la formule suivante : $N_0=F/(Eff^{Ct})$. Comme les niveaux d'expression dans les différents tissus étaient très différents, les normalisations ont été effectuées tissu par tissu, à l'aide de deux gènes de ménage : *elongation factor 1-alpha* (*EF1-α*) et *Glyceraldehyde-3-phosphate dehydrogenase* (*G3PDH*), dont l'expression est supposée être stable quelque soit le traitement ou le génotype (Nicot *et al.*, 2005). La normalisation a été effectuée par le calcul d'un facteur de normalisation qui est la moyenne géométrique des niveaux d'expression mesurés pour chaque gène de ménage. Afin d'obtenir les niveaux d'expression normalisés, le niveau d'expression du gène cible a été divisé par le facteur de normalisation de l'échantillon considéré (Vandesompele *et al.*, 2002).

Les niveaux d'expression entre chaque réplicat ont ensuite été moyennés, afin d'effectuer les ANOVAs sur les 4 répétitions biologiques utilisées par génotype et traitement, prélèvement par prélèvement, selon le modèle suivant :

$N = \mu + G + T + GxT$, ou N est le niveau d'expression normalisé, μ est la moyenne de l'expression normalisée, G est l'effet du génotype et GxT est l'effet de l'interaction G et T.

3 Résultats

3.1 Identification des différents membres de familles multigéniques

La recherche dans les banques de séquences 454 de contigs homologues à 63 protéines, identifiées chez Arabidopsis par Xu *et al.* (2009) et associées à 10 enzymes de la voie de biosynthèse des lignines, a permis de mettre en évidence 268 séquences homologues (colonne I du tableau 2). Parmi celles-ci, 115 contigs présentaient le domaine associé aux familles recherchées (colonne II).

Malgré l'identification de 46 séquences homologues (colonne I) aux 2 membres de la famille *F5H* chez *Arabidopsis*, il n'a pas été possible de trouver de séquences exprimées dans les banques d'apex correspondant à cette famille et portant le domaine conservé (colonne II). Les familles *4CL*, *CAD*, *CCoAOMT*, *CCR*, *HCT* et *PAL* présentaient un nombre similaire de séquences homologues (colonne I) et de contigs portant le domaine protéique (colonne II), contrairement aux familles *C3H*, *C4H*, *COMT* et *F5H*. Ceci indique que le domaine cherché pour ce premier lot de protéines reste très conservé, alors que des duplications puis spécialisation à de nouvelles fonctions enzymatiques au cours de l'évolution ont pu avoir lieu pour le second. Xu *et al.* (2009) ont observé les mêmes tendances de conservation du domaine pour *CCoAOMT*, *CCR* et *PAL* au cours de la recherche dans les séquences génomiques de 15 espèces de plantes. De même, parmi ces 15 espèces, le domaine des familles *C3H*, *C4H*, *COMT* et *F5H* ne semble pas très conservé. Seules les familles *4CL*, *CAD* et *HCT* présentent des différences par rapport aux résultats obtenus par Xu *et al.* (2009) : elles semblent porter un domaine très conservé chez *Eucalyptus*, alors que l'inverse est observé parmi les 15 espèces étudiées.

Tableau 2 - Résultats de la recherche de 10 gènes de la voie de biosynthèse des lignines dans les banques de séquences 454 d'apex d'eucalyptus. Pour chaque enzyme, sont indiqués, les e-value seuil pour la recherche d'homologie de séquence, ainsi que le domaine recherché. Le nombre de séquences homologues dans la banque (colonne I),le nombre de contigs portant le domaine protéique d'interêt (colonne II), ainsi que le nombre de contig restant après élimination des variants allèliques et issus d'épissage alternatif (colonne III) sont comparées avec les données publiées par Xu et al. (2009) pour 15 espèces, pour Arabidopsis et pour Populus (colonnes indiquées par *).

Enzymes	**e-value**	**Domaine recherché**	**Sequences Homologues *Eucalyptus* (I)**	**Contigs Homologues *Eucalyptus* (II)**	**Genes Homologues *Eucalyptus* (III)**	**Sequences Homologues sur 15 espèces***	**Genes Homologues sur 15 espèces***	**Genes Homologues *Arabidopsis* ***	**Genes Homologues *Populus* ***
4CL	1E-30	PTHR11968:SF43	13	11	10	2212	120	13	22
C3H	1E-30	PTHR19383:SF44	49	3	1	993	14	3	4
C4H	1E-30	PTHR19383:SF33	8	1	1	530	18	1	3
CAD	1E-30	PTHR11695:SF38	27	27	9	1022	110	9	21
CCoAOMT	1E-30	PTHR10509	20	20	10	61	55	7	7
CCR	1E-24	PTHR10366:SF9	25	25	14	207	207	7	40
COMT	1E-08	PIRSF005739	60	9	7	187	187	16	35
F5H	1E-30	PTHR19383:SF46	46	0	0	990	15	2	4
HCT	1E-30	PF02458	14	14	7	279	37	1	7
PAL	1E-30	PF00221/TIGR01226	6	5	4	55	55	4	6
Total		**11**	**268**	**115**	**63**	**6536**	**818**	**63**	**149**

Etant donné que des contigs différents peuvent refléter plusieurs variants d'épissage alternatif ou éventuellement différents allèles d'un même gène, considérer chaque contig comme un membre différent aurait pour conséquences de retrouver un plus grand nombre de membres transcrits qu'il n'en existe au niveau du génome. Pour en savoir plus, nous avons aligné, pour chacune des 10 familles de gènes étudiées, les séquences protéiques et nucléotidiques, pour discriminer les différents membres de chaque famille multi-géniques des variants d'épissage alternatif ou de différentes formes allèliques d'un même gène mal assemblées. Le regroupement des variants d'épissage alternatif et des formes allèliques a finalement permis d'identifier 63 contigs uniques (colonne III), chacun devant correspondre à un locus au niveau du génome. Ainsi les gènes *CAD*, *CCoAOMT*, *CCR* et *HCT* ont présenté une grande proportion de contigs assimilés à des allèles du même gène mal assemblés (plus de 50% de « faux contigs »). Très souvent ces « faux contigs » étaient constitués d'un très faible nombre de lectures, inférieur à 10, mais étaient présents à la fois dans les banques du génotype 1-41 et celles du 18-50, indiquant un problème au niveau des seuils d'assemblage plutôt qu'une trop grande variabilité entre les deux hybrides utilisés.

Certaines familles, telles que *C3H*, *C4H* et *PAL*, présentaient moins de 4 membres porteurs du domaine spécifique recherché (colonne III). Ce sont également des familles comportant peu de membres chez Arabidopsis (moins de 4 membres différents) et chez Populus (moins de 6 membres). A l'inverse, les familles *4CL*, *CAD*, *CCoAOMT* et *CCR* présentaient plus de 9 membres portant le domaine. Ces dernières pourraient inclure des membres ayant été dupliqués un grand nombre de fois dans le génome. A l'exception de *CCoAOMT*, il a été trouvé un grand nombre de membre pour ces mêmes familles chez *Arabidopsis* (plus de 7 membres) et *Populus* (plus de 20 membres).

Il a été trouvé une plus grande proportion de membres appartenant aux familles *CCoAOMT* dans les banques de séquences d'eucalyptus, que dans les génomes de peuplier ou d'*Arabidopsis*, et inversement, très peu de membres correspondant à la *COMT*. Pour *COMT*, ces différences peuvent s'expliquer par le fait que les génomes entiers d'*Arabidopsis* et de *Populus* ont été étudiés dans l'étude de Xu *et al.* (2009), alors que pour eucalyptus, on observe seulement les séquences exprimées dans les parties apicales des tiges.

3.2 Abondance des transcrits dans les banques d'ADNc

L'analyse des différences d'abondance de transcrit dans les apex caulinaires a permis de mettre en évidence 13 contigs différentiellement exprimés : 12 sont différentiellement exprimés entre les deux génotypes, 2 entre les deux traitements et 3 présentent un effet d'interaction, c'est-à-dire que les normes de réaction d'un traitement à l'autre sont différentes entre les génotypes (tableau 3).

Concernant les effets génotype, 4 contigs (codant *HCT* et *COMT*) sont surexprimés chez le 18-50, dont un n'est pas identifié dans les banques de séquences du génotype 1-41. A l'inverse, 8 contigs sont surexprimés chez le 1-41. Ils appartiennent aux familles *CAD*, *CCoAOMT*, *COMT* et *HCT*. Deux contigs codant *CAD* et *COMT* sont surexprimés dans le traitement irrigué. Les contigs présentant un effet GxT, codant *CCoAOMT* et *HCT* présentent le même patron d'expression : surexprimé en IR pour le 18-50, alors que les différences d'abondance entre traitements ne sont pas significatives pour le génotype 1-41.

Les contigs présentant des effets significatifs ont une profondeur moyenne de 211 lectures, alors que la profondeur moyenne parmi les contigs étudiés est de 62 lectures par contig, ce qui permet d'observer l'importance de la profondeur de séquençage pour la mise en évidence de différentiels d'expression.

Malheureusement, un seul contig (codant *CCoAOMT2*) de cette étude a été utilisé pour la validation par RT-qPCR des profils d'abondance prédits par RNA-Seq, et son profil a d'ailleurs été validé. En effet, les gènes utilisés pour cette validation ont été choisis dans différentes catégories fonctionnelles pour être représentatifs des résultats de l'analyse RNA-Seq.

3.3 Expression de certains membres dans différents tissus

La comparaison des différents tissus (feuilles, tiges, xylème) montre que l'expression de tous les gènes est nettement supérieure dans le xylème, tissu en différenciation à l'origine du bois. Les gènes *CAD*2 et *CCR* (gènes de la voie spécifique des monolignols) sont plus exprimés dans les feuilles que dans les tiges, alors que les gènes *CCoAOMT*, *COMT*, *F5H* et *C3H* (gènes de la voie commune des phénylpropanoïdes) sont plus exprimés dans les tiges que dans les feuilles.

D'autre part, il n'a pas été possible de mettre en évidence un gène appartenant à la famille *F5H* dans les banques de séquences d'apex. Ce résultat est cohérent avec le fait que ce

gène est également peu exprimé dans les feuilles. En effet, les parties apicales utilisées étant constituées en grande partie de primordia foliaires, le niveau d'expression de *F5H* dans ces tissus devait être trop faible pour pouvoir être détecté. Campbell et Sederoff (1996) expliquent que les séquences promotrices de certains de ces gènes contiennent des éléments de séquence qui permettraient leur expression spécifique dans certains tissus.

Les niveaux d'expression étant très différents d'un tissu à l'autre (figure 5), les analyses ont donc été réalisées tissu par tissu. Dans le xylème, le gène *CCR* est le plus exprimé alors que les gènes *C3H* et *PAL2* sont les plus faiblement exprimés. Dans les tiges, ce sont les gènes *CCoAOMT1* et *CCoAOMT2* sont les plus exprimés alors que *F5H* et *PAL2* sont très peu exprimés, et enfin dans les feuilles, ce sont les gènes *CAD2* et *CCR* qui sont les plus exprimés, les autres gènes étant très faiblement exprimés.

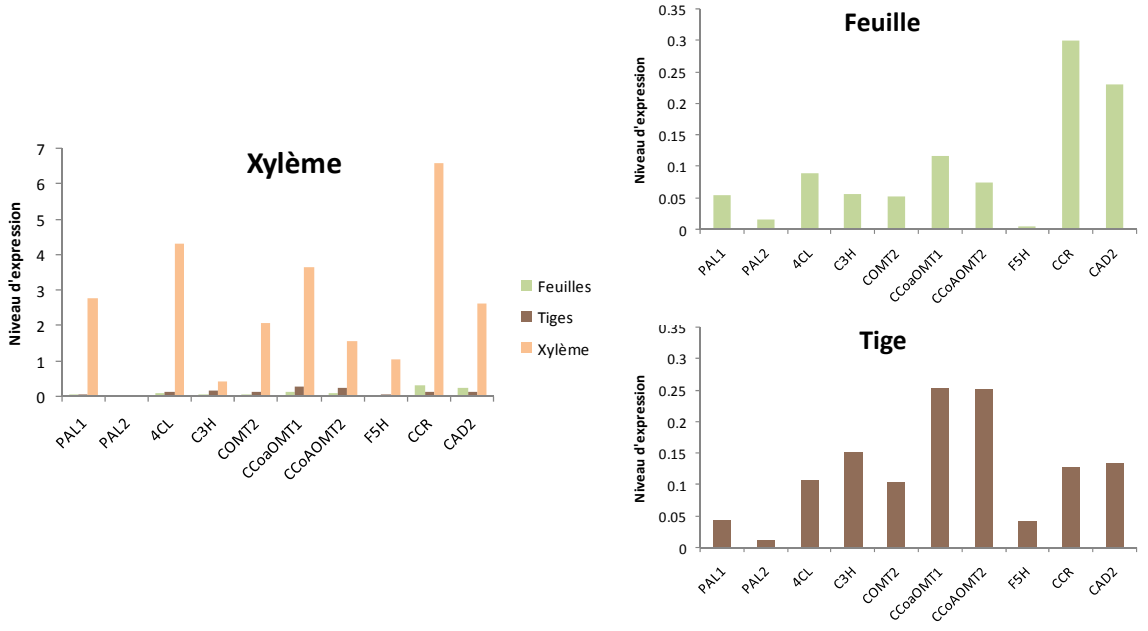


Figure 5 - Expression moyenne des différents gènes dans les différents tissus étudiés, exprimé en niveau d'expression normalisé.

L'analyse par tissu a été réalisée à différents points de prélèvement: pour le xylème au cours de la 1ère saison des pluies (P4), de la 2nde saison sèche (P5) et de la 2nde saison des pluies (P6), les prélèvements n'ayant pas été effectués auparavant. Pour les feuilles, les analyses ont été effectuées pour la 1ère saison sèche (P2) et aux P5 et P6 et, enfin pour les tiges nous avons réalisé les analyses d'expression sur tous les prélèvement disponibles.

Dans le xylème, les gènes *CAD2*, *CCoAOMT1*, *COMT2* et *F5H* sont plus fortement exprimés au point de prélèvement 5 (seconde saison sèche). Dans les tiges, tous les gènes étudiés sont plus exprimés au cours du prélèvement 3 (début de la 1ère saison des pluies), alors qu'aucune différence d'expression entre les points de prélèvement n'a pu être observée dans les feuilles (figure 6).

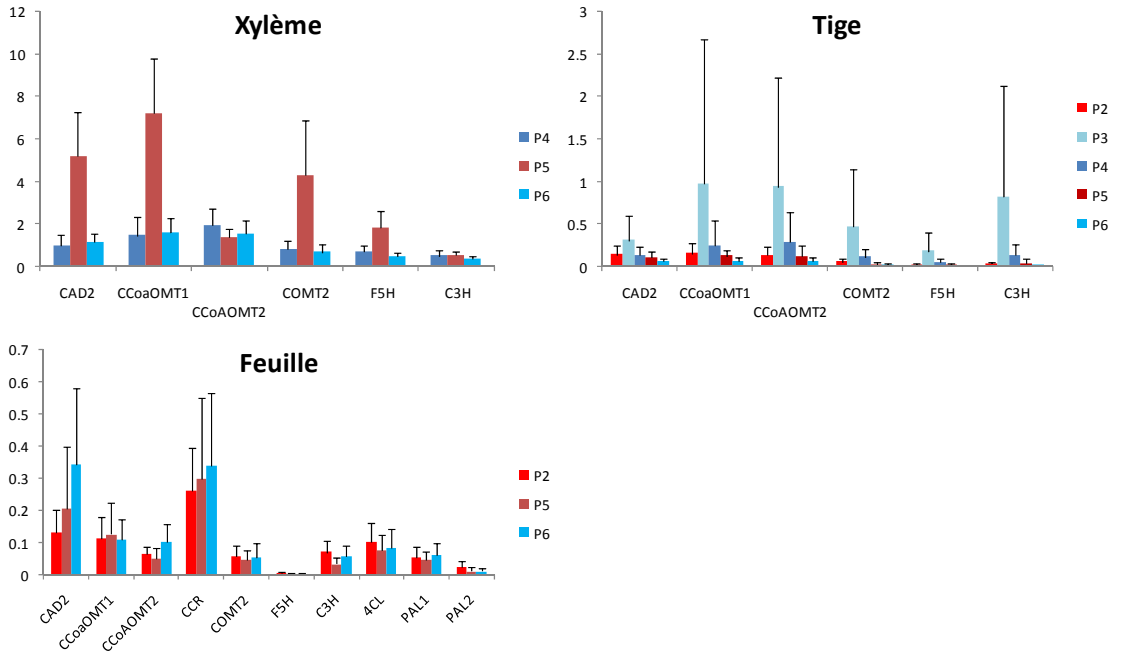


Figure 6 - Expression moyenne des différents gènes pour chaque point de prélèvement dans les différents tissus étudiés, exprimé en niveau d'expression normalisé.

Les ANOVA effectuées pour chaque point de prélèvement ont révélé que peu de gènes présentaient des différences d'expression (tableau 3). Dans les tiges, les gènes de la voie de biosynthèse semblent avoir une expression relativement stable quelque soit le génotype et le traitement, excepté *F5H* et *PAL2*. Les feuilles et le xylème présentent un plus grand nombre de gènes présentant des expressions différentes en fonction des conditions testées.

Tableau 3 - Résultats de l'ANOVA effectuée pour chaque point de prélèvement. Les réponses plastiques (présentant soit un effet T, soit un effet GxT) sont en gris.

	Xylème			Tige					Feuille		
Genes	P4	P5	P6	P2	P3	P4	P5	P6	P2	P5	P6
PAL1		x		x			x		x	x	T*
PAL2		x		x			T*		T*	x	G**
4CL		G*		x			x		x	T*	T*
C3H	G**; T*; GxT*	x	x	x	x	x	x	x	x	x	x
COMT2	G*; GxT*	x	G*	x	x	x	x	x	G*	x	x
CCoAOMT1	x	T*	x	x	x	x	x	x	x	x	x
CCoAOMT2	T**	x	x	x	x	x	x	x	x	x	G*
F5H	x	x	x	x	x	G*	x	G*; T*; GxT**	x	x	x
CCR		x		x			x		x	G*	x
CAD2	x	G*	x	x	x	x	x	x	x	T*	x

Dans les tiges, des effets du traitement sont observables au P5 (fin de la 2nde saison sèche) et au P6 (saison des pluies). Dans les feuilles, les effets du traitement sont visibles pour tous les points de prélèvement testés (P2, P5, P6), tandis que dans le xylème, les effets du traitement sont observables au P4 et au P5.

Nous avons porté plus d'intérêt aux gènes présentant de la plasticité, c'est à dire soit un effet T, soit un effet GxT (tableau 4). Dans les tiges, les gènes *F5H* au P6 et *CAD*2 au P5 sont surexprimés dans le traitement NI. Au P6, le gène *F5H* présente de l'interaction GxT car l'augmentation de l'expression en traitement NI est plus importante chez le génotype 1-41 que chez le 18-50.

Dans les feuilles, pendant les saisons sèches (P2 et P5), les gènes *PAL*2, *CAD*2 et *4CL* sont surexprimés en NI, alors qu'au cours de la saison des pluies les gènes *4CL* et *PAL*1 sont surexprimés en IR. Enfin dans le xylème, la même tendance est observée : pendant la saison des pluies (P4), les gènes *CCoAOMT*2 et *C3H* sont surexprimés en IR, alors qu'au cours de la saison sèche, *CCoAOMT*1 est surexprimé en NI. Le gène *COMT*2 présente un patron d'expression plus singulier, avec une surexpression en IR pour 1-41, tandis qu'il est surexprimé en NI chez 18-50, ce qui donne lieu à un effet GxT significatif, traduisant une plasticité liée à des effets de classement.

Tableau 4 - Niveaux relatifs d'expression par génotype et par traitement (moyennes ± écart-type), p-value de l'ANOVA (*<0.05 ; **<0.01), et indice plasticité (IP = (NIR-NNI)/NIR pour chaque génotype).

			Niveau d'expression relatif moyen				p-value ANOVA			Indice de plasticité	
Gène	**Tissu**	**Prélèvement**	**1-41IR**	**1-41NI**	**18-50IR**	**18-50NI**	**G**	**T**	**GxT**	**1-41**	**18-50**
PAL1	Feuille	P6	0.09 ± 0.04	0.03 ± 0.01	0.07 ± 0.03	0.06 ± 0.04	0.692	0.032*	0.148	0.72	0.20
PAL2	Tige	P5	0.0008 ± 0.0006	0.0035 ± 0.0032	0.0011 ± 0.0003	0.0044 ± 0.0037	0.653	0.031*	0.805	-4.26	-4.07
PAL2	Feuille	P2	0.012 ± 0.011	0.023 ± 0.007	0.019 ± 0.018	0.043 ± 0.023	0.112	0.044*	0.424	-0.98	-1.31
4CL	Feuille	P5	0.08 ± 0.06	0.11 ± 0.03	0.03 ± 0.02	0.1 ± 0.03	0.205	0.038*	0.291	-0.32	-2.19
4CL	Feuille	P6	0.15 ± 0.08	0.05 ± 0.03	0.06 ± 0.03	0.06 ± 0.02	0.085	0.044*	0.055	0.65	0.05
C3H	Xylème	P4	0.93 ± 0.2	0.44 ± 0.06	0.37 ± 0.14	0.37 ± 0.1	0.004**	0.014*	0.014*	0.53	0.00
COMT2	Xylème	P4	0.8 ± 0.02	0.39 ± 0.05	0.81 ± 0.11	1.22 ± 0.49	0.026*	0.794	0.044*	0.51	-0.50
CCoAOMT1	Xylème	P5	6.44 ± 1.14	7.77 ± 2.98	5.14 ± 0.67	10.36 ± 2.9	0.738	0.012*	0.101	-0.21	-1.02
CCoAOMT2	Xylème	P4	3.08 ± 0.9	1.3 ± 0.06	2.14 ± 0.54	1.52 ± 0.61	0.321	0.009**	0.129	0.58	0.29
F5H	Tige	P6	0.0009 ± 0.0005	0.0096 ± 0.0049	0.0022 ± 0.0027	0.0008 ± 0.0006	0.021*	0.024*	0.004**	-10.92	-0.35
CAD2	Feuille	P5	0.12 ± 0.05	0.32 ± 0.2	0.07 ± 0.05	0.36 ± 0.26	0.888	0.021*	0.661	-1.77	-4.38

4 Discussion

L'objectif de cette étude était d'observer si les gènes de la voie de biosynthèse des lignines pouvaient être impliqués dans la réponse à la contrainte hydrique, en lien avec la variation de certains caractères mis en évidence à d'autres échelles (anatomique ou chimique). Dans cette voie de biosynthèse aussi complexe, l'existence de gènes paralogues a semblé être un paramètre clé pour le déploiement de la plasticité moléculaire au niveau expressionnel, et pourrait résulter de l'application de certaines forces évolutives ayant permis aux génotypes de s'adapter aux contraintes environnementales.

4.1 Les gènes de la voie de biosynthèse de la lignine sont différemment exprimés en fonction des génotypes, de la contrainte hydrique, mais aussi au cours du temps.

Dans les banques d'ADNc d'apex caulinaires, les membres identifiés sont majoritairement différentiellement exprimés selon les génotypes, ce qui pourrait indiquer que ces différences d'expression sont liées à des mécanismes de spéciation (Pavey *et al.*, 2010). Néanmoins, lorsque l'analyse d'abondance révèle de la plasticité (effet T ou GxT), ces gènes sont surexprimés dans la condition irriguée, indiquant que leur expression serait plutôt liée à des processus de développement plutôt qu'à des mécanismes de défense vis-à-vis de la contrainte hydrique.

Lorsqu'on regarde l'expression au cours des 2 années de croissance, dans les feuilles et dans le xylème, les gènes qui présentent des effets T ou GxT sont surexprimés lors de la saison des pluies chez les arbres ayant été irrigués (IR) au cours des saisons sèches (*PAL*1, *C3H*, *COMT*2 et *CCoAOMT*2). Au moment où la contrainte est appliquée, c'est-à-dire en saison sèche, ce sont d'autres gènes (*PAL*2, *CCoAOMT*1, *CAD*2) qui sont surexprimés chez les non-irrigués. A l'exception de *4CL*, qui est surexprimé pendant la saison des pluies en IR et pendant la saison sèche en NI, il semble que les membres impliqués dans les mécanismes de développement (surexprimés en IR) ou dans des processus de réponse à la contrainte (surexprimé en NI) soit différents. C'est particulièrement le cas pour les familles *PAL* et *CCoAOMT*, pour lesquelles il a été possible d'étudier l'expression de deux membres différents. Ces résultats permettent de penser qu'il pourrait exister des voies de biosynthèse parallèles : soit constitutives, soit induites par la contrainte hydrique. Dans les feuilles, l'augmentation de la teneur en hydroxycinnamates aux propriétés antioxydantes dans la condition NI pendant la saison sèche

(Chapitre 4) associée à la surexpression de *4CL*, pourrait confirmer que la voie de biosynthèse des phényl-propanoïdes pourrait être activée en cas de contrainte hydrique, afin de protéger les cellules des dommages causés par le manque d'eau.

4.2 Les familles de gènes étudiées semblent avoir évolué différement

Au cours de cette recherche dans les banques d'ADNc d'apex, nous avons mis en évidence 40 contigs uniques, qui ne présentaient aucune similarité avec un autre des 115 contigs identifiés comme membre de la voie de biosynthèse des phényl-propanoïdes. 17 contigs ont été mis en évidence comme des variants d'épissage alternatif, et ont été regroupés au final en 8 contig uniques. Enfin, 58 contigs présentaient des similarités de séquences avec au moins un autre contig. Ces derniers ont été assimilés à des allèles différents mal assemblés, et regroupés en 16 contigs uniques. Cette analyse révèle la difficulté de trouver les paramètres optimaux lors de l'assemblage des différentes lectures en contigs. Ces paramètres résultent toujours d'un compromis entre un seuil de similarité très stringent afin de ne pas assembler des lectures provenant de différents membres d'une même famille multigénique ensemble, et un seuil de similarité entre séquences moins stringent afin d'assembler dans le même contig les lectures provenant des différents allèles d'un même gène. Ce problème est particulièrement important dans le cadre de notre étude car les génotypes séquencés sont des hybrides d'espèces différentes, donc a priori très hétérozygotes. De plus, le genre *Eucalyptus* montre des niveaux de diversité nucléotidique très élevés (Myburg *et al.*, 2007 ; Kulheim *et al.*, 2009), ce qui pourrait renforcer les difficultés à assembler différents allèles dans un même contig. Les familles *CAD*, *CCoAOMT*, *CCR* et *HCT* présentaient le plus de contigs correspondant à différents allèles mal assemblés, ce qui pourrait suggérer que ces familles présentent un niveau de polymorphisme plus élevé que les autres. Ces tendances ont déjà été observées par Mandrou (2010) : lors d'une recherche de SNP par séquençage de 8 gènes de structure de la voie de biosynthèse des phényl-propanoïdes dans une population d'*E. urophylla*, il a montré que les densités de SNP pouvait varier d'un gène à l'autre d'1SNP/15pb à 1SNP/80pb dans les régions codantes.

Différentes hypothèses ont été formulées pour comprendre la signification biologique des différents membres pour les familles de gènes de la voie de biosynthèse des phényl-propanoïdes. Dixon *et al.* (2003) proposent que la production de différents produits de la voie de biosynthèse doit être régulée de façon indépendante que ce soit dans les mêmes cellules ou dans des cellules différentes. Dans ce cas les différents membres conduiraient à la production de différentes classes de composés. De la même façon, Xu *et al.* (2009) propose que le

nombre de membres pour une famille donnée pourrait être liée à la diversité des substrats qu'elle peut utiliser. Ainsi les familles *4CL*, *CCR*, *COMT* et *CAD* pouvant utiliser plusieurs substrats présentent un grand nombre de membres différents, tandis que les gènes *PAL*, *C4H*, *C3H*, *HCT* et *CCoAOMT* qui peuvent utiliser un ou deux substrats différents présentent peu de membres. Ces différences en terme de conservation de la famille renseignent sur l'histoire évolutive du génome. Cela a donc amené à émettre des hypothèses en terme de pression de sélection les plus importantes pour l'évolution d'une famille de gènes. En comparant la conservation des familles multigéniques à travers différentes espèces, Xu *et al.* (2009) suggèrent que l'expansion récente et rapide des gènes de la voie de biosynthèse des lignines pourrait être due principalement à des forces évolutives liés à des mécanismes de défense plutôt qu'à des processus de développement.

5 Conclusion

Cette étude focalisée sur les gènes de la voie de biosynthèse des lignines a montré :

1/ que le séquençage en haut-débit a permis un accès quasi immédiat aux différents membres des familles multigéniques, voire aux allèles et aux variants d'épissage alternatif.

2/ que les membres d'une même famille multigénique peuvent avoir des patrons d'expression différents, suggérant l'existence de voies de biosynthèse parallèles, certaines étant constitutives, et d'autres induites.

L'analyse de l'expression de plusieurs des membres identifiés dans les banques d'ADNc sur les différents tissus à différents temps de prélèvement pourrait confirmer ou infirmer la tendance observée dans cette première étude. Cette analyse pourrait être ciblée particulièrement sur les membres des familles *CAD*, *CCoAOMT* et *COMT* car de précédentes études ont montré leur implication dans la réponse à la contrainte hydrique chez d'autres espèces. Ainsi, Hu *et al.* (2009) ont montré que l'expression des gènes *CAD* et *COMT* augmente dans les feuilles chez des lignées de maïs tolérantes à la sécheresse lorsqu'elles sont soumises stress hydrique. De plus, Berta *et al.* (2010) ont montré que les gènes *CCoAOMT* et *CAD* étaient surexprimés dans le cambium de peuplier durant les périodes de récupération de la contrainte hydrique. On pourrait également penser à mesurer différents caractères phénotypiques en lien avec cette voie de biosynthèse, pour essayer d'observer des covariations entre caractères et expression de gènes. Par exemple, la quantification des dépôts de lignines dans les tiges et dans les feuilles pourrait être effectuée par analyse anatomique

grâce à l'autofluorescence qu'émettent les lignines. On pourrait également essayer de quantifier ces lignines et éventuellement le rapport des monomères Syringyl/Guaiacyl dans les différents organes par des méthodes chimiques (méthode Klason)..

REFERENCES BIBLIOGRAPHIQUES

Antal, M. J., S. G. Allen, *et al.* (2000). "Attainment of the theoretical yield of carbon from biomass." Industrial & Engineering Chemistry Research **39**(11): 4024-4031.

Baucher, M., C. Halpin, *et al.* (2003). "Lignin: Genetic engineering and impact on pulping." Critical Reviews in Biochemistry and Molecular Biology **38**(4): 305-350.

Berta, M., A. Giovannelli, *et al.* (2010). "Transcriptome changes in the cambial region of poplar (Populus alba L.) in response to water deficit." Plant Biology **12**(2): 341-354.

Boerjan, W., J. Ralph, *et al.* (2003). "Lignin biosynthesis." Annual Review of Plant Biology **54**: 519-546.

Campbell, M. M. and R. R. Sederoff (1996). "Variation in lignin content and composition - Mechanism of control and implications for the genetic improvement of plants." Plant Physiology **110**(1): 3-13.

Chaumeil, P. (2006) Plasticité moléculaire de deux écotypes de pin maritime soumis à un stress osmotique. Thèse de doctorat.

Dixon, R. A. and M. S. Srinivasa Reddy (2003). "Biosynthesis of monolignols. Genomic and reverse genetic approaches." Phytochemistry Reviews **2**(3): 289-306.

Foucart, C., A. Jauneau, *et al.* (2009). "Overexpression of EgROP1, a Eucalyptus vascular-expressed Rac-like small GTPase, affects secondary xylem formation in Arabidopsis thaliana." New Phytologist **183**(4): 1014-1029.

Gion, J. M., P. Rech, *et al.* (2000). "Mapping candidate genes in Eucalyptus with emphasis on lignification genes." Molecular Breeding **6**(5): 441-449.

Goicoechea, M., E. Lacombe, *et al.* (2005). "EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis." Plant Journal **43**(4): 553-567.

Hu, Y., W. C. Li, *et al.* (2009). "Differential expression of candidate genes for lignin biosynthesis under drought stress in maize leaves." Journal of Applied Genetics **50**(3): 213-223.

Hunter, S., R. Apweiler, *et al.* (2009). "InterPro: the integrative protein signature database." Nucleic Acids Research **37**: D211-D215.

Kim, D. G., Y. J. Kim, *et al.* (2005). "Effect of wounding and chemical treatments on expression of the gene encoding cinnamate-4-hydroxylase in Camptotheca acuminata leaves." Journal of Plant Biology **48**(3): 298-303.

Kulheim, C., S. H. Yeoh, *et al.* (2009). "Comparative SNP diversity among four Eucalyptus species for genes from secondary metabolite biosynthetic pathways." Bmc Genomics **10**.

Lauvergeat, V., C. Lacomme, *et al.* (2001). "Two cinnamoyl-CoA reductase (CCR) genes from Arabidopsis thaliana are differentially expressed during development and in response to infection with pathogenic bacteria." Phytochemistry **57**(7): 1187-1195.

Legay, S., E. Lacombe, *et al.* (2007). "Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway." Plant Science **173**: 542-549.

Mandrou, E. (2010) Variabilité fonctionnelle de gènes candidats de la lignification chez l'eucalyptus. Thèse de doctorat.

Moura, J., C. A. V. Bonine, *et al.* (2010). "Abiotic and Biotic Stresses and Changes in the Lignin Content and Composition in Plants." Journal of Integrative Plant Biology **52**(4): 360-376.

Myburg, A. A., B. M. Potts, *et al.* (2007). Eucalypts. Genome Mapping and Molecular Breeding in Plants, Forest Trees. C. Kole. Berlin Heidelberg, Deutschland, Springer-Verlag. **7:** 115-160.

Nicot, N., J. F. Hausman, *et al.* (2005). "Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress." Journal of Experimental Botany **56**(421): 2907-2914.

Pavey, S. A., H. Collin, *et al.* (2010). "The role of gene expression in ecological speciation." Annals of the New York Academy of Sciences **1206**(1): 110-129.

Ramakers, C., J. M. Ruijter, *et al.* (2003). "Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data." Neuroscience Letters **339**(1): 62-66.

Reid, K. E., N. Olsson, *et al.* (2006). "An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development." BMC Plant Biology **6**(27): (14 November 2006).

Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-86.

Ruijter, J. M., C. Ramakers, *et al.* (2009). "Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data." Nucleic Acids Research **37**(6).

Savidge, R. A. (2000). "Biochemistry of seasonal cambial growth and wood formation - an overview of the challenges." Cell and Molecular Biology of Wood Formation: 1-30.

Shi, R., Y. H. Sun, *et al.* (2010). "Towards a Systems Approach for Lignin Biosynthesis in Populus trichocarpa: Transcript Abundance and Specificity of the Monolignol Biosynthetic Genes." Plant and Cell Physiology **51**(1): 144-163.

Thumma, B. R., M. R. Nolan, *et al.* (2005). "Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in microfibril angle in Eucalyptus spp." Genetics **171**(3): 1257-1265.

Thumma, B. R., S. G. Southerton, *et al.* (2010). "Quantitative trait locus (QTL) analysis of wood quality traits in Eucalyptus nitens." Tree Genetics & Genomes **6**(2): 305-317.

Vandesompele, J., K. De Preter, *et al.* (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biology **3**(7).

Vincent, D., C. Lapierre, *et al.* (2005). "Water deficits affect caffeate O-methyltransferase, lignification, and related enzymes in maize leaves. A proteomic investigation." Plant Physiology **137**(3): 949-960.

Wang, L. K., Z. X. Feng, *et al.* (2009). "DEGseq: an R package for identifying differentially expressed genes from RNA-seq data." Bioinformatics **26**(1): 136-138.

Xu, Z. Y., D. D. Zhang, *et al.* (2009). "Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom." BMC Bioinformatics **10**.

Chapitre VII - Discussion générale et perspectives

Lac de Lalla, Takerskout, Matthieu

1 Mise en place d'un dispositif expérimental pour étudier la plasticité phénotypique et moléculaire de l'eucalyptus

Le dispositif expérimental mis en place a permis d'obtenir une vision détaillée du comportement de deux génotypes (utilisés en plantation industrielle en République du Congo) face une contrainte hydrique au cours de leur deux premières années de croissance. Ce dispositif a fédéré pendant 3 ans différentes disciplines scientifiques présentes au sein d'équipes du CIRAD, de l'INRA et du CRDPI.

La mise en place d'un tel dispositif sur savane au Congo était très ambitieux, et il nous a fallu faire face à de nombreuses contraintes techniques. Nous cherchions un sol de savane, le plus plan possible, proche d'un point d'eau indispensable au système d'irrigation, et dans une zone facile d'accès. Le dispositif expérimental a finalement été installé au bord du lac Yanika, à 60 km de la ville de Pointe Noire. Du fait de son éloignement des stations CRDPI, le site expérimental nécessitait une relative autonomie, en termes d'alimentation électrique, de matériel à disposition et de surveillance.

Cet essai était constitué d'un factoriel croisant deux génotypes avec deux traitements, répété autant de fois que de prélèvements destructifs à effectuer. Comme nous l'avons vu dans le chapitre 3, des différences de croissance entre les placeaux ont été observés, rendant plus difficile l'interprétation des variations des caractères mesurés d'un point de prélèvement à un autre, puisqu'ils correspondaient à des placeaux différents. Pourtant, d'après les mesures de spectres proche infra-rouge, les sols semblent être homogènes sur tout l'essai. Pour parer à ce problème on aurait pu envisager une disposition aléatoire sur l'essai des arbres à prélever pour chaque point de prélèvement. Néanmoins, cela aurait rendu très difficile l'application des traitements d'irrigation, et le choix des individus à prélever au fur et à mesure des prélèvements: en effet, l'abattage des arbres au cours des prélèvements aurait diminué la densité de l'essai et donc la concurrence entre arbres, modifiant par voie de conséquence le comportement global des arbres.

La gamme environnementale testée au cours de cette étude consistait en une levée par irrigation ou non de la contrainte hydrique subie au cours de la saison sèche. Cela nous a donc donné un aperçu des effets que pouvaient avoir la limitation de l'alimentation hydrique en

condition naturelle, mais pas des effets que pourraient avoir des saisonnalités plus marquées, c'est-à-dire des saisons sèches plus longues, telles qu'elles sont prédites pour les années à venir aux vues des changements climatiques en cours.

Dans le traitement irrigué en saison sèche, les accroissements en biomasse de bois sont plus importants, mais diminuent quand même par rapport aux accroissements moyens pendant la saison des pluies. Ce phénomène peut être expliqué soit par le fait que l'apport en eau n'était pas suffisant pour conserver un accroissement élevé pendant la saison sèche, soit que d'autres contraintes réduisant la croissance n'ont pas été levées. On peut par exemple penser que la diminution des radiations lumineuses en saison sèche influence le rendement photosynthétique, et donc, limite la croissance. Comme il a été soulevé dans le chapitre 3, les eucalyptus sont connus pour avoir un enracinement profond (Drake et al., 2009; Christina et al., 2011), ce qui pourrait laisser penser qu'ils puissent accéder à la nappe phréatique, notamment au cours de la seconde saison sèche. Pourtant, les différentes mesures de croissance et de discrimination isotopique du carbone montrent des différences significatives entre traitement irrigué et non-irrigué, la contrainte hydrique a donc eu des répercutions sur la physiologie et la croissance des plants. Des mesures de potentiel hydrique foliaire, ou de teneur en eau relative de certains organes (la quantité d'eau que contient l'organe par rapport à la quantité maximale d'eau qu'elle peut contenir) auraient pu permettre de mieux caractériser le stress hydrique subi par les arbres.

La stratégie d'échantillonnage utilisée proposait de choisir, parmi les 24 arbres de chaque placeau, les 11 arbres les plus représentatifs, c'est-à-dire ceux dont la hauteur était la plus proche de la hauteur médiane du placeau. Si cela a rendu plus délicate la modélisation des relations allométriques entre production de biomasse et croissance, puisque la hauteur était peu variable pour l'échantillon mesuré, cette stratégie a permis de réduire la variance des caractères au sein du placeau, et ainsi de mettre en évidence des effets de l'interaction GxE. On aurait pu améliorer la stratégie d'échantillonnage en prélevant 4 arbres dont la hauteur était proche de la médiane pour les disciplines où l'on ne pouvait pas analyser un grand nombre d'individus (anatomie, transcriptomique et protéomique), et pour les 7 individus restant, choisir des arbres plus représentatifs de la variation au sein de chaque placeau, en échantillonnant une plus grande proportion dans les classes supérieures, car ce sont elles qui présentent le plus de variabilité (pour les caractères de croissance).

Finalement, malgré les difficultés techniques qu'il a fallu contourner, les résultats obtenus en termes de caractérisation de la plasticité mise en œuvre par les génotypes testés en réponse à

Tableau 1 : Récapitulatif des principaux résultats obtenus dans cette thèse. Pour chaque génotype sont données pour chaque traitement (Irrigué, IR et non-irrigué, NI), les valeurs moyennes obtenues pour chaque caractère mesuré. Les indices de plasticité sont calculés de la façon suivante : IP = $-(X_{IR} - X_{NI})/X_{IR}$. Un code couleur a ainsi pu être attribué : en rouge foncé les caractères augmentant fortement en NI (IP>25%), en rouge clair les caractère augmentant plus faiblement en NI (0<IP<25%) ; en bleu clair les caractères augmentant faiblement en IR (-25%<IP<0) et en bleu foncé, les caractères augmentant fortement en IR (PI<-25%).

			Génotype 1-41			Génotype 18-50			
Catégorie	**Caractère**	**Effets**	**Condition IR**	**Condition NI**	**IP (%) IR → NI**	**Condition IR**	**Condition NI**	**IP (%) IR → NI**	**Chapitre**
Croissance	volume sous écorce à 20 mois en m3/ha	T***	31.8	16.6	-48	31.3	18.8	-40	3
	accroissement en saison des pluies (dec07-juil08) en m3/ha/an	G*** T***	22.4	19.2	-14	32.1	25.2	-22	3
	accroissement en saison sèche (juil08-sept08) en m3/ha/an	T***	28.4	13.5	-53	31.4	12.4	-61	3
	accroissement en saison des pluies (sep08-janv09) en m3/ha/an	G*** T*** GxT*	40.2	17.5	-56	29.8	14.7	-51	3
Biomasse de bois	biomasse cumulée à 20 mois (T/ha) en T/ha	G** T***	20.1	11.3	-44	23.7	14.8	-38	3
	accroissement en saison des pluies (dec07-avr08) en T/ha/an	G*** T*** GxT*	20.7	14.5	-30	30.0	19.0	-37	3
	accroissement saison sèche (avr08-sept08) en T/ha/an	T***	13.9	10.7	-23	13.7	12.0	-12	3
	accroissement en saison des pluies (sep08-janv09) en T/ha/an	G* T***	21.2	10.1	-52	24.9	15.8	-37	3
Biomasse foliaire	biomasse cumulée en T/ha		2.87	1.91	-33	3.39	2.43	-28	3
	accroissement en saison des pluies (dec07-avr08) en T/ha/an	T***	7.35	7.65	4	6.95	7.78	12	3
	accroissement saison sèche (avr08-sept08) en T/ha/an	G*** T***	1.27	0.81	-36	0.93	0.66	-29	3
	accroissement en saison des pluies (sep08-janv09) en T/ha/an	G** T*	-0.24	-0.83	-250	0.57	0.03	-95	3
Biomasse racinaire	biomasse cumulée en T/ha		1.99	1.66	-17	2.26	2.01	-11	3
	accroissement en saison des pluies (dec07-avr08) en T/ha/an	G*** T*	1.70	1.46	-14	2.34	2.09	-11	3
	accroissement saison sèche (avr08-sept08) en T/ha/an	G*** T*** GxT*	1.98	2.43	23	1.37	2.09	52	3
	accroissement en saison des pluies (sep08-janv09) en T/ha/an	G** T*	1.24	0.73	-41	1.90	1.10	-42	3
Allocation de biomasse	ratio S/R saison des pluies (avril 08)	G** T***	10.4	10.9	5	10.2	9	-12	3
	ratio S/R saison sèche (sept 08)	G*** T***	8.9	7.3	-18	9.9	8.3	-16	3
	ratio S/R saison des pluies (janv 09)	G* T*** GxT**	11.5	7.7	-33	10.3	9.2	-11	3
Morphologie / Physiologie	SLA fin de saison des pluies (avril 08) en m2/kg	G*** T** GxT*	9.93	9.7	-2	12.4	10.6	-15	3
	SLA fin de saison sèche (sept 08) en m2/kg	G** T*** GxT*	10.5	9.5	-10	11.8	9.8	-17	3, 4
	δ13C fin de saison des pluies (avril 08) en ‰	G*** T** GxT**	-30.8	-30.8	0	-29.4	-30.3	-3	3
	δ13C fin de saison sèche (sept 08) en ‰	G*** T***	-31.6	-30.6	3	-30.6	-29.5	4	3, 4
	MLA (20 mois) en cm²	G*** GxT*	66.8	70.2	5	34	30.2	-11	3
	Densité stomatique saison sèche en nb/mm²	T*	122	115	-6	130	113	-13	4
Anatomie	Collenchyme, face supérieure en %	GxT*	39	35	-10	37.3	44.4	19	4
	Epiderme, face supérieure en µm	G* T*	16.2	11	-32	10.8	10.6	-2	4
	Epiderme, face inférieure en µm	G* T* GxT*	15	11.7	-22	11.6	10	-14	4
	Cuticule (2 faces) en µm	G* T***	26.4	37.2	41	17.6	32.3	84	4
	Vaisseau du xylème en %	T*	5.1	4.7	-8	6	5	-17	4
Polyphénols	Acides hydroxycinnamiques en mAu/min	G** T** GxT**	38.9	41.3	6	41.6	76.9	85	4
	Flavones en mAu/min	G***	23.4	24.3	4	0	0	0	4
	Falvonols en mAu/min		123.4	115	-7	103.3	101.6	-2	4

Catégorie	Caractère	Effet	Génotype 1-41 Condition IR	Génotype 1-41 Condition NI	Génotype 1-41 IP (%) IR → NI	Génotype 18-50 Condition IR	Génotype 18-50 Condition NI	Génotype 18-50 IP (%) IR → NI	Chapitre
Expression dans les apex	Hormones		+ NI			++ IR			5
	Messagers secondaires		+ NI (signalisation lumière)			++ NI (kinases réceptrices)			5
						++IR (signalisation lumière)			5
	Autres gènes liés au stress		+ NI (protéines de choc thermique)			++ NI (protéines PR)			5
						++IR (protéine de choc thermique)			5
	Facteurs de transcription		0			++ NI			5
	Protection des cellules		+ NI			++ IR			5
	Réparation des dommages		+ NI (carbohydrates, metabolisme			++ NI (organisation cellulaire,			5
						++ IR (métabolsime II)			5
	Activité cellulaire		0			++ IR (photosynthèse, transport)			5
	phenylpropanoïdes		0			++ IR			6
Expression dans les feuilles matures	aquaporines - dTIP	G** T** GxT*	513	102	-80	124	37	-70	4
	isoprènes - DXR	T**	112	32	-71	79	21	-73	4
	polyamines - DCAM	G** T***	524	136	-74	246	52	-79	4
	signalisation de stress de lumière - ELIP	G** GxT*	161	242	50	131	43	-67	4
	monolignols - CAD2 (saison sèche)	T*	117	325	177	66	355	438	6
	phénylpropanoides - 4CL (saison sèche)	T*	81	107	32	32	102	219	6
	phénylpropanoides - PAL2 (saison sèche)	T*	12	23	98	19	43	131	6
	phénylpropanoides - 4CL (saison des pluies)	T*	155	54	-65	62	59	-5	6
	phénylpropanoides - PAL1 (saison des pluies)	T*	89	25	-72	71	57	-20	6
Expression dans le xylème	phénylpropanoides -C3H (saison des pluies)	G** T* GxT*	935	442	-53	366	366	0	6
	phénylpropanoides - COMT2 (saison des pluies)	G* GxT*	800	393	-51	813	1220	50	6
	phénylpropanoides - CCoAOMT1 (saison sèche)	T*	6436	7774	21	5135	10363	102	6
	phénylpropanoides - CCoAOMT2 (saison des pluies)	T**	3082	1296	-58	2140	1516	-29	6

la contrainte hydrique, font de ce projet, une excellente base de travail pour la thématique de recherche abordée.

2 Une vision intégrative de la réponse à la contrainte hydrique pour deux clones d'eucalyptus

Ce projet a permis de dresser le profil de deux génotypes soumis à deux régimes hydriques différents au cours de la saison sèche, aussi bien au niveau de leurs caractéristiques de croissance, mais également du point de vue de leur physiologie, anatomie et réponse moléculaire. Les résultats montrent que pour la majorité des caractères étudiés, les différences entre génotypes expliquent une grande partie de la variation du caractère. Il a également été possible de mettre en évidence de la plasticité phénotypique pour un grand nombre de caractères. Si la plasticité mise en évidence reflète, pour la plupart des caractères étudiés, des réponses communes aux deux génotypes étudiés (voire commune avec d'autres espèces), des réponses spécifiques aux deux génotypes ont aussi été mises en évidence pour certains caractères. Un récapitulatif des principaux résultats obtenus est présenté dans le tableau 1 et s'intéresse principalement à la 2ème saison sèche, car les arbres avaient alors atteint une croissance suffisante pour permettre d'observer les effets de la contrainte hydrique à différentes échelles.

2.1 Des différences marquées entre génotypes

Les deux clones commerciaux, hybrides issus d'espèces différentes, avaient été choisis parce qu'ils présentaient des caractéristiques très différentes au stade mature en termes de productivité et de paramètres physiologiques.

Sur le dispositif expérimental, ces différences en terme de productivité ont été confirmées très rapidement après la plantation. En effet, en moyenne, dès 4 mois, le génotype 18-50 présente une croissance supérieure et produit plus de bois que 1-41 (chapitre 3). Le génotype 18-50 possède une biomasse foliaire plus importante que le génotype 1-41 quelque soit le traitement. Cela doit lui permettre d'avoir une surface photosynthétique plus élevée, et donc d'avoir des meilleures capacités de croissance.

Les caractéristiques foliaires de ces deux génotypes sont de prime abord particulièrement différentes (chapitre 4): les feuilles du 18-50 sont plus petites et plus fines, et vraisemblablement en plus grand nombre que pour le génotype 1-41 étant donné que les

biomasses foliaires sont plus faibles pour ce génotype. Le génotype 18-50 a en moyenne une surface foliaire spécifique plus élevée que 1-41, et dispose de stomates sur la face inférieure uniquement alors que le génotype 1-41 en possède sur les deux faces. Pendant la 2nde saison sèche, la production de feuilles diminue plus chez le 18-50, tandis qu'après la reprise des pluies, elle est plus faible chez le génotype 1-41, qui d'ailleurs, perd plus de feuilles qu'il n'en produit au cours des trois premiers mois de la 2nde saison des pluies, car l'accroissement est négatif pour ce génotype. Le génotype 18-50 montre également des niveaux de composition isotopique du carbone plus faibles, ce qui suggère qu'il est plus efficient vis-à-vis de l'alimentation hydrique. De plus, la réserve en eau du sol est toujours plus élevée pour le 18-50, ce qui suggère que le 18-50 serait plus économe en eau. Au niveau anatomique pourtant, mis à part des différences dues à des effets d'échelle, les proportions entre les différents tissus semblent être similaires entre les deux génotypes. Par contre les feuilles du génotype 1-41 présentent de plus fortes teneurs en polyphénols, particulièrement des flavones qui n'ont pas pu être détectées chez le 18-50.

Au niveau moléculaire, l'analyse du protéome de feuille et de tiges émises au cours de la 1ère saison sèche a mis en évidence 316 spots dont le volume était différent en fonction des génotypes parmi 593 spots différentiellement représentés. Au cours de l'analyse du transcriptome d'apex formé pendant la 2nde saison sèche, une grande proportion de gènes différentiellement exprimés en fonction des génotypes a également été mise en évidence. En effet, parmi les 1445 contigs différentiellement exprimés, 1280 présentaient un effet génotype (chapitre 5). On a observé le même nombre de contigs surexprimés par le génotype 1-41 de contigs surexprimés par 18-50, et aucune signature génotypique n'a pu être mise en évidence par rapport aux catégories fonctionnelles auxquelles ils appartiennent. On peut penser que ces différences d'expression révèlent des processus qui résultent du mécanisme de spéciation, car les hybrides utilisés proviennent d'espèces relativement différentes. En effet, si les deux hybrides étudiés appartiennent au sous-genre *Symphyomyrtus*, les espèces *E. urophylla* et *E. grandis* (parents du génotype 18-50) appartiennent à la section *Latoangulatae*, alors que l'espèce *E. alba* ("mère" du génotype 1-41) appartient à la section *Exsertaria*. Cet éloignement phylogénique a d'ailleurs eu des répercutions importantes sur l'assemblage des différentes lectures obtenues par séquençage 454 pour l'analyse du transcriptome. En effet, comme nous l'avons vu dans le chapitre 6, il est difficile d'obtenir un assemblage optimal des différentes lectures, étant donné la forte diversité génétique en présence chez les hybrides utilisés et le

niveau de complexité du génome d'eucalyptus (présence de nombreux membres pour une même famille multigénique).

2.2 Des réponses "génériques" à la contrainte hydrique ?

Un certain nombre des réponses à la contrainte hydrique mis en évidence sont communes aux deux génotypes étudiés. Parmi celles-ci, l'accroissement en biomasse de bois en traitement non-irrigué diminue drastiquement au cours de la saison sèche par rapport à la saison des pluies précédente chez les deux génotypes. Ce résultat était attendu, étant donné que plusieurs études ont montré l'importance de l'alimentation hydrique pour la croissance sur *Eucalyptus* (Campion et al., 2006; Stape et al., 2010) comme chez d'autres plantes (Chaves et al., 2003). L'accroissement en biomasse foliaire est également considérablement réduit au cours de la saison sèche pour les deux génotypes, même dans le traitement irrigué, bien qu'à une plus faible ampleur que dans le traitement non-irrigué. Au niveau des traits fonctionnels, les valeurs de composition isotopique du carbone augmentent dans le traitement non-irrigué pendant la saison sèche de la même façon pour les deux génotypes. Ce type de réponse a déjà été observé chez l'eucalyptus (Osorio et al. 1998), et chez un grand nombre d'autres plantes (revues par Warren et Adams, 2006). De même, au niveau anatomique, l'augmentation de l'épaisseur de la cuticule, la diminution de la densité stomatique ainsi que celle de l'épaisseur des vaisseaux est une réponse commune aux deux génotypes, voire à d'autres espèces (Riederer et Schreiber, 2001; Muchow et Sinclair, 1989; Hacke et Sauter, 1996).

Au niveau moléculaire, l'analyse de protéome effectuée sur les tissus récoltés au cours de la 1ère saison sèche a permis de mettre en évidence, parmi 1848 spots analysés, 394 spots présentant un effet du traitement et seulement 32 spots présentant de l'interaction GxT. A l'inverse, l'analyse du transcriptome en seconde saison sèche n'a mis en évidence que 155 contigs présentant un effet du traitement et 274 contigs présentant de l'interaction GxT, parmi 14460 contigs analysés. Ces différences peuvent être dues au fait qu'un certain nombre de régulations post-transcriptionnelles ou épigénétiques interviennent entre la synthèse d'ARN et la formation des protéines, tels qu'il a été mis en évidence en comparant les profils d'expression et d'accumulation protéiques (chapitre 5).

Au niveau du transcriptome, l'analyse des fonctions des gènes représentant un effet du traitement permet de distinguer des tendances générales. Ainsi, les gènes liés à la voie de biosynthèse de l'éthylène, majoritairement surexprimés en condition non-irriguée semblent être majoritairement impliqués dans la transduction du signal de stress. D'un autre côté, la sur-

expression de gènes liés à la dégradation de l'amidon, associée à la sous-expression de gènes liés à la synthèse de protéines ribosomales et de gènes liés au développement cellulaire, laisse penser que les arbres puisent dans leur ressources carbonées stockées et ont tendance à diminuer leur métabolisme sous contrainte hydrique. Cette tendance peut être confirmée par l'analyse protéomique qui a révélé que deux ATP synthases, sont sous-représentées en traitement non-irrigué. De telles réallocations de ressources semblent être un phénomène relativement commun à de nombreuses espèces, comme le montre la revue réalisée par Chaves *et al*. (2003).

2.3 Des réponses spécifiques, sous contrôle génétique ?

Dans certains cas, nous avons observé des intensités de réponse différentes entre les deux génotypes: la réponse au traitement a des effets similaires, mais les degrés de plasticité sont différents entre les deux génotypes. Cela suggère un contrôle génétique de la plasticité du caractère considéré.

Que ce soit pour les arbres des placeaux non-irrigués pendant les deux saisons sèches ou seulement pour la 1ère saison sèche, les accroissements en biomasse ré-augmentent rapidement lorsque les pluies reprennent pour le génotype 18-50, mais restent faibles pour 1-41. Ainsi, le génotype 18-50 pourrait avoir une plus grande capacité à retrouver ses capacités de croissance après avoir subit la contrainte que représente la saison sèche. Le génotype 1-41 présente des fluctuations plus importantes de l'indice de plasticité du ratio de biomasse aérienne/racinaire au cours des saisons que le génotype 18-50. Par contre, en condition non irriguée, la reprise des pluies n'induit pas une augmentation du ratio chez le génotype 1-41, confirmant que ce génotype doit avoir un temps de latence plus long que 18-50 pour récupérer ses capacités de croissance après la contrainte hydrique saisonnière. Dans ce cas, le délai de la réponse plastique est plus court pour le génotype 18-50.

Les différences de délai de réponse pourraient être expliquées par la mise en place de mécanismes spécifiques de protection des feuilles chez le 18-50. Ainsi, nous avons observé que la réduction des surfaces foliaires spécifiques pendant la saison sèche chez les arbres non-irrigués est plus importante chez 18-50. L'épaississement du collenchyme dans le traitement non-irrigué semblent également être spécifiques à 18-50, et pourrait avoir un rôle dans l'osmorégulation, pour limiter les effets délétères d'un chute de pH due à la diminution de l'eau dans l'arbre. Ces tendances sont confirmées par l'analyse du transcriptome: il semble que le génotype 1-41 modifie peu l'expression de ses transcrits, alors que le génotype 18-50

semble mettre en place un plus grand nombre de mécanismes pour se protéger de la contrainte hydrique. L'analyse du protéome confirme cette tendance: le génotype 1-41 présente des différences d'accumulation de protéines d'intensité plus faible que 18-50. La réduction de l'expression des gènes liés à la photosynthèse, au métabolisme secondaire, et l'augmentation de l'expression liée à la réorganisation cellulaire semblent ainsi être spécifiques au génotype 18-50. Au niveau du transcriptome d'apex, les gènes liés au photosystème semblent être particulièrement sous-exprimés chez 18-50 en condition non-irriguée. Parmi ceux-ci, les gènes *PGR5* et *ELIP* sont également sous-exprimés dans les feuilles matures lors de la contrainte hydrique. Enfin, l'augmentation du contenu en acides hydroxycinnamiques, en lien avec la sur-expression de gènes impliqués dans la voie de biosynthèse des phényl-propanoïdes et des terpenoïdes pourrait également permettre aux feuilles de mieux se protéger contre les espèces d'oxygènes activées formés lors du déficit hydrique.

Dans cette étude, le choix de génotypes très différents apparait assez pertinent, car il a permis de mettre en évidence un certain nombre de caractères présentant des effets d'interaction GxE. Les caractères tels que la surface foliaire spécifique, l'épaisseur de collenchyme et l'accumulation de phénols semblent être d'intéressant critères à mesurer sur une base génétique plus large, pour pouvoir estimer le potentiel adaptatif des populations. Etant donné que les génotypes utilisés proviennent d'espèces d'hybrides d'espèces différentes, il est encore difficile de savoir si ces différences de plasticité seront observables en intraspécifique ou au sein de différents hybrides de la même espèce. Cette étude aura ainsi permis de fournir de bonnes bases pour la description de la plasticité phénotypique et moléculaire à de plus grandes échelles, qui pourront aider à l'étude du déterminisme génétique de cette plasticité.

2.4 Conclusion

Cette étude a révélé deux stratégies d'acclimatation à la contrainte hydrique différentes chez deux hybrides d'espèces différentes. Le génotype le plus performant (18-50) semble mettre en place un grand nombre de mécanismes à la fois physiologiques et structuraux, qui lui permettent d'être plus réactif lorsque les pluies reprennent. La mise en place de mécanismes spécifiques pour répondre à la contrainte hydrique ont irrémédiablement un coût : le carbone fixé au cours de la photosynthèse n'est plus (ou à une moindre mesure) alloué à la croissance, mais dérouté pour la synthèse de composés intervenant dans la mise en place de tissus spécifiques tels que le collenchyme, ou dans la protection contre les espèces activées d'oxygène telles que les phénols. On peut donc penser que la stratégie adoptée par ce génotype est plus coûteuse au cours de la saison sèche, mais que c'est une stratégie payante

dans le cadre de la gamme environnementale testée. En effet, les feuilles ainsi protégées de la contrainte subie au cours de la saison sèche devraient encore être fonctionnelles quand les conditions environnementales redeviennent favorables. Ainsi, ce génotype pourrait alors être capable de réallouer directement le carbone fixé à la croissance. Par contre, à la reprise des pluies, le génotype le moins performant 1-41, ayant peu mis en place de mécanismes de protection des feuilles, devra renouveler le stock de feuilles endommagées lors de la saison sèche, donc allouer une plus grande partie à la biomasse foliaire plutôt qu'à la biomasse de bois. Cette hypothèse pourrait expliquer le retard de croissance et le temps de latence après la saison sèche observée sur les répartitions de biomasse.

3 Perspectives

3.1 Perspective relatives aux données générées dans cette étude

Le séquençage du transcriptome par la méthode haut-débit 454/Roche a permis de générer un catalogue de transcrits très large (chapitre 5). Néanmoins, comme nous l'avons vu au chapitre 6, il est difficile d'obtenir un assemblage optimal des différentes lectures, et la quantité de séquences générées mériterait d'être ré-assemblée, d'autant plus que le génome d'*Eucalyptus grandis* est en passe d'être finalisé. Dans le cadre du projet Tree For Joules[2], il est ainsi prévu de construire un nouvel unigène en alignant sur le génome d'*Eucalyptus grandis* l'ensemble des lectures obtenues dans le cadre de cette thèse et d'autres disponibles dans les bases de données, . Cela permettra d'analyser les variations d'abondances (au moins les variations de type présence/absence, car les banques ne sont pas toutes non-normalisées) en fonction d'une grande variété d'organes et de conditions.

Dans le chapitre 5, nous avons également soulevé le problème de validation des mesures d'abondances effectuées *in silico* par des mesures d'expression *via* RT-qPCR. Cette analyse montre que la technique haut-débit peut conduire à une approximation des niveaux d'expression des gènes et qu'une validation par RT-qPCR est nécessaire pour sélectionner des gènes d'intérêt. La sélection de gènes d'intérêt sur la base de leur profil d'expression devra donc passer par une recherche de séquences homologues au contig d'intérêt pour vérifier que son abondance mesurée *in silico* est robuste, et privilégier des transcrits montrant des forts

[2] Ce projet est inscrit dans le cadre d'une action quadrilatérale France – Allemagne – Espagne – Portugal (Appel d'offre ERA-PG/KBBE – ANR) avec 9 partenaires étrangers.

différentiels d'expression. Néanmoins, la connaissance des séquences et surtout du polymorphisme dans les génotypes étudiés permet de générer très rapidement des amorces utilisables en RT-qPCR, et finalement de valider relativement rapidement les gènes intéressants.

Au cours de cette thèse, certains caractères phénotypiques se sont révélés être clés pour la comparaison des potentiels de plasticité des génotypes en réponse à la contrainte hydrique. Par exemple, la surface spécifique foliaire, la discrimination isotopique du carbone, certains paramètres anatomiques tels que l'épaisseur du collenchyme, des vaisseaux du xylème ou de la cuticule foliaire, l'accumulation de phénols semblent être des caractères intéressants pour la caractérisation (voire la sélection) de génotypes adaptés à un environnement tel que celui du Congo. Au niveau moléculaire, nous avons mis en évidence une faible proportion de gènes différentiellement exprimés qui pourraient être à l'origine de telles variations. Il reste donc encore un intéressant travail alliant recherches bibliographiques et exploitation des résultats du pyroséquençage, pour cibler des gènes potentiellement liés à la variation des caractères d'intérêt. On pourrait particulièrement s'intéresser à des gènes liés avec la variation des caractères anatomiques. Par exemple, rechercher dans les banques de séquence des gènes impliqués dans la synthèse de pectine et de cellulose, et observer par hybridation *in situ* s'ils sont exprimés dans le collenchyme, le gène *POLG* (chapitre 4) constituerait un bon point de départ à une telle étude. Concernant la diminution du diamètre des vaisseaux du xylème, il devrait être possible de compléter l'analyse anatomique en mesurant les dépôts de lignine dans les cellules grâce à l'autofluorescence émise par ces composés. Si des différences peuvent être mises en évidence en fonction du traitement, les gènes de la voie de biosynthèse des lignines identifiés comme différentiellement exprimés pourraient également faire l'objet d'une analyse par hybridation *in situ*. Il serait également intéressant de compléter l'analyse de la voie de biosynthèse des phényl-propanoïdes en analysant l'expression des différents membres identifiés pour quelques familles.

Enfin, ce projet faisait partie d'une étude plus large faisant intervenir d'autres espèces comme le caféier et le pin maritime. La comparaison des résultats obtenus pour chaque espèce permettra éventuellement de mettre en évidence des mécanismes à la réponse à la contrainte hydrique communs à plusieurs espèces pérennes.

3.2 Perspectives relatives à la thématique de recherche

Pour la suite de cette étude, il sera important d'élargir la gamme environnementale et le pool génétique utilisé.

Comme nous l'avons vu précédemment, le traitement testé représente une levée de contrainte par rapport aux conditions naturelles subies en plantations industrielles. Pour savoir si la stratégie d'investissement dans des mécanismes de protection des feuilles déployé par le génotype 18-50 serait toujours avantageuse dans le cadre d'une contrainte hydrique plus importante, on pourrait imaginer appliquer une privation d'eau plus importante. Cela pourrait se faire au champ par une exclusion des eaux de pluies à l'aide de gouttières ou de films plastiques. De tels dispositifs ont été réalisés en Italie, en Espagne, au Portugal et en France pour étudier les effets d'une augmentation de la sécheresse sur le fonctionnement des écosystèmes méditerranéens (Limousin, 2009), mais également en forêt tropicale d'Amazonie (Nepstad et al., 2002). Il est également possible d'appliquer une contrainte plus importante en utilisant des plantes en pots, sous serre . Cette méthode, moins coûteuse et plus facile à mettre en place, permettra de contrôler un maximum de paramètres environnementaux. Néanmoins, étant donné les capacités exceptionnelles de croissances des eucalyptus, elle devra être cantonné à quelques mois de croissance et reflétera ainsi moins bien la réalité de la réponse en conditions naturelles.

Dans notre étude, les points de prélèvements effectués ont permis de caractériser les trajectoires de croissance cumulés sur des périodes allant de 2 à 4 mois. Sur de telle périodes, il n'a pas été possible d'identifier finement le délai de réponse à la contrainte hydrique, comme par exemple à partir de quel potentiel hydrique du sol les arbres arrêtent leur croissance. De la même façon que l'état hydrique du sol est mesuré en continu, une caractérisation plus régulière des trajectoires de croissance permettraient d'avoir une vision des relations "instantanées" entre la croissance et l'environnement. C'est l'objet d'un dispositif expérimental installé dans le cadre du projet ABIOGEN[3], qui se propose de mesurer en continu la croissance en circonférence par la mise en place de capteurs sur une descendance de cartographie en République du Congo.

Parmi les caractères identifiés pouvant être pris en compte dans les programmes d'amélioration parce qu'ils apportent un avantage en terme d'interaction de l'arbre avec son

[3] Projet financé par le conseil régional d'Aquitaine.

environnement, la surface spécifique foliaire et le composition isotopique en carbone semblent être non seulement des caractères intégratifs mais également les plus facilement mesurables à grand échelle. Il reste donc à évaluer si ces caractères sont variables, héritables et positivement corrélé avec les traits cibles de la sélection, et pour cela, augmenter le pool génétique à analyser.

Une trentaine de clones issus des dernières sélections du CRDPI ont déjà fait l'objet de mesures de caractères écophysiologiques au niveau foliaire tels que la surface spécifique en saison des pluies et en saison sèche, la concentration en nutriments, la discrimination isotopique du carbone foliaire (Maurice, 2010). Il ressort de cette étude que ces caractères semblent être relativement variables et ont une héritabilité au sens large élevée (supérieure à 0,5), ce qui constitue un bon pré-requis pour l'utilisation de ces caractères dans le programme d'amélioration. Chez le peuplier, Monclus et al. (2005) ont également montré que de tels caractères présentaient de fortes héritabilités au sens large. Ils suggèrent l'utilisation de la surface foliaire maximale et du contenu en azote foliaire comme critère de sélection chez le peuplier, car ces caractères sont fortement corrélés avec la productivité et la discrimination isotopique du carbone.

Afin de continuer à rechercher les déterminants génétiques de la plasticité phénotypique en réponse à une contrainte hydrique, on pourrait également mesurer un certain nombre de ces caractères sur la famille de plein-frères utilisée pour la détection de QTL. Ainsi une descendance des parents *E. urophylla* (génotype 14-144) et *E. grandis* (9-21) qui a servi à établir une première carte génétique (Gion et al., 2000) vient d'être mis en place sur un dispositif expérimental à Pointe Noire[4]. On pourrait alors rechercher des régions génomiques à effets majeurs (QTLs) liées à la variation de ces caractères, comparer les régions trouvées pour des mesures effectuées en saison sèche et en saison des pluies, et éventuellement identifier des colocalisations entre ces régions d'intérêt et des gènes candidats. L'étape suivante serait alors l'identification de polymorphismes de ces gènes candidats qui pourraient être liés à la variations des caractères d'intérêt, tel qu'il a été réalisé chez l'eucalyptus pour les gènes de la voie de biosynthèse des lignines et la composition en lignine (Mandrou, 2010).

L'avènement des nouvelles générations de séquençage devrait permettre de rendre ce type d'approches de plus en plus abordable. En effet, comme nous l'avons vu dans le chapitre

[4] Une thèse financée par le CIRAD vient de débuter sur le suivi en continue de la croissance de ces arbres.

introductif, les capacités en termes de reséquençage ne sont plus limitantes. Cela devrait permettre à la fois d'établir des cartes génétiques plus fines, de recenser le polymorphisme sur de plus grandes régions pour un nombre important de génotypes, pour pouvoir plus aisément mettre en évidence les mécanismes génétiques sous-jacent à la variation d'un caractère par clonage positionnel des QTL. De plus, ces techniques devraient également permettre d'étudier les phénomènes épigénétiques qui interviennent dans la régulation de l'expression des gènes.

REFERENCES BIBLIOGRAPHIQUES

Campion, J. M., M. Nkosana, et al. (2006). "Biomass and N and P pools in above- and below-ground components of an irrigated and fertilised Eucalyptus grandis stand in South Africa." Australian Forestry **69**(1): 48-57.

Chaves, M. M., J. P. Maroco, et al. (2003). "Understanding plant responses to drought - from genes to the whole plant." Functional Plant Biology **30**(3): 239-264.

Christina, M., J. P. Laclau, et al. (2011). "Almost symmetrical vertical growth rates above and below ground in one of the world's most productive forests." Ecosphere **2**(3): art27.

Drake, P. L., D. S. Mendham, et al. (2009). "A comparison of growth, photosynthetic capacity and water stress in Eucalyptus globulus coppice regrowth and seedlings during early development." Tree Physiology **29**(5): 663-674.

Gion, J. M., P. Rech, et al. (2000). "Mapping candidate genes in Eucalyptus with emphasis on lignification genes." Molecular Breeding **6**(5): 441-449.

Hacke, U. and J. J. Sauter (1996). "Drought-induced xylem dysfunction in petioles, branches, and roots of Populus balsamifera L and Alnus glutinosa (L) Gaertn." Plant Physiology **111**(2): 413-417.

Limousin, J.M. (2009) Vulnérabilité du chêne vert (*Quercus ilex* L.) à une augmentation de la sécheresse: Quels ajustements fonctionnels? Thèse de doctorat.

Mandrou, E. (2010) Variabilité fonctionnelle de gènes candidats de la lignification chez l'eucalyptus. Thèse de doctorat.

Maurice, D. (2010). Etude exploratoire de la variabilité génétique de morphologie foliaire, de l'état nutritionnel et hydrique chez l'eucalyptus : effet sur la croissance. Agrocampus Ouest. Diplôme d'ingénieur en agriculture.

Monclus, R., E. Dreyer, et al. (2005). "Productivity, leaf traits and carbon isotope discrimination in 29 Populus deltoides x P-nigra clones." New Phytologist **167**(1): 53-62.

Muchow, R. C. and T. R. Sinclair (1989). "Epidermal Conductance, Stomatal Density and Stomatal Size among Genotypes of Sorghum-Bicolor (L) Moench." Plant Cell and Environment **12**(4): 425-431.

Nepstad, D. C., P. Moutinho, et al. (2002). "The effects of partial throughfall exclusion on canopy processes, aboveground production, and biogeochemistry of an Amazon forest." Journal of Geophysical Research-Atmospheres **107**(D20).

Osorio, J., M. L. Osorio, et al. (1998). "Effects of water deficits on ¹³C discrimination and transpiration efficiency of Eucalyptus globulus clones." Functional Plant Biology **25**(6): 645-653.

Riederer, M. and L. Schreiber (2001). "Protecting against water loss: analysis of the barrier properties of plant cuticles." Journal of Experimental Botany **52**(363): 2023-2032.

Stape, J. L., D. Binkley, et al. (2010). "The Brazil Eucalyptus Potential Productivity Project: Influence of water, nutrients and stand uniformity on wood production." Forest Ecology and Management 259(9): 1684-1694.

Résumé :

Dans le contexte des changements climatiques, la capacité d'adaptation à la contrainte hydrique des arbres de plantation devient un enjeu majeur pour le maintien de leur productivité. La plasticité phénotypique des génotypes, facteur majeur de l'adaptation aux changements environnementaux, reste encore insuffisamment décrite chez les arbres forestiers, plantes pérennes à long cycle de révolution. Cette thèse se propose i / de caractériser la plasticité phénotypique pour deux clones commerciaux d'eucalyptus soumis à une contrainte hydrique au champ, ii/ d'identifier les caractères potentiellement adaptatifs (*i.e.* ceux présentant de l'interaction GxE) et iii/ de mettre en évidence les mécanismes moléculaires sous-jacents.

Un dispositif expérimental mis en place en république du Congo a permis de comparer, à travers une approche intégrant différents caractères allant de l'expression de gènes à la production de biomasse, la réponse de deux génotypes d'eucalyptus soumis à des régimes hydriques différents au cours de la saison sèche sur les deux 1ères années de croissance.

Les résultats montrent que si l'effet de la saison sèche est relativement similaire pour les deux génotypes (réduction des accroissements relatifs), le clone plus productif présente des accroissements très élevés en saison des pluies, et semble donc mieux optimiser des conditions environnementales favorables. La capacité d'ajustement physiologique, anatomique et moléculaire au niveau foliaire de ce génotype semble être un atout au maintien de ses capacités photosynthétiques.

Cette thèse a permis de mettre en évidence certains critères (surface foliaire spécifique, accumulation de phénols, épaisseur de collenchyme dans les feuilles), qui pourraient permettre d'évaluer le potentiel adaptatif des populations d'eucalyptus, nécessaire à leur gestion durable des plantations. D'autre part, certains gènes impliqués dans la photosynthèse et le métabolisme secondaire, dont l'expression pourrait être liée à la variation de caractères phénotypiques ont été mis en évidence. L'identification de ces gènes constitue une première étape vers la compréhension des mécanismes moléculaires sous-jacents à la plasticité phénotypique chez l'eucalyptus.

Abstract :

In the context of climate change, the ability of industrial forest plantation to cope with water scarcity is becoming a challenge for productivity maintenance. The phenotypic plasticity of genotypes, a major factor of adaptation to environmental changes, is still insufficiently described for long-lived species such as trees. This thesis proposes i/ to characterize the phenotypic plasticity for two commercial eucalyptus clones that differ in terms of productivity, subjected to water stress in the field, ii/ to identify potentially adaptive traits (i.e. those revealing GxE interaction) and iii/ to highlight the underlying molecular mechanisms.

A field trial installed in the Republic of Congo was used to compare, responses of two contrasted eucalyptus genotypes subjected to different watering regimes during the dry season on the first two years of growth. We used an integrative approach involving different traits from gene expression to biomass production.

The results show that if the effects of the dry season were relatively similar for both clones (lower relative increments), the most productive clone displayed a higher growth increment during the rainy season, and seemed to take more benefits when environmental conditions become more favorable. The ability to adjust leaf physiological, anatomical and molecular traits of this genotype seems to be an asset to maintain its photosynthetic capacity.

This thesis allowed to highlight some criteria (specific leaf area, carbon isotope discrimination, thickness of collenchyma and cuticle in leaves), which could help to assess the adaptive potential of *Eucalyptus* populations for sustainable management of planted stands. On the other hand, some genes whose expression may be related to variation in phenotypic characters were revealed. This set of genes is resource first step toward the understanding of the molecular mechanisms underlying phenotypic plasticity in eucalyptus.