



HAL
open science

Analyse génétique et écophysiologique de la tolérance à la sécheresse et au stress thermique chez le blé tendre (T. Aestivum L.)

Gaëtan Touzy

► **To cite this version:**

Gaëtan Touzy. Analyse génétique et écophysiologique de la tolérance à la sécheresse et au stress thermique chez le blé tendre (T. Aestivum L.). Génétique des plantes. Université Clermont Auvergne [2017-2020], 2019. Français. NNT : 2019CLFAC011 . tel-03270914

HAL Id: tel-03270914

<https://theses.hal.science/tel-03270914v1>

Submitted on 25 Jun 2021

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Clermont Auvergne

Années 2016-2019

Ecole Doctorale Sciences de la Vie, Santé, Agronomie, Environnement

THESE DE DOCTORAT

Présentée à l'Université Clermont Auvergne pour l'obtention du grade de

DOCTEUR D'UNIVERSITE

Spécialité : Biologie végétale

ANALYSE GENETIQUE ET ECOPHYSIOLOGIQUE DE LA TOLERANCE A LA SECHERESSE ET AU STRESS THERMIQUE CHEZ LE BLE TENDRE (*T. AESTIVUM* L.)

Soutenue le 7 mai 2019 par :

Gaëtan TOUZY

Directeur de thèse : Sébastien PRAUD

Co-directeur de thèse : Jacques LE GOUIS

Co-encadrante de thèse : Katia BEAUCHENE

Composition du jury :

Marie-Odile BANCAL	Rapporteuse
Jacques DAVID	Rapporteur
Delphine LUQUET	Rapporteuse
Anne LAPERCHE	Examinatrice
Laurence MOREAU	Examinatrice
Gilles CHARMET	Président
Jacques LE GOUIS	Co-directeur de thèse
Sébastien PRAUD	Directeur de thèse

Travaux réalisés dans le cadre d'une convention CIFRE (2015/0686), avec l'aide de L'ANRT

ARVALIS - Institut du végétal

Biopole Clermont Limagne, 63360 Saint-Beauzire, France.

BIOGEMMA

Centre de recherche de Chappes, Route d'Ennezat, Chappes, France

UMR 1095 INRA / UCA GDEC

5 chemin de Beaulieu, Clermont-Ferrand, France

Université Clermont Auvergne, Aubière, France

Cette thèse (2015/0686) a été financée par l'ANRT (Association Nationale de la Recherche et de la Technologie) et Arvalis Institut du Végétal. Les données génotypiques ont été obtenues grâce au financement du projet PIA (Projet d'Investissement d'Avenir) BreedWheat (ANR-10-BTBR-03). Les données de phénotypage ont aussi été obtenues grâce au financement du projet BreedWheat, mais aussi le support du projet Phénome (ANR-11-INBS-0012) projet financé par l'ANR, du fond de soutien FranceAgriMer, du Fond de Soutien à l'Obtention Végétale (FSOV) et de l'INRA.

Remerciements

Nous sommes en 2019 après J-C, toute la recherche végétale est occupée par le maïs... Toute ? non ! Une petite troupe d'irréductibles chercheurs tentent encore et toujours de sauver le *WheatSpirit* face à l'envahisseur. Et la vie n'est pas facile pour les garnisons de chercheurs des camps retranchés d'Arvalis-SVIG, de Biogemma-GGC et de l'INRA-GDEC.

Au milieu de ces camps tranchés et retranchés, le légionnaire Touzyx tient à remercier l'ensemble des acteurs de cette vaste aventure.

Les aventures de Touzyx chez Arvalis :

Merci à Katia Beauchêneis pour m'avoir suivi dans cette thèse et aiguillé pour aller à la recherche des compétences éparpillées dans les différents camps d'Arvalis, un merci à Matthieu Bogartis (le premier à m'avoir baptisé sous le nom Touzyx) et qui m'a apporté des critiques sincères (pas toujours écoutées) sur mes travaux, merci à Jean-Charles Deswartis et Benoît De Solanis, les physiologistes du camp de CHN. Enfin merci à David Gouacheis et Jean-Pierre Cohanis qui m'ont fait confiance pour attaquer ce grand projet sur la tolérance à la sécheresse.

Les aventures de Touzyx chez BiogemmaX :

C'est ici que j'ai fait mes armes et que j'ai lancé mes premières attaques, donc il faudrait plus d'une page pour vous remercier tous et toutes. Mais quand même pour l'histoire, il faut remercier tous les gens qui sont passés dans le camp de l'openSpace Blé. Tout d'abord, un grand merci à Madame HébertX et Mister ThroudeX (« -il est frais mon marqueur ; -non il n'est pas frais ; -il est frais... la bagarre ? »), il ne faut pas oublier LénaïcX avec sa moustache dans son aquarium et tous les X-men and X-girls qui sont passés bien sûr Fabien CormieX *alias* panoramix, Elise, Florian, Nicolas, Cécile, Anaïs, Adeline... Après l'openSpace Blé, un merci à l'openSpace Maïs dont l'effectif a bien bougé jusqu'à devenir vide aujourd'hui (Victoire) et merci quand même aux gens du MaïX, ClémenX, JérémyX, HerveX, JohanX, ArianX, NadineX. Merci à Pierre DubreuilX pour les prises de tête en statistiX et en autre X. D'ailleurs merci aux adeptes du 10h30 et de 16h30 pour les discussionX. Enfin, il faut aussi dire un merci à Sébastien PrauX (le abracourcix de l'équipe GGC) toujours en tête sur son bouclier pour attaquer les problèmes de la recherche et motiver les troupes. Un grand merci à Stéphane LafargeX la force tranquille et rassurante.

Les aventures de Touzyx chez l'INRAc

Un grand merci à Renaud Rincen^{ta}c surement futur traître au *WheatSpirit*, mais bon, sans qui cette thèse n'aurait pas été aussi aboutie (si elle l'est). Jacques Le Gouis^{ac} le César de la recherche sur la génétique du blé grand empereur de BreedWheat qui est toujours disponible. Et merci à l'ensemble des autres personnages de l'INRA.

Les aventures de Touzyx à Wageningen

Thank a lot to Fred van Eeuwijk for welcoming me in your lab. I also would like to thanks Willem Kruijer, Marcos Malosetti and Martin Boer who took times for explaining me how statistic workd or not... And special dedicas to Emilie Millet (51.965424, 5.662401)

Les aventures de Touzyx dans la science

Merci à Fred et Jamy parce que c'est pas sorcier et merci à deux ou trois profs quand même qui ont réussi à me captiver pour leur matière. Merci aux rapporteurs et rapportrices, ainsi qu'aux examinateurs et examinatrices pour avoir pris le temps de lire mes travaux et d'avoir apportés leurs critiques très constructives qui m'aideront sûrement dans la suite de mes aventures.

Les vraies aventures de Touzyx

Un grand merci à la Tranche-sur-cère, he oue ça tranche dans tous les coins et s'il y en a un peu trop ils vous le mettent quand même. He aussi à tous les sankitorpeurs. He aussi à tous les cantalous (qu'ils soient chasseurs de sangliers ou pas).

Enfin merci à ma mère et à mon père sans qui je ne serais pas là. He pardis, merci à la famille, dont les Roumega Louisa, Marius et les parents pardis.

Et surtout ne jamais oublier que la Mini, et les Marguerites, sont les plus belles.

PUBLICATIONS

Ly D., Huet S., Gauffreteau A., Rincant R., **Touzy G.**, Mini A., Jannink JL, Cormier F., Lafarge S., Le Gouis J. and Charmet G., Genomic random regression for adaptation breeding: Whole-genome prediction of the reaction norms to environmental stress in bread wheat (*Triticum aestivum* L.). **Field Crops Research** 12/2016

Rincant R., Malosetti M., Ababaei B., **Touzy G.**, Mini A., Bogard M., Martre P., Le Gouis J., van Eeuwijk F., Using crop growth model stress covariates and AMMI decomposition to better predict genotype by environment interactions, Expected submission: 2019

Touzy G., Rincant R., Bogard M., Lafarge S., Dubreuil P., Mini A., Deswarte J.-C., Beauchêne K., Le Gouis J., 2018, Praud S., Identifying drought resistance specific QTLs in bread wheat (*T. aestivum* L.) using environmental clustering. Submitted to **Theory Applied Genetics** 10/2018 accepted on 05/2019

Touzy G., Le Gouis J., Lievin V., Xavier D., Redondo E., Praud S., Lafarge S., 2019, A genome-wide identification of QTLs by heat stress interaction in wheat (*T. aestivum* L.), Expected submission: 2019

PRESENTATION

Communications orales :

Touzy G., van Eeuwijk F., Rincant R., Kruijjer W., Lafarge S., Beauchene K., Le Gouis J., Praud S. and The BreedWheat Consortium. Improving water use efficiency in bread wheat by multi-trait multi-environment genome –wide association studies. EUCARPIA cereal section, 19-22 March 2018 Clermont-Ferrand, France

Posters :

Touzy G., Lafarge S, Beauchene K, Rincant R, Praud S, Le Gouis J., Phenotypic characterization of French elite germplasm in response to drought and heat stress. 13th IGWS (International Wheat Genetics Symposium), 24-28 April 2017, Tulln, Autriche

Touzy G., van Eeuwijk F., Rincant R., Kruijjer W., Lafarge S., Beauchene K., Le Gouis J., Praud S. and The BreedWheat Consortium. Genome-Wide Association Study of Winter Bread Wheat (*Triticum aestivum* L.) in Response to Drought in a Multi-Environment European Network. Plant and Animal Genome XXVI Conference, 13-17 January 2018 San Diego, CA, Etats-Unis

Table des matières

Introduction Générale

Contexte de l'étude.....	1
Positionnement de la thèse dans ce contexte.....	3

Chapitre I : Synthèse bibliographique

1. L'importance du blé dans le monde.....	6
1.1. Le blé : <i>Triticum aestivum</i> L.....	6
1.2. Production et importance économique dans le monde.....	7
1.3. Importance nutritionnelle et utilisation industrielle du blé tendre.....	8
2. Les stades de développement du blé et la mise en place du rendement.....	9
2.1. L'échelle de Zadoks pour définir les stades de développement du blé.....	9
2.2. Le rendement se met en place tout au long du cycle du blé.....	10
2.2.1. La densité de plantes au mètre carré.....	11
2.2.2. Le nombre d'épis par plante.....	11
2.2.3. Le nombre de grains par épis.....	12
2.2.4. Le poids des grains ou PMG (poids de mille grains).....	13
2.3. Le développement racinaire.....	14
3. Du concept de stress à l'impact de ceux-ci sur le développement des plantes.....	15
3.1. Impact des stress thermiques.....	16
3.1.1. Impact du stress thermique sur la croissance végétative.....	17
3.1.2. Impact du stress thermique sur le développement du grain de blé.....	18
3.2. Impact des stress hydriques.....	20
3.2.1. Description des flux d'eau.....	20
3.2.2. Caractérisation des flux d'eau au cours du cycle d'une culture grâce au modèle de culture CHN.....	23
3.2.3. Impact du stress hydrique sur la physiologie de la plante.....	28
4. Les différences d'adaptation du blé en réponse aux stress hydriques et thermiques.....	30
4.1. Les concepts de traits physiologiques associés à l'adaptation des stress hydriques et thermiques.....	30
4.1.1. Modèle de mise en place du rendement en condition stressante présenté par Passioura, (1977).....	31
4.1.2. Description des traits les plus connus permettant une adaptation aux stress hydriques et thermiques.....	32
4.2. Méthode d'analyse de réseau d'essai pour sélectionner les variétés en lien avec un stress.....	36
4.2.1. Le problème des interactions génotype x environnement (GxE).....	36
4.2.2. L'importance du GxE en expérimentation variétale.....	37
4.2.3. Comment analyser et modéliser le GxE.....	38
4.2.4. Caractérisation de l'environnement pour expliquer le GxE.....	39
4.2.5. L'utilisation des modèles mixtes pour modéliser les interactions GxE.....	41

5. Les composantes de variabilité génétique en réponse aux stress et méthode de détection	42
5.1. Les gènes majeurs chez le blé	43
5.2. La détection de nouveaux gènes et/ou allèles avec la GWAS	44
5.2.1. La structuration des panels	46
5.2.2. Choix du niveau de significativité	48
5.3. Dissection des composantes génétiques de l'interaction génétique par environnement	49
5.4. Les QTL de tolérance à la sécheresse	50

Chapitre II : Le panel BREEDWHEAT

1. Analyse des données moléculaires	53
2. Etude du déséquilibre de liaison pour définir des QTL	54
3. Analyse de la structuration et de l'apparentement du panel élite	55
La caractérisation du panel élite face aux stress hydriques et thermiques	57

Chapitre III : Tolérance aux stress hydriques

Using environmental clustering to identify specific drought resistant QTLs in bread wheat (*T. aestivum* L.)

1. INTRODUCTION	61
2. MATERIALS AND METHODS	63
2.1. Plant materials and field experiments	63
2.2. Drought stress pattern characterization	64
2.2.1. WSI simulation and clustering of the multi-environment trial	64
2.2.2. WSI simulation and frequency analysis at the scale of France	65
2.3. Phenotypic Traits	65
2.4. Statistical analyses	66
2.4.1. The two-Stage GxE analysis	66
2.4.2. Statistical analyses of the MET	66
2.4.3. Heritabilities	67
2.4.4. Genotypic value extraction from Environment Types	67
2.5. Genetic progress	68
2.6. GWAS analysis	68
3. RESULTS	70
3.1. Environmental clustering identified four types of drought pattern in the MET	70
3.2. Pattern and frequencies of drought scenarios in France	70
3.3. Environment types helped to understand grain yield structure and GxE interactions	70
3.4. Relationship between agronomic traits varies within environment types	71
3.5. Genetic gains for grain yield varied within environment types	72
3.6. GWAS results	72

3.7. Colocalization network	72
4. DISCUSSION	73
4.1. Environment types improved understanding of grain yield structure and GxE interaction	73
4.2. Genetic progress between 1992 and 2011	75
4.3. GWAS	76
CONCLUSION	79

Chapitre IV : Tolérance aux stress thermiques

Genome-wide identification of QTLs affecting terminal heat stress responses in wheat (*Triticum aestivum* L.)

1. INTRODUCTION	82
2. MATERIALS AND METHODS	84
2.1. Plant material and growth conditions	84
2.2. Split-plot design	85
2.3. Heat treatment	85
2.4. Measurement of yield components	86
2.5. Measurement of chlorophyll content and estimation of senescence traits	86
2.6. Statistical analysis	88
3. RESULTS	90
3.1. Phenotypic analysis and variance decomposition	90
3.2. Multi-year phenotypic analysis and variance decomposition	91
3.3. An overview of senescence traits	92
3.4. Genetic correlations	92
3.5. GWAS	92
4. DISCUSSION	94
4.1. Experiment validation and stress impact	94
4.2. GWAS	95
4.2.1. QTL associated with traits	95
4.2.2. QTLs associated with a response to heat stress	97
CONCLUSION	98

Chapitre V : Prospection multi-stress

Prospection des effets alléliques en fonction de facteurs de stress dans un réseau d'essais multi-environnemental

1. INTRODUCTION	101
2. MATERIEL ET METHODES	102
2.1. Matériel végétal et expérimentation	102
2.2. Données moléculaires	103
2.3. Description des analyses multi-environnementales	104
2.4. Calcul des covariables environnementales	105

2.5. Sélection des covariables	106
2.6. Détection des SNP associés aux différents traits phénotypiques dans le réseau multi-environnemental	107
2.6.1. GWAS classique par trait et par environnement	107
2.6.2. GWAS multi-environnements par locus et par trait	107
2.6.3. Multi-environnements multi-locus et par trait	108
3. RESULTATS	109
3.1. Description du MET	109
3.2. Sélection des covariables environnementales	110
3.3. GWAS et identification de QTLxE	111
4. DISCUSSION	112
4.1. Description du MET	112
4.2. Sélection des covariables environnementales	113
4.3. Identification des QTLxE et relation avec les covariables environnementales	114
CONCLUSION	117

Discussion Générale et Perspectives

Comment le matériel qui est à notre disposition peut-il répondre à nos objectifs ?	119
Les ressources moléculaires sont-elles suffisantes ?	120
Les résultats sont-ils en accord avec les objectifs qui avaient été déterminés ?	121
Le stress hydrique	121
Le stress thermique	123
La prospection multi-stress	125
Réponses aux objectifs	126
Comment répondre aux objectifs qui n'ont pas été atteints ?	127

FIGURES

Figure 1 : Evolution du rendement en France de 1961 à 2017 (FAO, 2018)

Figure 2 : Les prédictions de changement climatique du fait du réchauffement climatique, adapté de Pachauri et al., (2015)

Figure I.1 : Les différentes ploïdies chez les Triticum, d'après Bonjean, (2001)

Figure I.2 : Répartition des productions moyennes entre 2008 et 2017, en millions de tonnes (FAO 2018)

Figure I.3 : Les différents stades de développement du blé et l'initiation des différentes composantes du rendement

Figure I.4 : Développement du grain en nombre de jours après floraison et coupe longitudinale d'un grain mature

Figure I.5 : Impact d'un stress environnemental sur les plantes, d'après Gaspar et al., (2002)

Figure I.6 : Schéma de l'impact possible d'un stress thermique sur la photosynthèse, d'après Farooq et al., (2011)

Figure I.7 : Les échanges gazeux entre la plante et l'atmosphère, figure adaptée de Taiz and Zeiger, (2002)

Figure I.8 : Triangle des textures de la carte des sols du département de l'Aisne (France) et réserve utile maximum associée (en mm d'eau pour 1m de sol), d'après Jamagne, (1969) et Baize and Jabio, (1995)

Figure I.9 : Les différents réservoirs d'eau d'un sol, adapté de Bouffier, (2014)

Figure I.10 : Synoptique générale du modèle CHN, d'après Le Bris et al., (2016)

Figure I.11 : Variables d'entrées et de sorties du modèle CHN, d'après Le Bris et al., (2016)

Figure I.12 : Répartition classique des assimilats envers les différents organes pendant le cycle des céréales, par Savary and Willocquet, (2014)

Figure I.13 : Représentation des différentes formes des relations entre Facteur de stress hydrique et l'eau dans le sol d'après les équations présentés par Sinclair (1986)

Figure I.14 : Répartition des productions et des méga-environnements définis par le CIMMYT d'après Sonder, (2018)

Figure I.15 : Impact d'un stress environnemental sur les plantes, d'après Gaspar et al., (2002)

Figure I.16 : Modèle conceptuel représentant les traits physiologiques associés à l'adaptation des plantes à un environnement stressé hydriquement et/ou thermiquement, d'après Reynolds et al. (2017)

Figure I.17 : Illustration de l'effet génotypique (G), de l'environnement (E) et de l'interaction Génotype x Environnement (GxE), d'après van Eeuwijk et al., (2016)

Figure I.18 : Illustration du calcul de l'effet génotypique (G), de l'environnement (E) et de l'interaction Génotype x Environnement (GxE)

Figure I.19 : Les différents stades de développement du blé et l'initiation des différentes composantes du rendement

Figure I.21 : Différents scénarios de LD, d'après Flint-Garcia, (2003)

Figure I.22 : Représentation possible du LD le long du chromosomes 5A, d'après Breseghello, (2005)

Figure II.1 : Répartition de différents clusters de catégories de qualité des SNP de TaBW280K spécifique au panel élite

Figure II.2 : Représentation des polymorphismes aux marqueurs et données manquantes pour l'ensemble des SNP 'PolyHighResolution' du panel élite

Figure II.3 : Visualisation des recombinaisons le long des chromosomes

Figure II.4 : Structuration et apparentement dans le panel élite

Figure III.1 : Environment types dendrogram based on water stress index patterns

Figure III.2 : Environment types water stress index throughout the crop cycle

Figure III.3 : Dendrogram based on the water stress index (WSI) patterns simulated for more than 40000 scenarios occurring in France plus ETs patterns

Figure III.4 : Biplots based on agronomic traits and earliness for each environment type

Figure III.5 : Boxplot of Grain Yield performance corrected for earliness as a function of registration years of 205 wheat lines grown in the four ETs

Figure III.6 : Heatmap of the standardized effects of the 53 QTLs of grain yield in the four ETs

Figure IV.1 : Example of the fitted logistic curve and the estimation of stay-green traits

Figure IV.2 : Boxplot of TKW for 199 wheat cultivars grown over two years (2016 and 2017) with two heat stress treatments

Figure IV.3 : Change in senescence kinetics over time presented for all genotypes in the two heat-treatments

Figure IV.4 : Biplots based on a subset of traits measured on 199 wheat varieties grown in two heat treatment situations

Figure IV.5 : Manhattan plot of GWAS conducted on the SNP main effect and the SNP by treatment interaction for D.Z65 and TKW

Figure IV.6 : gBLUP values for SNP main effect and SNP by treatment interaction effect

Figure V.1 : Présentation de la sélection des environnements d'après la classification d'essais réalisée dans Touzy et al. (en prep.)

Figure V.2 : Distribution des covariables environnementales et clusterisation du MET

Figure V.3 : Corrélation phénotypique et génétique entre environnement

Figure V.4 : Corrélation entre la matrice de VarCov Σ_G et la matrice de distance Σ_{EC}

Figure V.5 : Visualisation de l'effet des SNP en fonction de l'évolution de covariables environnementales dans le MET

Présentation des fréquences de stress hydrique en France

TABLES

Table I.1 : Production et utilisation commerciale du blé dans le monde ces dix dernières années

Table I.2 : Seuils de température nécessaires pour le développement des plantes, d'après Porter et Gawith, (1999)

Table I.3 : Tableau récapitulatif de l'effet des traits sous différents stress hydriques, d'après Tardieu (2012)

Table I.4 : Récapitulatif de l'ensemble des publications où des QTL sont associés à des phénotypes dans des environnements optimaux, en stress hydrique et/ou thermique

Table II.1 : Répartition et recouvrement des marqueurs sur la carte physique (Appels et al., 2018) et génétique (Rimbert et al., 2018)

Table III.1 : Statistical indicators of experiment quality and the genetic variability of grain yield (t/ha) over the trial network

Table III.2 : Variance components of the different models for grain yield and its components

Table III.3 : Variance components within environment types

Table III.4 : QTL detected on the wheat association panel for 6 traits in the 4 environment types

Table III.5 : Number of common QTLs between two traits

Table IV.1 : Variance components from the multi-year analysis

Table IV.2 : List of associated SNP following a GWAS conducted on 199 wheat varieties grown for two years with two post-flowering temperature treatments: SNP with no interaction effect. Position is the position on Chinese Spring refSeq1.0 (IWGSC et al 2018)

Table IV.3 : List of associated SNP following a GWAS conducted on 199 wheat varieties grown for two years with two post-flowering temperature treatments: SNP with a significant interaction effect. "Position" is the position on Chinese Spring refSeq1.0 (IWGSC et al 2018)

Table V.1 : Indicateur statistique de la qualité des expérimentations et variabilité du rendement en grain dans le MET

Table V.2 : Les différentes formes de la matrice Σ_G

Table V.3 : Comparaison des matrices des BIC pour choix des variance/covariances

Table V.4 : Description des SNP sélectionnés dans le modèle multi-environnement multi-locus (Mod [4]) pour le rendement et relation avec les covariables environnementales

Sigles et Abréviations

°C	Degré Celsius	HSP	Heat Shock Proteins
qtx	quintaux	HWD	High Water Deficit
qtx.ha-1	quintaux par hectare	INN	Indice de Nutrition azotée
%	Pourcent	IWGSC	International Wheat Genome Sequencing Consortium
Gb	Giga base	LAI	Leaf Area Index
AIC	Akaike Information Criterion	LD	Linkage Disequilibrium
AMMI	Additive Main effects and Multiplicative Interaction	LI	Light Interception
ANOVA	Analysis of Variance	LRT	Likelihood Ratio Test
BAF	Blé Améliorant de Force	LWD	Late Water Deficit
BAU	Blé Autres Usages	MAF	Minor Allele Frequency
BB	Blé Biscuitier	ME	Méga-Environnements
BIC	Bayesian Information Criterion	MET	Multi-Environment Trials
BLUE	Best Linear Unbiased Estimator	MWD	Medium Water Deficit
BLUP	Best Linear Unbiased Predictor	NDVI	Normalized Difference Vegetation Index
BP	Blé Panifiable	OPT	Optimal Conditions
BPS	Blé Panifiable Supérieur	PAR	Photosynthetically Active Radiation
CHL	Chlorophyll content	PAV	Presence Absence Variation
CIMMYT	International Maize and Wheat Improvement Center	PCA	Principal Component Analysis
cM	CentiMorgan	PH	Plant Height
CNV	Copy Number Variation	PIC	Polymorphism Index Content
CSPA	Continuum Sol/Plante/Atmosphère	PLS	Partial Least Square
ET	Environment Types	QTL	Quantitative Trait Loci
ETP	Evaporation potentielle	QTN	Quantitative Trait Nucleotide
FA	Factorial Analysis	REML	Restricted Maximum Likelihood
FaPAR	Fraction of Absorbed Photosynthetically Active Radiation	RFU	Réserve d'eau Facilement Utilisable
FDR	False Discovery Rate	RG	Rayonnement Global
FiPAR	Fraction of Intercepted Photosynthetically Active Radiation	RU	Réserve Utile
FTSW	Fraction of Transpirable Soil Water	RUE	Radiation Use Efficiency
GPS	Number of Grains Per Spike	SA	Spikes per Area
GWAS	Genome Wide Association Study	SNP	Single Nucleotide Polymorphisme
GxE	Interaction Génotype-Environnement	TKW	Thousand Kernel Weight
GY	Grain Yield	TPE	Target Population of Environments
HCC	Humidité à la Capacité au Champ	WSI	Supply/demand ratio for Water
HD	Heading Date	WU	Water Absorption
HI	Harvest Index	WUE	Water Use Efficiency
HPFP	Humidité au Point de Flétrissement Permanent		

Introduction Générale

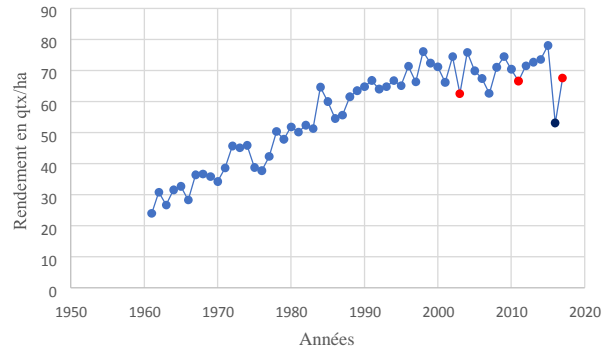
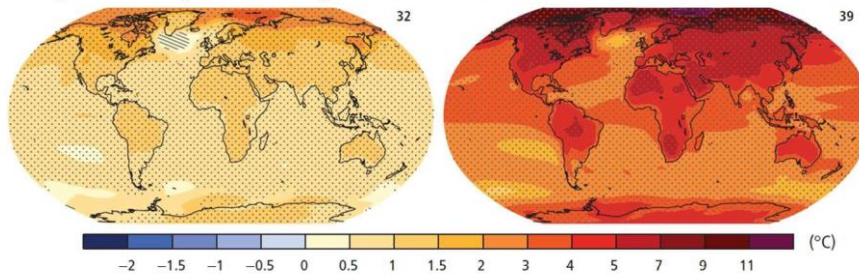


Figure 1 : Evolution du rendement en France de 1961 à 2017 (FAO, 2018)

Changement moyen des températures de surface entre 1986-2005 et 2081-2100



Changement moyen des précipitations de surface entre 1986-2005 et 2081-2100

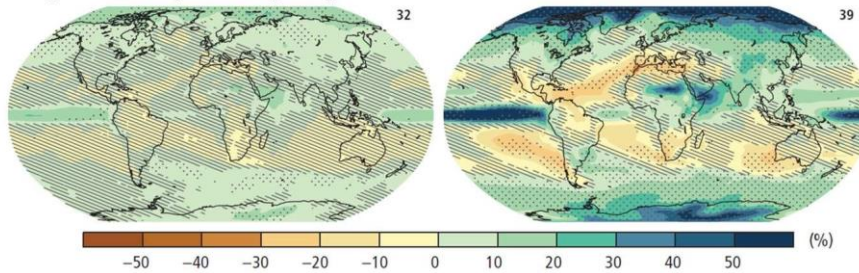


Figure 2 : Les prédictions de changement climatique du fait du réchauffement climatique, adapté de Pachauri et al., (2015)

Introduction Générale

Contexte de l'étude

Depuis les années 50 jusqu'au milieu des années 90, la production française de blé a suivi une augmentation régulière d'environ 1.18 qtx par hectare et par année, passant d'un peu plus de 20 qtx par hectare à plus de 70, aujourd'hui (Brisson et al., 2010). Cette progression peut être attribuée à trois principaux facteurs : (1) le perfectionnement des outils de travail et notamment des machines agricoles, (2) une meilleure utilisation des produits de synthèses (pesticides, régulateurs ou fertilisants, par exemple le fractionnement des apports azotés), et (3) la mise sur le marché de variétés plus performantes, (Brisson et al., 2010). Cependant, nous pouvons constater un ralentissement de l'augmentation des rendements en blé depuis le milieu des années 90, comme nous le montre la [Figure 1](#). Ce ralentissement, voire cette absence de progression, a été observé dans de nombreux pays occidentaux (Calderini et Slafer, 1998; Cassman, 1999; Finger, 2010; Peltonen-Sainio et al., 2009), et ce, même si les rendements mondiaux continuent d'augmenter (Lobell et al., 2011). En 2010, Brisson et al. indiquent deux pistes majeures pour expliquer ce ralentissement : la désintensification des systèmes de production (réduction de 15-20% des apports azotés, des fongicides, moindre travail du sol...) et l'impact d'évènements climatiques limitant la croissance des plantes. Par exemple, sur la [Figure 1](#), les trois points rouges représentent trois années où les rendements ont été plus faibles à cause d'évènements de sécheresse. A l'opposé, la récolte de 2016 (point bleu foncé) a été historiquement basse en raison d'un surplus de pluie et un manque de radiations lumineuses pendant des phases critiques d'établissement du rendement, notamment pendant la floraison. Ceci est compatible avec les modèles qui prédisent à la fois une augmentation globale des températures et une augmentation de la fréquence (ou de l'intensité) de la plupart des phénomènes extrêmes (Ben-Ari et al., 2018).

De nombreuses études montrent l'impact des variations de température et des vagues de sécheresse sur l'absence de progression des rendements chez le blé (Brisson et al., 2010). Les conditions hydriques pendant la montaison et les fortes chaleurs pendant le remplissage des grains expliquent une grande partie des variations interannuelles du rendement observées dans la plupart des régions de France (Gouache et al., 2012, 2015). Il a été estimé qu'en moyenne, le blé perd entre 1.5 et 2 qtx par hectare à chaque fois que le déficit hydrique s'accroît de 10 mm (Gate, 1995). Or, le projet ANR CLIMATOR a évalué un accroissement probable des déficits hydriques en blé tendre de 10 à 20 mm d'ici le milieu du siècle, et de 50 à 70 mm d'ici la fin du siècle (Bancal et al., 2010). De plus, Liu et al., (2016) ont estimé une perte de rendement entre 4.1 et 6.4% pour chaque degré d'augmentation de température, alors que les prédictions les plus optimistes, qui ont été réalisées par le GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat (Pachauri et al., 2015)), font part d'une élévation des températures globales de plus de 2°C avant la fin du siècle ([Figure 2](#)). Ceci alors que la consommation de blé devrait augmenter d'environ 30% dans les quarante prochaines années (Weigand, 2011).

Les stress hydriques et thermiques sont souvent concomitants, mais les stratégies d'adaptation à l'un ou l'autre de ces stress peuvent être différentes voir complètement opposées (Tardieu, 2012). Si l'investissement sur la recherche des adaptations aux stress a été jusqu'à maintenant moins marqué en Europe, du fait d'une présence moindre de ces types de stress, des avancées importantes sur le plan international ont été faites sur la compréhension physiologique de la réponse du blé aux stress hydriques et thermiques en comparant différents géotypes sensibles et résistants. Des différences au niveau physiologique, métabolique, transcriptomique ont été étudiées afin d'émettre des hypothèses sur les meilleures stratégies d'adaptation. Des approches de génétiques quantitatives ont également été menées, permettant l'identification de différents QTL (quantitative trait loci). Malgré ces nombreux résultats et leurs applications en amélioration variétale, comme par exemple la stratégie de conservation de l'eau qui s'est avérée judicieuse dans certains programmes australiens et qui a mené à l'obtention de variétés plus productives en situations de fortes sécheresses (Condon, 2004), leur transposition en Europe reste limitée voire difficile en raison (i) de la nature des stress observés sous nos latitudes qui ne correspondent pas vraiment aux environnements étudiés (Ober et al., 2010), (ii) et de la spécificité du matériel étudié, souvent du matériel de printemps non adapté à nos marchés.

Jusqu'à maintenant, l'amélioration génétique européenne n'a pas été spécifiquement orientée vers la sélection de variétés tolérantes à la sécheresse, mais plutôt vers une augmentation de la régularité des performances. Cette sélection de variétés est dite sélection indirecte, car elle se base sur l'estimation du rendement en grains dans de nombreux environnements. Si les événements de sécheresse sont de plus en plus fréquents, un processus de sélection dite directe de variétés tolérantes peut être envisagé, mais pour cela il est important de multiplier nos connaissances sur le sujet et notamment développer nos connaissances sur du matériel disponible pour la sélection européenne. Nous devons donc orienter la recherche vers une compréhension des mécanismes de résistance et de sensibilité des variétés à ces types de stress, ainsi qu'à une identification de matériel pertinent pour la sélection et d'outils moléculaires pour le suivi des allèles favorables dans le germoplasme.

La création de nouveaux cultivars ayant de bonnes performances et un niveau élevé de tolérance à un déficit hydrique et/ou de hautes températures peut être accélérée par une meilleure compréhension des facteurs génétiques impliqués dans des traits adaptatifs. Différents traits physiologiques sont susceptibles d'être associés avec un certain niveau de tolérance : sénescence tardive, efficacité de la photosynthèse sous stress, architecture racinaire, comportement stomatique, croissance expansive du feuillage... L'utilisation conjointe de matériel génétique bien caractérisé au niveau moléculaire et de nouvelles méthodes de phénotypage et d'analyse est le point d'entrée pour comprendre et exploiter ces traits. Cette thèse s'inscrit dans ce cadre précis.

Positionnement de la thèse dans ce contexte

Pour répondre à une partie des objectifs présentés précédemment, cette thèse est à la jonction entre recherche privée et publique : une partie du financement a été faite par l'ANRT (Association Nationale de la Recherche et de la Technologie) et l'autre par Arvalis Institut du Végétal, l'institut technique dédié aux céréales et dont les expertises en phénotypage et en modélisation sont reconnues. Cette thèse a été hébergée au sein de Biogemma, dont les travaux de recherche en blé étaient financés par Arvalis et par Limagrain, coopérative agricole et 4^e semencier mondial, porteur d'expertises complémentaires dans les domaines de la génétique et de la sélection. En outre, la thèse a été co-encadrée par une équipe de l'INRA de Crouel, dont les projets sont focalisés sur le blé tendre (UMR GDEC) et notamment la tolérance aux stress abiotiques. Pour finir, la thèse a pu bénéficier de la dynamique et de nombreux résultats obtenus dans le projet d'investissement d'avenir BREEDWHEAT, un partenariat public/privé dans lequel les trois parties prenantes de la thèse sont associées.

Les ambitions initiales de cette thèse étaient d'explorer la variabilité génétique naturelle pour la sensibilité au stress hydrique et thermique dans un contexte Européen de stress, au travers de la dissection de ces caractères complexes en caractères sous-jacents. Lors de l'élaboration du sujet, différents axes de recherche avaient été envisagés et proposés en fonction des jeux de données qui étaient disponibles ou qui étaient prévus d'être générés au cours des trois ans de la thèse ainsi que des hypothèses faites alors. Voici la liste de ces axes d'intérêt : (1) une étude des interactions génotype x environnement (GxE) au champ ; (2) une étude génétique du système racinaire ; (3) une étude génétique de l'architecture végétative de la plante ; (4) la mise en place de méthodes d'associations multi-traits ; (5) une approche intégrative de la tolérance chez le blé tendre.

Au final, au cours de ces trois ans de thèse, la stratégie a évolué, confrontée au pragmatisme de la production de données, (les données de phénotypage racinaire produites dans le projet FSOV Archirac n'ont finalement été accessibles que depuis l'automne 2018) et au pragmatisme du travail de thèse (certaines priorités ont été faites au fur et à mesure des analyses et des questions rencontrées et discutées lors des comités de thèse notamment) et certains axes ont été abandonnés dans ce travail, comme l'aspect racinaire, au bénéfice de l'étude du stress thermique et de l'analyse multi environnements.

Nous avons choisi de présenter cette thèse suivant quatre axes majeurs :

- (1) Le premier et le deuxième chapitre ont été rédigés dans le but de présenter l'état de l'art sur la tolérance aux stress hydriques et thermiques, ainsi que les outils à notre disposition pour répondre aux problématiques posées.

- (2) Dans le chapitre III, nous avons analysé de façon indépendante l'impact des stress hydriques sur une collection de blés élités européens, tout en plaçant les événements de stress dans un contexte français.
- (3) Le chapitre IV, s'intéresse quant à lui à l'impact d'un stress thermique post-floraison sur le remplissage des grains.
- (4) Enfin, dans le cinquième chapitre nous avons choisi d'analyser de façon intégrative l'ensemble des stress abiotiques subis par les plantes, et ceci à l'aide d'un réseau d'essais qui a été préalablement sélectionné pour maximiser les chances d'identifier des contraintes hydriques et thermiques affectant la réponse des plantes.

Pour finir, les résultats de ces travaux seront discutés en fin de manuscrit pour tenter d'apporter des éléments aux problématiques de recherche qui ont été définies. Aussi, nous confronterons les résultats qui étaient attendus lors de la rédaction du projet avec ceux présentés dans cette thèse, ou qui ont été présentés à la communauté scientifique *via* posters et présentations orales lors de congrès. Nous pensons que cette thèse contribuera à enrichir les connaissances à la fois sur la caractérisation des environnements stressés en France, sur la variabilité génétique existante et sur le déterminisme génétique (QTL) des caractères étudiés. De plus, nous espérons que ce manuscrit constituera un bon appui pour confronter et/ou valider les résultats d'actuels ou de futurs projets sur la tolérance de variétés européennes aux stress thermiques et hydriques.

Chapitre I : Synthèse Bibliographique

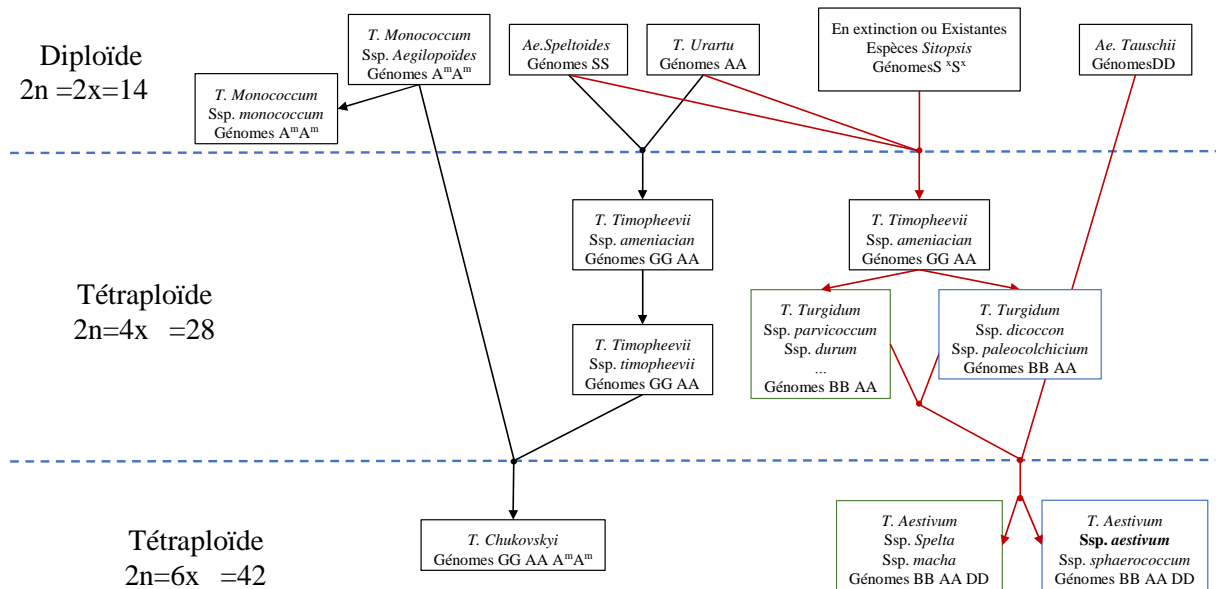


Figure I.1 : Les différentes ploïdies chez les *Triticum*, d'après Bonjean, (2001)

Chapitre I : Synthèse bibliographique

1 L'importance du blé dans le monde

1.1 Le blé : *Triticum aestivum* L.

Le "blé" est un terme générique qui désigne aussi bien la plante que la graine qui est produite par celle-ci. Du Moyen-Orient à l'Occident, en passant par le Maghreb, il est connu de tous car il est à la base de l'alimentation de ces régions. Toutefois, il existe plusieurs sortes de "blé", ils font tous partie de la famille des *Poacea* (anciennement les *Graminea*). Suivant les différentes classifications botaniques, cette famille peut être déclinée soit en une seule tribu, les *Triticum*, ou bien en deux tribus, *Triticum* et *Aegilops*. Le nom de blé est associé avec différentes espèces à l'intérieur du genre *Triticum* (Auriau et al., 1992), par exemple, le blé tendre (ou froment) "*Triticum aestivum* L." qui est le plus utilisé et qui permet de faire de la farine et du pain ; le blé dur "*Triticum durum*" qui est à l'origine des pâtes et des semoules de couscous, l'épeautre "*Triticum monococcum*" ou "*Triticum aestivum* ssp. *spelta*", l'engrain ou petit épeautre, l'amidonniér, etc... Dans l'ensemble du manuscrit, et de façon générale, le terme de "blé" est utilisé pour désigner le blé tendre. Il représente 95% de toutes les espèces de blé cultivées dans le monde (Shewry, 2009) et il est l'objet de notre étude.

Le blé a été domestiqué dans le croissant fertile (un des principaux centres d'origine de l'agriculture), berceau des sept cultures fondatrices de cette période (engrain, amidonniér, orge, lentille, pois, vesce et pois chiche). C'est dans cette région que les ancêtres du blé ont été retrouvés, ce qui a permis de retracer son histoire phylogénétique. Une première façon de différencier les groupes auxquels appartiennent les blés est de se référer à leur ploïdie, c'est-à-dire le nombre de copies de chromosomes contenues dans le noyau. Ainsi, pour un blé diploïde, le nombre de chromosomes est de $2n = 14$, pour un blé tétraploïde le nombre de chromosomes est de $4n = 28$ et le nombre de chromosomes pour un blé hexaploïde est de $6n = 42$. Toutefois, dans ces cas-là, la polyploïdie n'est pas issue de la duplication du génome, mais de l'hybridation de deux espèces proches. Il y a donc deux ou trois génomes qui coexistent dans un même noyau. Il est donc plus juste de parler d'allopolyploïdie, où les chromosomes correspondants qui n'appartiennent pas au même génome, sont dits homéologues. De façon générale et en dépit d'une certaine affinité, les chromosomes homéologues ne s'apparient pas à la méiose. L'appariement se fait principalement entre chromosomes homologues.

Il est donc possible de retracer l'histoire phylogénétique des blés et plus particulièrement du blé tendre (*Triticum aestivum* ssp. *aestivum*) (cf [Figure I.1](#) flèches rouges). Le premier événement d'allopoloïdisation a donné naissance au groupe tétraploïde BB AA suite à une hybridation entre *Aegilops spelta* (génome BB) et *Triticum urartu* (génome AA) qui aurait eu lieu il y a environ 300 000 ans. Alors que le second événement, qui a donné le groupe hexaploïde serait issu de l'hybridation

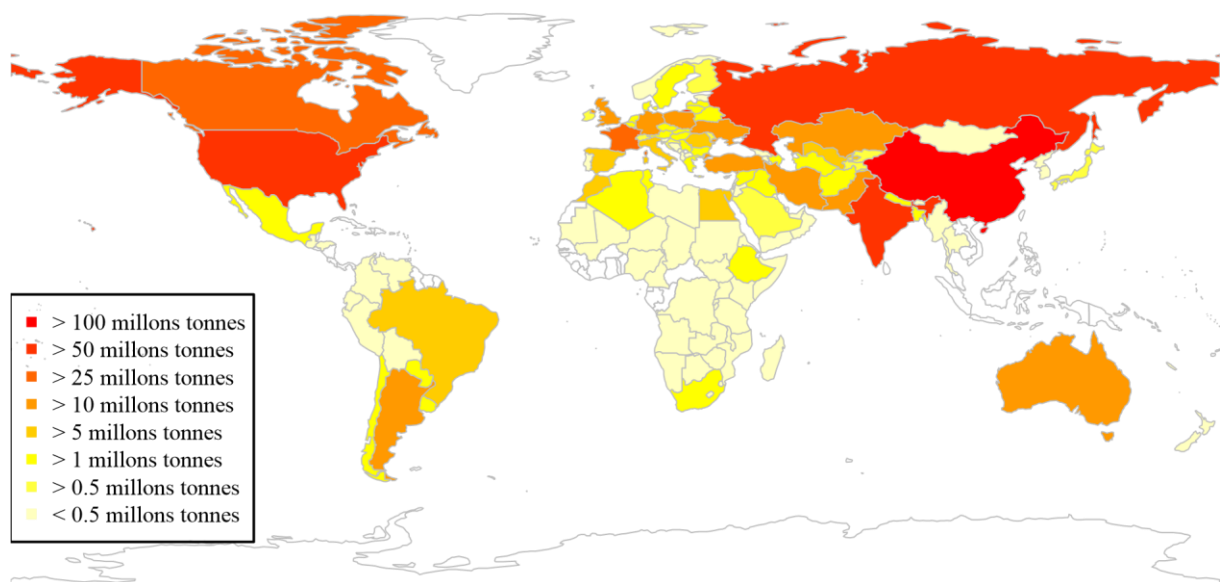


Figure I.2 : Répartition des productions moyennes entre 2008 et 2017, en millions de tonnes (FAO 2018)

entre un blé tétraploïde (génome BB AA) et *Aegilops tauschii* (génome DD) il y a environ 10,000 ans. Le blé tendre *Triticum Aestivum ssp. aestivum* est un blé allohexaploïde de génome AA BB DD, qui est très vraisemblablement apparu uniquement après la domestication des blés diploïdes et tétraploïdes (Goncharov et al., 2009).

Du fait de la grande complexité du génome du blé et de l'importance de cette céréale, plusieurs structures de recherche publique et privée et sociétés semencières se sont regroupées en 2005 pour séquencer le génome du blé. C'est ainsi qu'en 2018, le consortium international pour le séquençage du blé (IWGSC) a rendu public la première séquence annotée du blé (IWG et al., 2018). Ce séquençage a été réalisé sur les 21 chromosomes de la variété Chinese Spring (CS), il recouvre plus de 94% du génome, soit une taille de plus de 14.5 Gb. Sur l'ensemble de cette séquence, plus de 100 000 gènes ont été identifiés avec un niveau de confiance important. L'ensemble de ces ressources a et va avoir un impact significatif dans l'analyse des déterminants génétiques utiles à l'amélioration variétale.

1.2 Production et importance économique dans le monde

Le blé est cultivé à quasiment chaque endroit de la Terre où l'agriculture s'est développée. Il est cultivé à des latitudes très diverses, de la Scandinavie (67°N) à l'Argentine (45°S), Chili et Nouvelle-Zélande (Rajaram et al., 1995), et à des altitudes allant du niveau de l'océan à plus de 3000m sur les plateaux du Tibet (Chen et al., 2015). Cependant, il se trouve essentiellement dans les climats tempérés et subtropicaux.

Les chiffres présentés dans le paragraphe suivant proviennent de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO, 2018). De plus, pour éviter les variations interannuelles nous avons moyenné l'ensemble des chiffres sur les dix dernières années, *i.e.* de 2008 à 2017. Ainsi, c'est 771 millions de tonnes de blé qui ont été produites chaque année dans 124 pays. Plus de 52% de cette production vient de seulement cinq pays différents, avec en première position la Chine (17.4%), puis l'Inde (12.8%), la Russie (11.1%), les USA (6.1%) et la France (4.8%) ([Figure 1.2](#)). En fonction de leur type de production, extensif ou intensif, deux pays peuvent atteindre les mêmes niveaux de production avec des surfaces très différentes. Par exemple, l'Allemagne et le Pakistan ont produit en moyenne 24.8 et 24.2 millions, respectivement. Cependant, quand l'Allemagne réalise cette production sur seulement 3.2 millions d'hectare, il en faut près de 9 millions pour le Pakistan. Les hauts rendements de l'Allemagne lui permettent donc de produire autant sur des surfaces bien plus réduites.

Les quatre premiers pays producteurs de blé sont aussi les pays qui accumulent le plus de surfaces cultivées. En ajoutant l'Australie, ces pays représentent un peu plus de 50% des 220 millions d'hectares qui ont été cultivés, avec l'Inde (13.4%) en première position, suivie de la Russie (11.3%), la Chine (11.0%), des USA (8.6%) et l'Australie (5.8%) (FAO, 2018). Les rendements les plus élevés se trouvent quant à eux majoritairement en Europe. En effet, hormis la Nouvelle-Zélande, les huit pays

Table I.1 : Production et utilisation commerciale du blé dans le monde ces dix dernières années

Les données sont issues de la section « World Food Situation » et ont été mises à jour au 06/12/2018. Pour le blé, une saison commerciale s'étend du mois de juillet au mois de juin de l'année suivante. Le tableau montre l'évolution de la production, de la quantité de fourniture (production plus les stocks de l'année précédente), la quantité de fourniture exportée, le stock à la fin de l'année et les ratios entre les fournitures utilisées et le stock, enfin le ratio entre le stock et disparition des produits (utilisation domestique plus exports).

Saison	Production	Fourniture	Fourniture Utilisée	Fourniture Exportée	Stock à la fin de l'année	Ratio Fourniture utilisée / stockée	Ratio stock / Disparition
						Million de tonnes	
2009/10	684.8	865	654.9	130.9	209.9	32	22.8
2010/11	651	860.9	656.7	129.8	206	29.7	21.9
2011/12	697.8	903.9	694.7	149	203.9	29.9	18.8
2012/13	657	860.8	682.8	143.4	183	26.5	14.9
2013/14	712.9	896	689.9	159.5	196.8	27.6	15.7
2014/15	732.6	929.4	712.6	156.4	217.6	30.6	18.4
2015/16	734.6	952.2	710.2	167.4	236	32.1	18
2016/17	756.7	992.8	734.5	176.6	255.1	34.6	20.2
2017/18	760.3	1015.4	738	176.2	276.7	37.4	21.1
2018/19	725.1	1001.9	739.6	172.5	264.5	35	16.9

qui présentent les rendements les plus élevés de ces dix dernières années se trouvent en Europe du nord-ouest, avec en première position l'Irlande (92.2qtx/ha), suivie de la Belgique (87.7qtx/ha) et la Hollande (87.4qtx/ha). Ces niveaux de rendement sont souvent plus du double de ceux observés chez les quatre plus gros producteurs avec 50.6qtx/ha pour la Chine, 30.9qtx/ha pour les USA, 30.0qtx/ha pour l'Inde et 23.6qtx/ha pour la Russie.

Au cours des dix dernières saisons commerciales, la production annuelle, le stock, la disponibilité mondiale (les productions plus les stocks) et l'utilisation des ressources ont vu leurs quantités augmenter significativement ([Table I.1](#)). Ces dix dernières années, les exportations correspondaient à 16.8% des fournitures globales, ce qui fait chaque année 156.17 millions de tonnes. Les plus gros exportateurs sont, dans l'ordre, l'Argentine, l'Australie, le Canada et l'UE. Avec au sein de l'UE la France qui est le premier pays (FAO, 2018). Le stock est resté assez stable ces dernières années, avec en moyenne 224.95 millions de tonnes par an, ce qui représente plus de 24% des volumes.

Le prix de la tonne de blé est particulièrement volatile depuis le début du 21^{ème} siècle. Il a varié de façon importante passant de 152€ la tonne à plus de 375€, en moyenne il est de 229€ (macrotrends.net, 2018). Cette évolution est intrinsèque à tout produit agricole dans un système de commerce mondialisé, avec une augmentation de la demande liée aux évolutions démographiques, et une offre qui peut fortement varier en fonction des saisons (dû à l'influence du climat par exemple). La mise sur le marché a fortement augmenté cette fluctuation, notamment avec des phénomènes de spéculations qui ont pu être observés dans les années 2007-2008 et 2011-2012.

1.3 Importance nutritionnelle et utilisation industrielle du blé tendre

Le blé, le maïs (*Zea mays* L.) et le riz (*Oriza sativa* L.) sont les trois principales sources d'alimentation dans le monde. Ils permettent l'apport de 44% des calories et 37% des protéines. Avec 19% d'apport calorifique et 20% d'apport protéique, c'est le blé qui a la plus grande part dans ces apports (Braun et al., 2010).

Plusieurs débouchés industriels sont associés au blé tendre. Ils peuvent être séparés en cinq catégories différentes : (1) la farine, notamment pour faire du pain, qui est de loin le débouché le plus important ; (2) l'amidon, qui vient majoritairement des grains de maïs, mais peut aussi être issu des grains de blé (10% en Europe) ; (3) la nutrition animale ; (4) la production d'énergie via la production d'éthanol ou l'utilisation des pailles dans les bioréacteurs et (5) la production de semences. L'utilisation du blé pour l'un ou l'autre de ces secteurs est fortement orientée par l'accès aux différentes filières, mais aussi suivant la qualité de la production (FranceAgriMer).

La qualité d'un produit agricole peut faire référence à plusieurs critères, tous aussi importants les uns que les autres : la qualité sanitaire (e.g. présence de mycotoxine), la qualité nutritionnelle, la qualité

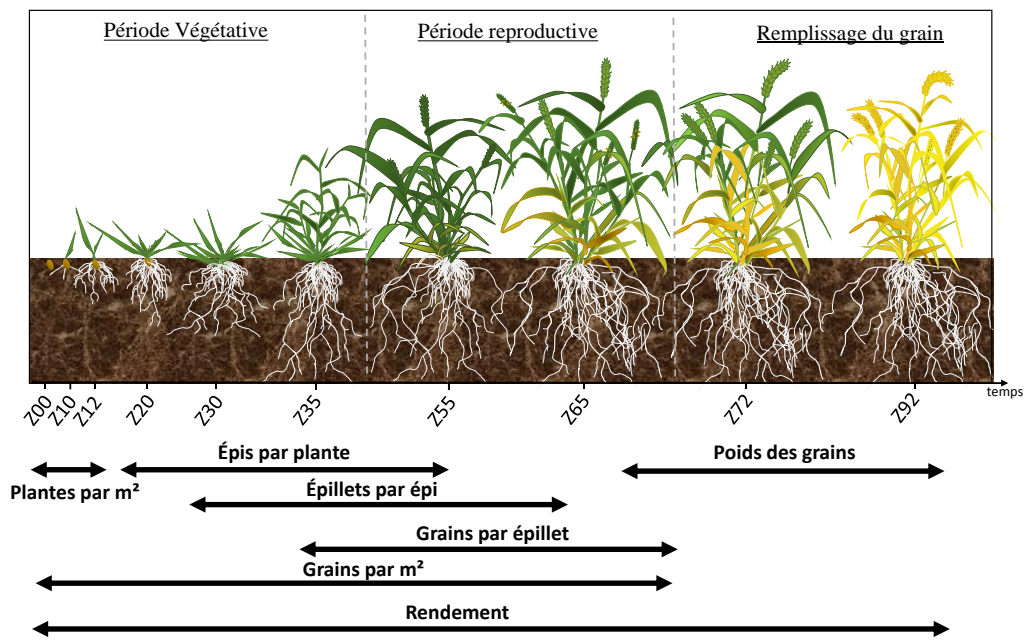


Figure I.3 : Les différents stades de développement du blé et l'initiation des différentes composantes du rendement

organoleptique (goût, odeur, aspect, couleur et consistance) et la qualité technologique qui permet la bonne transformation du produit en vue de remplir les deux critères précédents.

En France, le blé tendre est largement utilisé pour la production de farine en vue d'une transformation en boulangerie. Les produits issus de cette filière sont aussi bien consommés en France qu'exportés à l'international. Pour la fabrication de la farine, la qualité technologique du blé tendre est étroitement liée à la qualité boulangère de la farine. Chaque pays possède ses propres critères de qualité technologique, ils sont bien sûr guidés par leurs différentes utilisations (de la baguette croustillante aux pains de mie). La fabrication de la farine est dans la plupart des cas la résultante du mélange de plusieurs variétés de blé aux propriétés différentes. En France, cinq classes ont été distinguées pour le blé tendre : Blé Améliorant de Force (BAF), Blé Panifiable Supérieur (BPS), Blé Panifiable (BP), Blé Biscuitier (BB) et Blé Autres Usages (BAU). La qualité technologique de la farine va dépendre en partie de la qualité fermentaire et rhéologique de la pâte. La qualité fermentaire vient de la dégradation de l'amidon par les enzymes (alpha et bêta-amylase). Elles sont naturellement présentes dans le grain de blé et donneront des substrats assimilables par les levures. L'amidon du grain est présent sous forme de granules qui vont être endommagés lors de la moulure du grain, facilitant la dégradation de l'amidon par les enzymes. Il a été montré que la dureté (résistance à l'écrasement) d'un grain est positivement corrélée à la teneur en protéines, notamment lié au locus *Ha* (Jolly et al., 1996; Giroux and Morris, 1998). La qualité rhéologique de la pâte est définie par l'alvéographe de Chopin qui va mesurer la ténacité, l'extensibilité et l'élasticité de celle-ci. Ces trois paramètres sont dépendants de la teneur en gluten des farines. Le gluten est une fraction protéique insoluble qui se forme suite au pétrissage de la farine. Il est composé de deux familles de protéines issues des protéines de réserves situées dans la couche à aleurone, il s'agit des gliadines et des gluténines. La qualité boulangère a connu une forte amélioration depuis la moitié du 20^{ème} siècle. En effet, la force boulangère mesurée par l'indice W (le « travail » de la pâte, mesuré avec un alvéographe de Chopin) était de 120 pour la variété "Vilmorin 27" inscrite en 1933, aujourd'hui celle-ci serait considérée impropre à la panification. Les blés cultivés actuellement ont un W moyen de 196 (FranceAgriMer).

2 Les stades de développement du blé et la mise en place du rendement

2.1 L'échelle de Zadoks pour définir les stades de développement du blé

Le développement du blé suit une succession d'étapes déterminées et répétées de la germination à la récolte. Schématiquement, il peut être divisé en trois périodes majeures : (1) la période végétative qui s'étend du semis à l'initiation de la floraison ; (2) la période reproductive qui s'étend de l'initiation à la floraison et (3) le remplissage du grain qui s'étend de la floraison jusqu'à la maturité du grain [Figure I.3](#). Plusieurs échelles ont été proposées pour décrire finement chacune de ces étapes, mais c'est Zadoks et al., (1974) qui ont publié l'une des plus utilisées en agronomie et en recherche. Cette échelle découpe le cycle de développement des céréales en différentes étapes de Z00 à Z99. Le premier des

deux chiffres représente les 10 stades les plus importants pour le blé, par exemple, (0) la germination, (1) la plantule, (6) la floraison... Le second chiffre symbolise une des sous-étapes du stade, par exemple, Z11 est la sortie de la première feuille au stade plantule et Z12 est la sortie de la deuxième feuille.

Bien définir les stades d'une culture est très important car ils représentent des étapes clés dans le développement de celle-ci, par exemple : Z40 est le stade de gonflement de l'épis et coïncide avec la méiose, Z55 est un stade facilement identifiable car c'est le moment où la moitié des épis d'un couvert de blé est sortie de la gaine, ou encore Z65 qui est le stade de floraison (sortie des anthères). Ces stades sont régulièrement utilisés par les agronomes car ils leur permettent de bien situer l'avancement d'une culture et leur permettent des prises de décision quant à sa conduite. Par exemple, en France, il est recommandé de fractionner les apports azotés, avec la première application avant le stade épis 1cm (Z30) et la deuxième pendant ce stade.

La date de réalisation d'un stade est contrôlée par différents opérateurs : génétiques et climatiques. Pour certains stades, il est possible de prédire leur apparition en connaissant l'un et/ou l'autre des opérateurs (Gate, 1995). Par exemple, la floraison (Z55) chez le blé à lieu entre 4 et 15 jours après l'épiaison. Mais comme le développement des plantes répond à une somme des températures (facteur climatique), on peut estimer de façon plus précise la floraison comme étant l'épiaison plus 145°jour. La somme de degrés jours (°jour) est la somme des températures moyennes au-dessus d'un seuil critique de développement, chez le blé ce seuil est égal à 0°C.

Le blé s'est adapté à de nombreux environnements en particulier en raison d'une grande plasticité dans son rythme de développement qui dépend à la fois des conditions du milieu et de son contenu en gènes d'adaptations. Les blés de type hiver nécessitent une longue période d'exposition à des températures froides pour que la transition végétative- reproductive se déclenche. Les blés de printemps peuvent se passer de cette période dite de vernalisation. De plus, pour fleurir, certains génotypes sont sensibles à la photopériode, c'est-à-dire qu'ils ont besoin de jours longs pour fleurir. Ces besoins en jours longs qui correspondent à la sortie de l'hiver en Europe, permettent de retarder les stades de différenciation en organes reproducteurs et de méiose, qui sont des étapes sensibles au froid. Les blés insensibles à la photopériode et par conséquent à forte précocité intrinsèque peuvent fleurir de manière précoce et permettent théoriquement d'éviter les stress thermiques en cours de remplissage ou les sécheresses de fin de cycle, mais peuvent être sensibles aux froids tardifs durant certains hivers. Les gènes majeurs de floraison et de photopériode ont été largement décrits dans la littérature et seront présentés dans la [partie I.5.1.](#)

2.2 Le rendement se met en place tout au long du cycle du blé

De façon générale, il est important de bien identifier les stades de la culture car à chaque stade peut être associé le développement d'une ou plusieurs composantes du rendement [Figure I.3.](#) En effet, le

rendement du blé est la multiplication du nombre de plantes par m², du nombre d'épis par plante, du nombre de grains par épis et du poids des grains. Chacune de ces composantes se met en place sous l'influence de facteurs génétiques, mais aussi sous l'influence de l'environnement. De plus, elles se mettent en place en cascade et donc une compensation entre les différentes composantes est couramment observée (Gate, 1995).

2.2.1 La densité de plantes au mètre carré

Le nombre de plantes par mètre carré est défini par la densité de semis, la faculté germinative des semences, le taux de levée et les pertes hivernales. La densité de semis est une variable d'ajustement sur laquelle l'agriculteur peut jouer. En France, elle peut varier entre 150 et 350 grains par m² en fonction du type de sol et de la région. La faculté germinative est quant à elle déterminée par la qualité des semences. Pour des semences certifiées, elle s'élève à plus de 90%, mais elle peut aussi chuter rapidement en fonction des conditions de conservation et des traitements qui lui ont été appliqués. Le taux de levée est la proportion de graines qui ont effectivement germé et qui sont sorties de terre. Le plus souvent, le taux de levée diminue en raison d'obstacles physiques, de semis trop profond, de la présence de mottes, de pierre ou d'une croûte de battance. Enfin, les pertes hivernales sont le plus souvent liées à des événements climatiques extrêmes. En France, ces facteurs climatiques sont souvent liés à des excès d'eau (ennoyage et submersion) mais aussi au froid (déchaussement, gel mécanique ou dépassement du seuil de résistance de la culture).

2.2.2 Le nombre d'épis par plante

Le nombre d'épis par plante s'établit en deux phases : (1) la phase de tallage (Z20) à proprement parlé, où les talles sont émises en quantité largement suffisantes (en France cette phase se passe pendant l'hiver) et (2) une phase de régression de talles (*i.e.* les bourgeons des talles meurent) qui se passe pendant le stade de montaison (Z30) (en France, cela se déroule au printemps).

Dans des conditions optimales, la première talle apparaît à l'aisselle de la première feuille entre le stade Z13 et Z14. Ensuite, une talle apparaît à chaque phyllotherme, qui est une durée exprimée en somme de température séparant l'apparition de deux feuilles successives. Chez le blé, un phyllotherme est d'environ 100°C.jour (Gate, 1995). La plupart des talles sont issues du brin maître, mais pour certaines variétés et dans certains cas il est possible que des talles se développent à partir de talles secondaires et tertiaires ou à partir du coléoptile (Kirby et al., 1987; Gate, 1995). Une plante isolée peut produire un nombre important de talles (jusqu'à 25-30) qui donneront des épis et des graines matures, ce nombre peut dépendre de la variété et aussi de l'environnement dans lequel elle pousse. En effet, le nombre de talles décroît en fonction de la densité du couvert végétal (Darwinkel, 1978). Pour la même variété, ce nombre peut décroître jusqu'à seulement une ou deux talles voire aucune. C'est la compétition pour la lumière qui va faire que la plante va plutôt développer de la biomasse en hauteur que de la biomasse en largeur.

L'arrêt du tallage est contrôlé par deux mécanismes : (1) le premier mécanisme est engendré par la réduction de la lumière nécessaire pour la photosynthèse (Simon and Lemaire, 1987) ; (2) le second est engendré *via* un signal lumineux perçu par la plante : l'équilibre entre les longueurs d'onde rouge clair et rouge sombre est modifié par la densité de végétation (Holmes and Smith, 1977). Ainsi, la plante quantifie la concurrence dans son entourage (autres plantes de la même culture ou adventices) et module son tallage en conséquence.

La régression de talles ne débute que lorsque le tallage est arrêté, et il n'y a pas de cohabitation entre l'émission de nouvelles talles et la régression d'autres (Gate, 1995). Dans des conditions optimales et de faible densité, l'arrêt du tallage correspond à la fin de la formation de l'épi du brin maître (Baker and Gallagher, 1983; Gómez-Macpherson, 1998), qui coïncide avec l'élongation de la tige (Hay, 1999). La régression des talles correspond à une concurrence trophique entre organes : les tiges les plus développées mobilisent les ressources (azote, eau, lumière) au détriment des tiges les plus récemment émises, qui régressent en premier. Selon les conditions de croissance et la disponibilité des ressources, la régression sera plus ou moins forte. Plus les semis sont denses, plus la régression sera forte. A l'inverse, lorsque la levée a été difficile ou que des plantes ont disparu pendant l'hiver, le tallage sera fort et la régression faible. Généralement, peu après l'épiaison, le peuplement d'épis fertiles final est définitif (Gate, 1995).

2.2.3 Le nombre de grains par épis

L'épi de blé est une inflorescence composée car elle rassemble plusieurs ensembles de structures florales, les épillets. Chaque épillet rassemble plusieurs fleurs, le plus souvent trois à quatre sont fertiles. La distribution du nombre de fleurs fertiles le long de l'épi n'est pas équivalente. Les épillets du centre de l'épi sont souvent ceux qui ont le plus de fleurs fertiles donc potentiellement de grains. C'est le nombre de fleurs fertiles par épi qui va être le plus important pour définir le nombre de grains (Kirby, 1988; Slafer and Andrade, 1993; Miralles et al., 1998, 2003).

La fertilité des épis est l'une des composantes les plus sensibles de la mise en place du rendement chez de nombreuses céréales. Elle s'élabore sur une période assez longue qui s'étend du milieu du tallage jusqu'à la fécondation [Figure I.3](#).

Si les stades définis par Zadoks décrivent très bien les changements visuels de la plante, ils sont limités pour décrire les changements au sein des tissus. En effet, la transition florale débute lorsque le bourgeon apical (ou primordia) ne produit plus de feuilles mais commence à produire des épillets qui eux-mêmes produiront des fleurs. Chaque épillet initie entre huit et douze bourgeons floraux. Ce changement se produit au milieu du tallage ou le brin maître possède entre quatre et huit feuilles. Dans des conditions optimales, moins de la moitié des fleurs seront complètes et fonctionnelles au moment de la fécondation (Kirby et al., 1987; Kirby, 1988; Hay and Kirby, 1991). Le nombre d'épillets varie de 20 à 30 par épis (Allison and Daynard, 1976; Kirby et al., 1987).

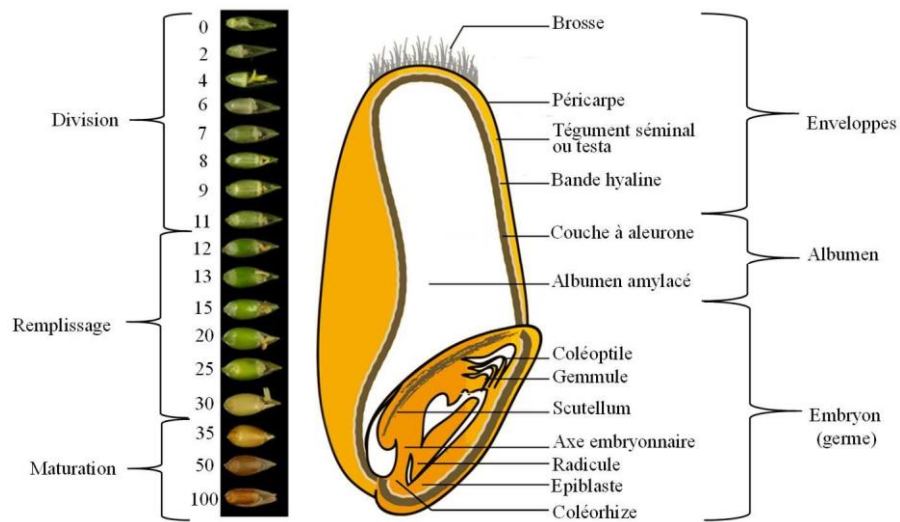


Figure I.4 : Développement du grain en nombre de jours après floraison et coupe longitudinale d'un grain mature

d'après Wikipedia (2018), photo grain présentation Jean Roche HeatWheat (2017)

A la fin de ce stade, l'épi complet qui est au niveau du sol commence à monter dès que l'élongation des entre-nœuds s'initie. Ce stade est très sensible à des stress environnementaux. Même s'il est difficile à repérer, il est identifié par les agriculteurs et agronomes comme étant le stade épis 1cm ou Z30. La tige va donc s'allonger et l'épi va grossir. La croissance de l'épi commence au moment où la feuille drapeau sort et continue jusqu'à dix jours après l'anthèse, qui est la période pendant laquelle une fleur est complètement ouverte et fonctionnelle (Kirby et al., 1987). L'avortement des fleurs commence au moment du grossissement de l'épi et se termine aussi à l'anthèse. Cet avortement est dû à des facteurs multiples, notamment à la compétition pour les ressources et/ou des facteurs environnementaux. La méiose commence aussi à ce stade. C'est elle qui est à l'origine de la fabrication du pollen dans les anthères et de l'ovule dans le sac embryonnaire. Chez le blé, la méiose commence au centre de l'épi puis se poursuit vers les extrémités (Zadoks et al., 1974).

Enfin, vient le stade de la floraison et fécondation. A ce stade, chaque épillet contient entre trois et six fleurs potentiellement fertiles (Kirby et al., 1987). Comme pour la méiose, la floraison commence au milieu des épis pour s'étendre aux extrémités de celui-ci pendant une durée de trois à cinq jours (Peterson, 1965; Baillot et al., 2018). Le blé est une plante dite autogame. En effet, dans plus de 95% des cas, le pollen de la même plante va féconder ses propres ovules, c'est l'autofécondation (Martin et al., 2006). Dans la majorité des cas, ce sont les grains du centre de l'épi qui seront les plus gros (Simmons and Crookston, 1979).

2.2.4 Le poids des grains ou PMG (poids de mille grains)

Le grain commence à se former après la fécondation. La durée du développement de celui-ci dépend de la variété et de facteurs environnementaux, notamment la température. Ce développement est classiquement décomposé en trois phases ([Figure I.4](#)) : (1) La phase de division cellulaire qui s'étend de la floraison au stade grain laiteux, et qui dure environ 20 à 30% de la durée de formation du grain (Acevedo et al., 2006). Pendant cette étape, les cellules de l'endosperme se multiplient et se différencient en cellules de stockage de l'amidon. Cette multiplication conditionnera en partie la capacité du grain à accumuler des réserves. Durant cette phase, on observe notamment une augmentation du volume du grain grâce à une forte absorption d'eau. L'embryon, quant à lui, entame son développement, puis entre en phase latente. (2) Le remplissage du grain ou l'accumulation d'amidon, s'étend du stade grain laiteux à grain pâteux, environ 50 à 70% de la formation du grain est réalisée (Acevedo et al., 2006). Les cellules qui ont été formées dans la phase précédente vont accumuler les sucres essentiels pour la mise en réserve sous forme d'amidon. Les sucres accumulés dans les grains à cette période sont issus à la fois de l'activité photosynthétique de la plante, mais surtout de la remobilisation via les feuilles et les tiges. Le poids de la matière sèche des grains augmente alors que le niveau d'humidité reste le même : c'est le pallier hydrique. Les enveloppes se modifient tant d'un point de vue morphologique que biochimique et l'embryon reprend son développement (Jones et al., 1985). (3) La phase de maturation du grain ou dessiccation, qui s'opère

lorsque la maturité physiologique est atteinte, c'est-à-dire juste après le stade grain pâteux jusqu'à récolte. A partir de ce moment-là, il n'y a plus de migration de réserves vers les grains et leur poids définitif est pratiquement atteint. Des phénomènes de polymérisation tardive des protéines peuvent apparaître pendant cette phase. Durant cette phase, s'acquiescent les qualités technologiques finales des farines.

2.3 Le développement racinaire

A ce stade nous n'avons parlé que du développement aérien de la plante, mais il ne faut pas oublier le développement de l'appareil racinaire, notamment dans le cadre de l'étude de la tolérance à un stress hydrique. Le développement de l'appareil racinaire n'est pas le plus évident à mesurer et à appréhender, mais il est l'une des composantes majeures de la croissance des plantes puisqu'il permet la prospection du sol pour l'approvisionnement en eau et en minéraux. L'ensemble du système racinaire des plantes peut être organisé en trois grandes classes : (1) racines pivotantes, (2) racines fasciculées et (3) un système de racines adventives. L'architecture racinaire diffère selon les espèces, mais aussi entre les génotypes d'une même espèce, notamment chez les céréales (Deak and Malamy, 2005). Chez le blé, l'architecture racinaire principale est un système de racines fasciculées. Cependant, il existe aussi un système adventif qui apparaît à la base de chaque talle, si bien que chacun d'eux peut se développer de façon indépendante de la plante mère (Lucas et al., 2000). Les racines séminales (provenant de la semence) sont associées dans le grain à différentes parties de l'embryon : une racine principale qui se développe suite à l'allongement de la radicule, puis quatre à six racines issues du scutellum et de l'épiblaste ([Figure I.4](#)). La première racine séminale qui est émise définit la germination au sens strict du terme (Z07), puis deux racines de scutellum et enfin les deux premières racines d'épiblaste apparaissent au moment de la levée (Z10) et la troisième à Z11. Les racines primaires adventives quant à elles sont émises régulièrement en fonction des degrés jours à partir du stade Z13-Z14 jusqu'à la fin de la montaison (Z39). Il existe une relation linéaire entre l'émission des racines primaires nodales et l'émission des feuilles, avec nombre de racines primaires nodales = 1.92 fois nombre de feuilles de brin maître moins 5.6 (Gate, 1995). Une plante peut posséder jusqu'à une trentaine de racines primaires nodales en condition optimale de culture (Klepper et al., 1984).

La structure des racines est largement déterminée par l'angle de croissance des racines séminales (Tottman, 1987) et des racines adventives. Ces dernière qui semblent aussi être associé à l'efficacité d'acquisition de l'eau et des nutriments phosphatés (Manschadi et al., 2013). Dans un sol homogène, la densité des racines est très différente selon que l'on observe la couche superficielle ou les couches profondes. La densité de longueur racinaire en surface est relativement importante et peut aller de de 3 à 5 cm.cm⁻³, pour la plupart des cultures céréalières. Cette forte densité excèderait la densité requise pour extraire l'eau disponible en surface (Passioura, 1983). En 2012, Wasson et al., ont montré que la diminution de la quantité des racines dans ces horizons superficiels pourrait s'avérer avantageuse en cas de stress hydrique en fin de développement, ceci au dépend de la nutrition minérale. Pour les

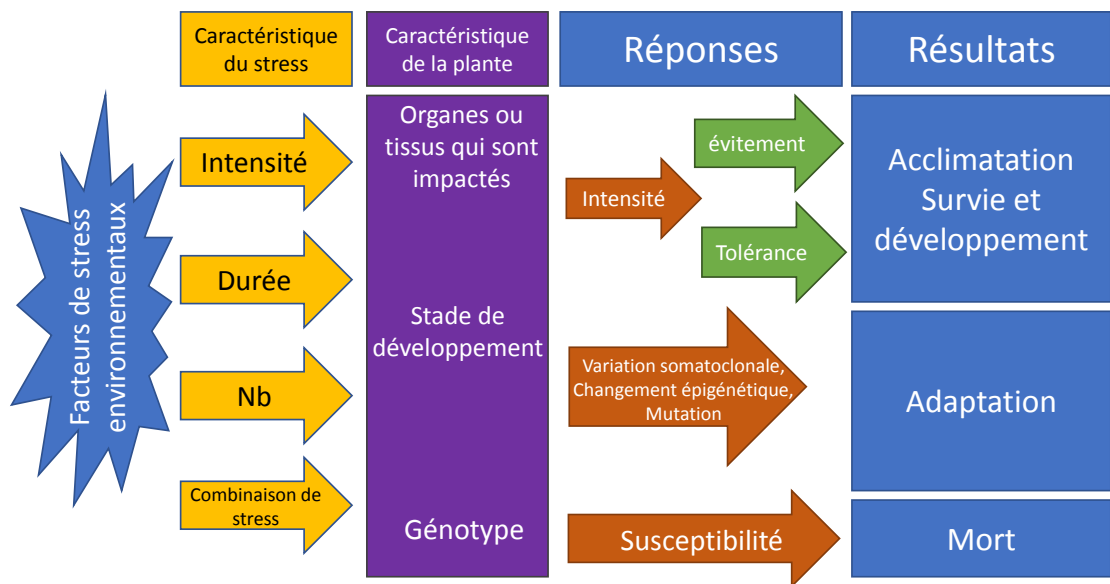


Figure I.5 : Impact d'un stress environnemental sur les plantes, d'après Gaspar et al., (2002)

couches plus profondes, l'accès à l'eau va de pair avec une augmentation du volume de sol prospecté, donc une augmentation de la profondeur d'enracinement. Les racines des céréales peuvent atteindre deux mètres de profondeur au maximum de leur développement, qui coïncide avec la floraison de la plante (Z65 ; Lucas et al., (2000)). Il a aussi été estimé qu'il faut 1cm de racines pour extraire l'eau disponible dans 1cm³ de sol, dans des conditions optimales. Comme pour les parties aériennes, la croissance des racines dépend de la température et de l'environnement dans lequel elles croissent (texture du sol).

3 Du concept de stress à l'impact de ceux-ci sur le développement des plantes

Nous avons vu que le développement du blé était régi par des facteurs génétiques et environnementaux. La notion de facteurs environnementaux fait donc tout de suite écho aux notions de facteurs limitants ou stressants. En biologie, le terme de "stress" n'est pas simple à définir. En 1980, Levitt propose de définir le terme de stress comme tout facteur environnemental non-favorable à un organisme bien défini. Sharkey, (1997), quant à lui, propose de définir le terme stress au sens physiologique comme étant l'altération d'un équilibre causé par un facteur.

En agronomie, les facteurs de stress sont souvent séparés en deux classes : (1) les facteurs biotiques qui comprennent les insectes et les pathogènes comme les champignons, bactéries et virus ; (2) les facteurs abiotiques qui sont associés aux conditions pédoclimatiques, comme le stress hydrique, stress thermique, manque d'azote ou manque de lumière, etc... Dans l'ensemble du manuscrit, nous utiliserons les termes de stress hydrique et/ou stress thermique pour définir respectivement un stress lié à un manque d'eau et/ou une température trop élevée. Cependant, *stricto sensu* le terme de stress hydrique peut également être utilisé pour définir de trop grandes quantités d'eau (provoquant l'anoxie des racines), de même que les stress thermiques peuvent être liés à des températures gélives.

Les plantes ne se déplacent pas et l'environnement dans lequel elles croissent est en constante évolution. Donc si l'environnement n'est plus favorable, pour de multiples raisons, elles ne peuvent s'en échapper en changeant d'environnement et doivent faire face à ces facteurs dit de stress. Les plantes ont une certaine flexibilité dans leur métabolisme qui leur permet d'être en adaptation constante (Gaspar et al., 2002). De plus, les plantes sont capables de s'acclimater à leur environnement ([Figure I.5](#)), ce qui signifie qu'une déviation d'un facteur par rapport à un optimum n'est pas systématiquement synonyme de stress et de perte de rendement (Gaspar et al., 2002; Taiz and Zeiger, 2002).

A l'échelle du cycle de développement d'une culture, les stress sont définis par trois éléments : le moment auquel la plante subit le stress, le niveau d'intensité du stress et la fréquence d'apparition de celui-ci. Ces trois éléments combinés forment un scénario de stress dans lequel la plante va devoir se développer. Comme nous avons vu précédemment, certains stades de développement de la plante sont plus ou moins sensibles et vont impacter plus ou moins fortement le rendement, qui est l'objet

Table I.2 : Seuils de température nécessaires pour le développement des plantes, d'après Porter et Gawith, (1999)

La table présente la moyenne des seuils de température (en °C) et les écarts-types définissant les seuils minimums (*Tmin*), le seuil de température optimal (*Topt*) et le seuil de température maximal (*Tmax*), d'après l'ensemble des publications qui ont été analysées par Porter et Gawith, (1999).

<i>Processus biologiques</i>	<i>Tmin</i>	<i>Topt</i>	<i>Tmax</i>
<i>Initiation des feuilles</i>	-1.0 (1.1)	22.0 (0.4)	24.0 (1)
<i>Croissance des tiges</i>	3.0 (0.4)	20.3 (0.3)	>20.9 (0.2)
<i>Croissance des racines</i>	2.0 (0)	<16.3 (3.7)	>25.0 (5)
<i>Phénologie stades du semis à l'épiaison</i>	3.5 (1.1)	22.0 (1.6)	32.7 (0.9)
<i>Vernalisation</i>	-1.3 (1.5)	4.9 (1.1)	15.7 (2.6)
<i>Epiaison</i>	1.5 (1.5)	10.6 (1.3)	>20.0 (0)
<i>Floraison</i>	9.5 (0.1)	21.0 (1.7)	31.0 (0)
<i>Remplissage du grain</i>	9.2 (1.5)	20.7 (1.4)	35.4 (2)

d'intérêt pour l'agriculteur. L'ensemble des impacts des stress hydriques et thermiques sur le développement de la plante sont présentés de façon séparée dans cette partie.

Dans un premier temps, nous décrirons l'impact d'un stress thermique sur les plantes. Le concept de stress dû à des températures élevées semble être le plus simple à présenter. C'est l'essence même du fonctionnement biochimique des cellules qui en est la raison. En effet, toutes les enzymes ont une température optimale d'activité et par transposition, l'organe en a une aussi. Les températures chaudes peuvent avoir deux effets : un effet réversible lié à des températures non-extrêmes qui interviennent en dehors des stades sensibles souvent pendant une longue durée et ralentissent la croissance de la plante ; un effet irréversible lié à des températures extrêmes se traduisant par un arrêt de la croissance ou du développement de l'organe (Hall, 2012).

Ensuite nous nous intéresserons à l'impact du stress hydrique sur la plante. Cependant, à l'inverse du stress thermique, il semble important de clarifier les différents concepts de nutrition en eau de la plante, avant de présenter les différents impacts d'un stress hydrique sur la plante et sur les cultures. Dans cette partie nous ferons une description théorique des flux d'eau ainsi qu'une description pratique du suivi des flux d'eau de la plante en condition de culture en s'appuyant sur le modèle de culture CHN développé par Arvalis.

3.1 Impact des stress thermiques

Comme nous l'avons souligné précédemment, les stress thermiques sont souvent séparés en deux classes différentes : (1) Les stress de forte intensité ($>40^{\circ}\text{C}$) qui sont souvent de courte durée (quelques heures), on parle alors de "choc thermique". Ces chocs thermiques causent des dommages irréversibles sur les organes des plantes. (2) Les stress thermiques qui sont de faible intensité sur une longue durée, souvent c'est l'accumulation de plusieurs jours de températures stressantes. Ces types de stress ont un impact sur le fonctionnement des organes de la plante sans pour autant causer de dommages irréversibles.

De multiples études de la réponse du blé à différents niveaux de températures, appliquées à différents stades de croissances, ont été réalisées. C'est à partir de plus de soixante-cinq études que Porter et Gawith, (1999) ont estimé plusieurs seuils de température qui caractérisent le développement du blé : un seuil minimal, un optimal et un maximal ([Table I.2](#)). Dans ce tableau, on peut voir que chaque stade de développement de la plante peut être impacté par un stress thermique (haute température) et que chacun d'eux peut avoir des niveaux de seuils de température différents. En 2012, Wasson et al., ont montré que pour 17 espèces cultivées, les processus physiologiques de développement de la plante en réponse à la température n'ont pas été impactés par la sélection intensive de ces dernières années.

Dans cette partie, nous allons présenter les différents mécanismes physiologiques qui sont mis en cause en cas de stress thermique. Ceci en présentant l'impact du stress sur les différents stades de

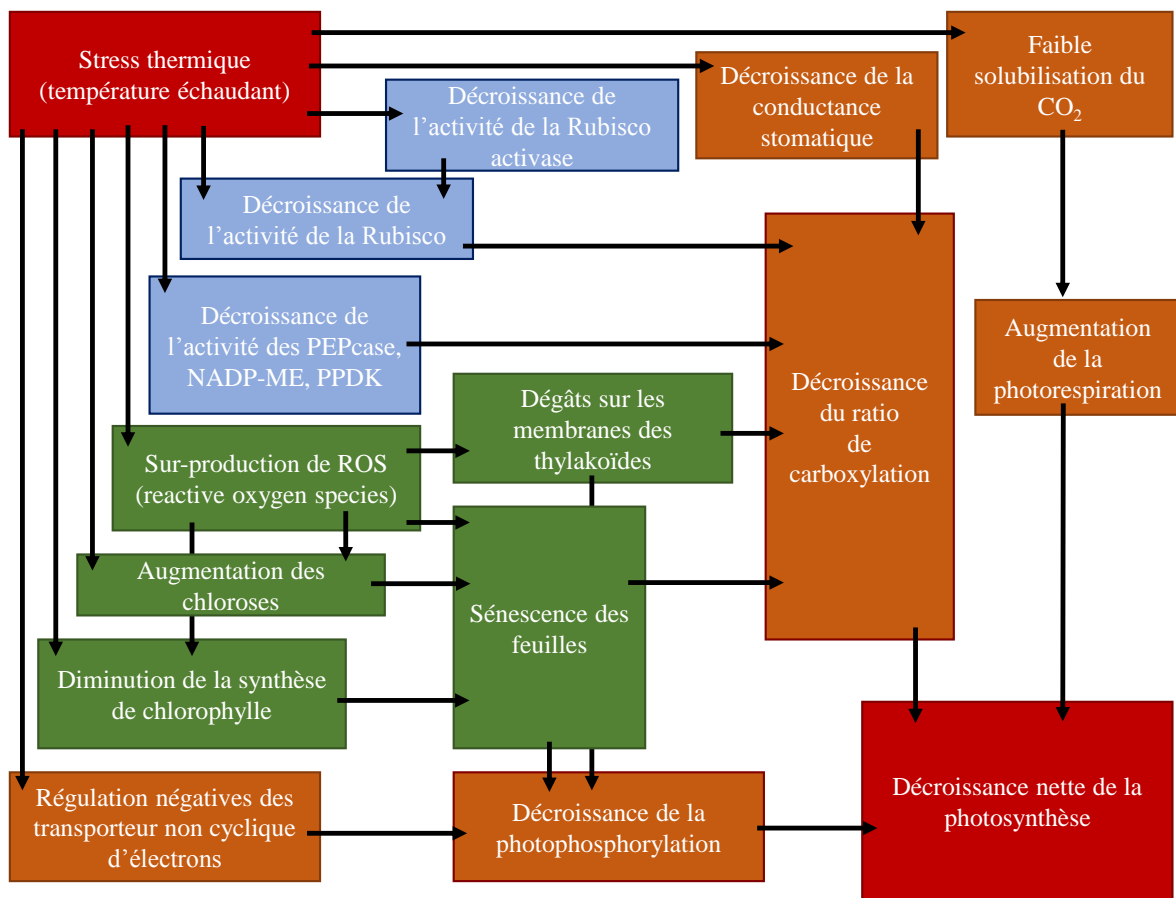


Figure I.6 : Schéma de l'impact possible d'un stress thermique sur la photosynthèse, d'après Farooq et al., (2011)

Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase), phosphoenolpyruvate carboxylase (PEPCase), nicotinamide adenine dinucleotide phosphate malic enzyme (NADP-ME) et pyruvate phosphate dikinase (PPDK).

développement comme la croissance végétative de la plante, l'appareil reproducteur et le remplissage du grain.

Nous allons nous focaliser sur des stress thermiques terminaux car ils sont les plus probables dans un contexte européen mais il en existe bien d'autres. Les scénarios de stress thermique au stade de la germination et au stade plantule ne sont pas très fréquents, voire quasiment nuls en Europe, car le blé tendre est semé durant l'automne où les températures gélives sont bien plus probables. Cependant, dans les régions plus chaudes, le blé est semé dans des environnements susceptibles de subir des températures chaudes et stressantes. En effet, la température maximale dans les premiers centimètres de sol peut dépasser de 10 à 15°C la température de l'air. Donc pour une température de l'air qui avoisine les 30°C, la température du sol peut atteindre 40°C à 45°C, ce qui peut entraîner des dégâts importants sur les plantules et donc sur le rendement (Acevedo et al., 2009).

3.1.1 Impact du stress thermique sur la croissance végétative

Tout processus physico-chimique ou biologique qui impacte la photosynthèse va impacter le développement de la plante, donc le rendement. Les températures élevées font partie de ces facteurs qui peuvent réduire l'activité photosynthétique (Wahid et al., 2007). Plusieurs mécanismes peuvent entraîner concomitamment une diminution nette de la photosynthèse. La [Figure I.6](#) présente un des schémas possibles de l'impact de températures trop élevées sur la photosynthèse. Pour résumer de façon succincte, deux voies peuvent être mises en avant : (1) la diminution de l'activité enzymatique. Par exemple, la Rubisco (Ribulose-1,5-bisphosphate carboxylase/oxygénase), qui est l'enzyme clé de la photosynthèse car elle permet la fixation du CO₂ pour initier le cycle de Calvin, va voir diminuer sa capacité à fixer le CO₂ en cas de températures trop élevées (Ogren, 1984; Long et al., 2004). (2) La dégradation des éléments responsables de la photosynthèse. Al-Khatib and Paulsen, (1999) ont montré que la diminution de la photosynthèse par un stress thermique est principalement dû à une dégradation des composantes du photosystème II, qui sont présentes dans la membrane des thylakoïdes des chloroplastes. Ils ont également montré que les variétés de blé des régions tempérées (dont le blé tendre d'hiver) ont une sensibilité plus élevée à ce type de stress que des espèces comme le riz ou le millet, qui elles, sont adaptées à des environnements plus chauds. D'autres effets peuvent être notés, par exemple, la réduction du contenu en chlorophylle va réduire la photosynthèse (Xu et al., 1995).

Il est aussi possible qu'à un certain seuil de température, la plante initie les processus de sénescence (Al-Khatib and Paulsen, 1999). En effet, chez les espèces monocarpiques, c'est-à-dire les plantes qui ne fleurissent qu'une fois avant de mourir, la sénescence des tissus végétatifs est la dernière étape du développement de la plante pour assurer la formation et la viabilité des graines et donc de sa descendance. La sénescence s'accompagne de la remobilisation des ressources vers la graine ou le fruit, ressources qui ont été accumulées tout au long du cycle de la plante. La sénescence est définie par la perte progressive et active de tissus verts qui survient à la suite de la reproduction (Nooden,

2014). Des effets de pertes de chlorophylle sous des températures trop hautes peuvent accélérer la sénescence ([Figure I.6](#)) (Tewari and Tripathy, 1998). Le même mécanisme est identifié suite à la dégradation de la membrane des thylakoïdes (Harding et al., 1990). Ces mécanismes sont néanmoins souvent identifiables sous l'effet de stress thermique de forte amplitude (>35°C). Zhao et al., (2007) ont, quant à eux, montré que des fluctuations importantes de température entre le jour et la nuit et des stress thermiques de moindre ampleur initient la sénescence des feuilles.

3.1.2 Impact du stress thermique sur le développement du grain de blé

Le nombre de grains et leur poids sont sensibles à une élévation de température (Ferris, 1998). Ces deux composantes sont les dernières à être mises en place. Elles sont la résultante de nombreuses compensations et donc peuvent être sensibles au stress à chacune des étapes de développement de la plante. Par exemple, un stress qui impacterait le nombre d'épillets par épi impacterait également le nombre de grains (Rawson and Bagga, 1979). Cependant, ce type de stress n'est pas très fréquent en Europe. De plus, en comparaison avec des stress sur l'appareil reproducteur du blé, ils sont souvent moins dommageables pour le rendement. Le remplissage du grain est quant à lui impacté par la capacité de la plante à remobiliser les assimilats. Le volume d'assimilats dépend de la vitesse, de la durée de la remobilisation et de la capacité de la plante à faire migrer les assimilats. Ces trois composantes peuvent être impactées par des températures trop élevées.

3.1.2.1 Effet du stress thermique sur l'appareil reproducteur et donc sur le nombre de grains

Des températures supérieures à 30°C pendant le développement de la fleur peuvent provoquer une stérilité complète (Saini and Aspinall, (1982)). C'est la nutrition carbonée des fleurons qui peut être compromise par ce stress (Abbate et al., (1995); Demotes-Mainard and Jeuffroy, (2004)). De plus, Saini and Aspinall, (1982) ont montré qu'une exposition des plantes à des températures supérieures à 30°C pendant la floraison peut réduire considérablement la fertilité des pollens. C'est la structure des sacs polliniques et des microspores qui est directement impactée par les stress thermiques (Anjum et al., 2008).

Le stress thermique réduit aussi la viabilité des ovaires et limite la fertilité de ceux-ci (Saini and Aspinall, (1982); Ferris, (1998)). Il existe des différences de sensibilité entre les génotypes (Calderini et al., (1999) ; Gibson and Paulsen, (1999) ; Zhao et al., (2008); Anjum et al., (2008)). Ces différences peuvent venir de différentes sources, par exemple, Mascarenhas and Crone, (1996) ont montré que la sensibilité du pollen à des températures trop élevées est associée à une incapacité de la plante à synthétiser certaines protéines de stress thermiques (heat shock proteins HSP). Les différences génotypiques et les moyens de les identifier et de les sélectionner seront présentés par la suite.

3.1.2.2 Effet du stress thermique sur le remplissage du grain

Le grain est composé de plus de 70% d'amidon et la réduction de la production d'amidon est une raison majeure de la diminution du poids des grains (Bhullar and Jenner, 1985). Comme pour tous les

processus biochimiques, la synthèse de l'amidon dépend d'une voie de biosynthèse complexe faisant intervenir de nombreuses enzymes. Si la température augmente au-dessus d'un certain seuil, l'activité de ces enzymes va être réduite et donc la production d'amidon aussi (Denyer et al., 1994; Jenner, 1994). Porter and Gawith, (1999) ont établi le seuil optimal du remplissage du grain à 20.7°C ([Table I.2](#)).

Pendant le remplissage du grain, les différents assimilats sont aussi bien transférés à partir des réserves de la plante que des nouveaux assimilats qui sont produits en même temps par la photosynthèse (Palta et al., 1994; Blum, 1997). Dans des conditions optimales, plus de 90% des ressources carbonées stockées dans le grain provient des assimilats produits durant le remplissage (Kobata et al., 1992). Comme les stress thermiques impactent la photosynthèse, cette proportion va changer en faveur de la remobilisation à partir de ressources déjà existantes. La plus grande partie de ces ressources est issue des tiges, mais aussi des feuilles. En fonction de l'intensité du stress sur la photosynthèse, Blum, (1997) a montré que le ratio de remobilisation à partir des tiges pouvait aller de 6% à 100%. Quelle que soit l'origine des ressources, la transition des assimilats va être réalisée à partir des voies apoplastiques et symplastiques (c'est-à-dire à travers la paroi cellulaire ou *via* la paroi cellulaire). Ce transport peut être impacté par des températures stressantes.

Wardlaw, (1974) a montré que ce transport est substantiellement réduit en cas de températures supérieures à 30°C si les assimilats sont issus de la feuille drapeau, mais qu'il n'est pas impacté par des températures allant jusqu'à 50°C si les assimilats sont issus des tiges. Ceci laisse penser que le remplissage du grain par la remobilisation des assimilats venant des tissus existants n'est pas très impacté par les stress thermiques, mais il serait plutôt impacté par la durée et la vitesse de remobilisation.

La vitesse de remobilisation des ressources des tissus végétatifs vers le grain est accélérée en cas d'augmentation de températures, ce qui engendre une durée de remobilisation réduite (Dias and Lidon, 2009). Streck, (2005) a estimé que pour chaque °C au-dessus de la température optimale, la durée du remplissage du grain est réduite de deux à trois jours. Des températures nocturnes ont aussi un impact important sur la durée du remplissage du grain. Prasad et al., (2008) ont montré qu'une température de 20 à 23°C peut réduire la période de remplissage de 3 à 7 jours. De plus, Sofield et al., (1977) ont montré que la vitesse de remplissage du grain est augmentée pour des températures entre 15 et 21°C. Cette augmentation compense la perte de poids de matière sèche du fait du ralentissement de la photosynthèse, mais elle n'est pas suffisante pour des températures élevées qui vont de 21 à 30°C. Plusieurs études confirment ces travaux (Wiegand and Cuellar, (1981); Jenner, (1994)), alors que Bhullar and Jenner, (1985) reportent aussi que la phase de remplissage survient plus rapidement en cas de stress thermique. Que les expositions à de hautes températures soient prolongées ou non, elles réduisent toutes la durée de sénescence dans le temps. Cette réponse est liée à des signaux hormonaux

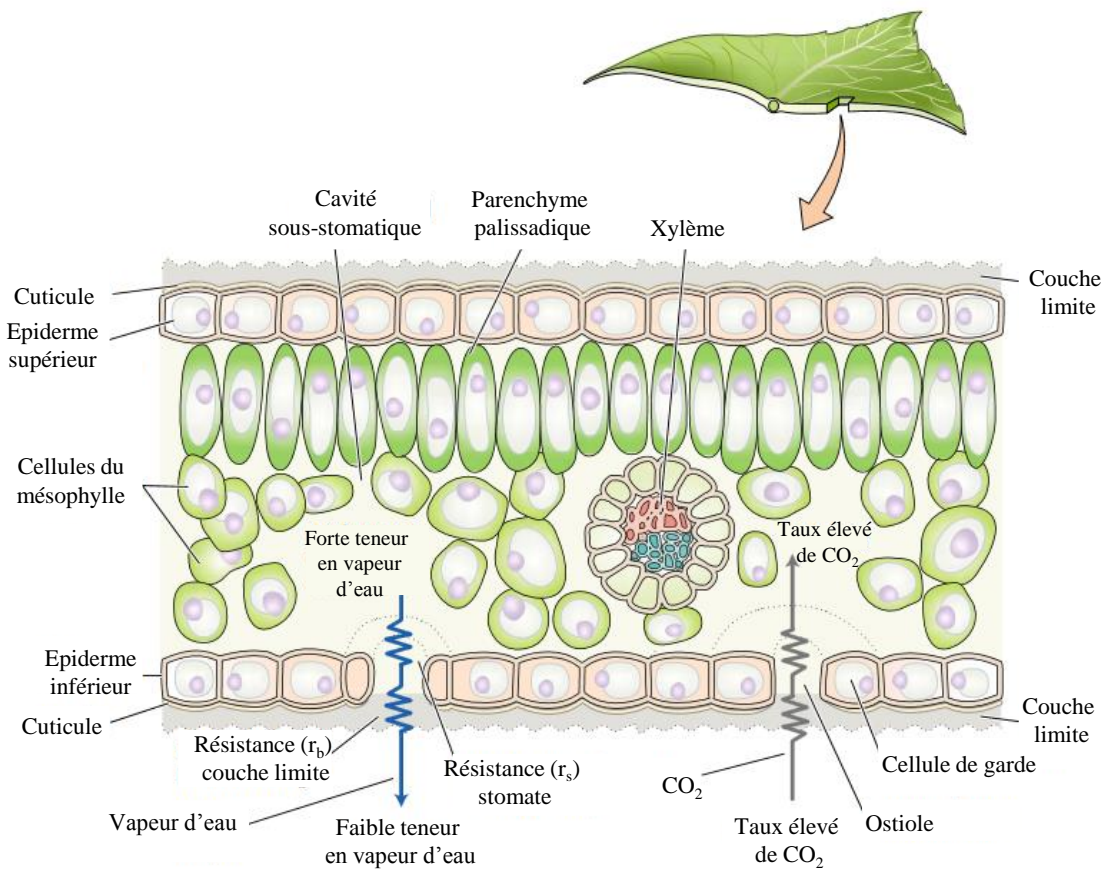


Figure I.7 : Les échanges gazeux entre la plante et l’atmosphère, figure adaptée de Taiz and Zeiger, (2002)

dans le grain, notamment l'éthylène qui est une hormone de sénescence (Hays et al., (2007). Ainsi, l'augmentation de la température va réduire le temps d'interception de la lumière.

3.2 Impact des stress hydriques

3.2.1 Description des flux d'eau

Comme pour l'ensemble des êtres vivants, l'eau est indispensable à la vie des plantes. Dans certains tissus, la concentration en eau peut s'élever à 70%. Elle est en perpétuel mouvement au sein de la plante et lui permet de croître et de survivre. La totalité des besoins de la plante en eau vient du sol grâce à l'absorption faite par les racines. Cependant, 97% de l'eau absorbée par les racines est perdue dans l'atmosphère, 2% est utilisée pour accroître le volume des cellules et seulement 1% utilisée pour des processus métaboliques (Taiz and Zeiger, 2002).

L'eau est à la base de plusieurs réactions biochimiques, dont la photosynthèse. La photosynthèse peut être définie par l'équation simplifiée faisant réagir du CO₂, de l'eau (H₂O) et un photon et qui produit des sucres simples et de l'oxygène. Pour capter les molécules de CO₂ de l'atmosphère, la plante ouvre ses stomates et donc s'expose à des pertes d'eau par évaporation ([Figure I.7](#)). Environ 400 molécules d'eau sont perdues pour chaque molécule de CO₂ captée (Taiz and Zeiger, 2002). L'eau permet aussi de transporter des nutriments grâce à la sève brute et à la sève élaborée. L'évaporation de l'eau au niveau des stomates va aussi jouer un rôle de thermorégulation.

L'ensemble des échanges d'eau, de gaz et d'énergie au niveau des couverts végétaux est défini comme le continuum sol/plante/atmosphère (CSPA). Et dans les cas où la plante est dans des conditions non limitantes en eau, alors le CSPA est réalisé de façon passive grâce au potentiel hydrique.

3.2.1.1 Le concept de potentiel hydrique

Le potentiel hydrique est la mesure de l'énergie libre, au sens thermodynamique du terme, de l'eau par unité de volume (J.m⁻³ ou Pa). Le potentiel hydrique est la différence de potentiel entre de l'eau pure (à température et à pression ambiante) et de l'eau dans un autre état, par exemple dans la sève de la plante. En effet, toutes les substances présentent dans la sève (sucre, minéraux...) exercent sur les molécules d'eau une force d'attraction. Plus la sève est concentrée et plus cette force d'attraction est forte, et moins les molécules d'eau peuvent s'extraire facilement de cette attraction. En physiologie végétale le potentiel hydrique est noté Ψ_w , il est l'addition de trois variables :

$$\Psi_w = \Psi_s + \Psi_p + \Psi_g$$

le potentiel osmotique (Ψ_s) qui est égal à zéro dans l'eau pure et est supérieur à zéro dans une solution aqueuse ; la pression hydrostatique ou de turgescence (Ψ_p) qui est due à la présence de la paroi de la cellule végétale, il en résulte que Ψ_p est supérieure à la pression atmosphérique et généralement comprise entre 0.3 et 0.8 MPa ; la gravité (Ψ_g) qui est significativement supérieure à zéro seulement si la plante dépasse les 10m (arbre).

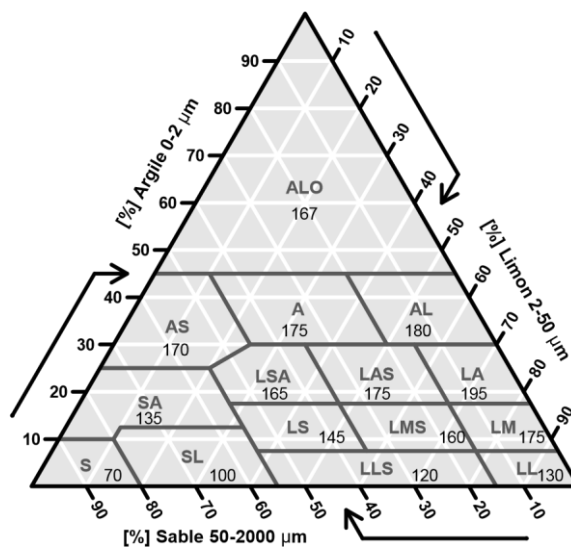


Figure I.8 : Triangle des textures de la carte des sols du département de l'Aisne (France) et réserve utile maximum associée (en mm d'eau pour 1m de sol), d'après Jamagne, (1969) et Baize and Jabio, (1995)

Avec ALO Argile lourde, A Argile, AL Argile limoneuse, AS Argile sableuse, LA Limon argileux, LAS Limon argilo-sableux, LSA Limon sablo-argileux, SA Sable argileux, LM Limon moyen, LMS Limon moyen sableux, LS Limon sableux, SL Sable limoneux, S Sable, LL Limon léger, LLS Limon léger sableux.

Le potentiel hydrique représente donc la capacité de l'eau à quitter un compartiment donné pour un autre qui a un potentiel hydrique plus faible. Par exemple, si une cellule turgescente ayant un potentiel $\Psi_{w.cell} = -0.244\text{MPa}$ est placée dans une solution de glucose à $\Psi_{w.sol} = -0.732$ alors de l'eau va sortir de la cellule jusqu'à ce que la valeur de $\Psi_{w.cell} = \Psi_{w.sol}$. Cette théorie peut être appliquée à chaque compartiment, sol, cellule, atmosphère, etc...

3.2.1.2 L'eau dans le sol

Le sol est un mélange de particules solides organiques ou minérales qui proviennent, soit de la dégradation d'êtres vivants, soit de la dégradation de la roche mère. La taille et le type de particules définissent la texture du sol. La porosité entre les particules est remplie d'air ou d'eau et donne au sol des propriétés physiques propres à chaque type de sol. La texture et le type de sol sont souvent définis par le triangle des textures, qui indique pour chaque côté le pourcentage de sable, de limon et d'argile [Figure I.8](#). Par exemple, en bas à droite du triangle des textures il y a le sable, qui est fait de particules de roche mère de 1mm et plus. Les sols sableux ont une surface par gramme de sol relativement faible, donc de grandes porosités qui ne retiennent pas l'eau. A l'opposé, un sol qui est majoritairement composé d'argile (particule de moins de 2 μm) a une surface très importante et donc présente une petite porosité. Avec l'aide de l'humus (décomposition de la matière organique) les particules d'argiles forment le complexe argilo-humique qui aide à l'aération du sol, l'infiltration et la conservation de l'eau.

D'un point de vue agronomique et lorsque l'on s'intéresse au suivi de l'eau, le sol est souvent défini comme un réservoir d'eau pour la plante. Savoir combien d'eau peut contenir au maximum un sol est l'une des premières questions que doit se poser un agriculteur qui ne veut pas que sa culture manque d'eau. Ce paramètre est communément appelé l'humidité à la capacité au champ (HCC). Il est caractérisé par le nombre de mm d'eau qui peut être contenu dans 1m de sol grâce à la force de capillarité. Au-delà de l'HCC, le sol est à saturation et l'eau percole du fait de la gravité, si cela est possible. L'humidité au point de flétrissement permanent (HPFP) est quant à elle l'opposé de la capacité au champ. C'est le point d'humidité du sol qui ne permet plus à la plante d'extraire l'eau du sol, du fait de la trop grande force de capillarité et d'adsorption, c'est-à-dire que le potentiel hydrique du sol est supérieur à celui de la plante. Pour connaître la réserve utile (RU) d'eau que la plante peut exploiter, il faut connaître les points précédents, ainsi que la densité apparente (Da) et la profondeur d'exploration du sol par les racines (Z). La RU sera donc définie par la formule suivante $RU = (HCC - HPFP) \times Da \times Z$. Elle est spécifique de chaque type de sol. La RU pour un mètre de profondeur d'enracinement pour différents types de sol est présentée dans la [Figure I.8](#) du triangle des textures. Suivant Allen et al., (1998), la réserve utile peut être séparée en deux parties : une réserve d'eau facilement utilisable (RFU) par la plante où le potentiel hydrique de l'eau dans le sol est inférieur à celui de la plante et une réserve difficilement utilisable (RDU) où le potentiel hydrique est supérieur à la plante, mais pour laquelle elle va mettre en place différents processus physiologiques

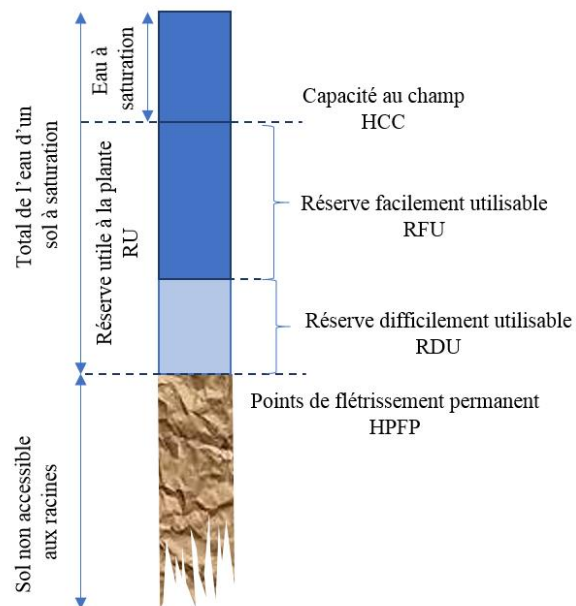


Figure I.9 : Les différents réservoirs d'eau d'un sol, adapté de Bouffier, (2014)

qui lui permettront quand même d'extraire de l'eau. C'est à partir de ce moment-là que la plante commence à entrer en phase de stress. Une représentation des différentes couches du sol en tant que réservoir d'eau est présentée dans la [Figure I.9](#). Lebon et al., (2003) introduisent la notion de fraction d'eau du sol utilisable par la plante (FTSW pour *Fraction of transpirable soil water* en anglais), qui est le rapport entre la quantité d'eau du sol utilisable à un instant donné (RU) et la quantité totale d'eau lorsque le sol est à son état maximum d'hydratation. Elle indique le pourcentage d'eau disponible pour la plante.

3.2.1.3 Les mouvements d'eau dans la plante

Nous avons vu que les mouvements d'eau à travers le continuum sol, plante et atmosphère (CSPA) étaient régis par le potentiel hydrique, exprimé en MegaPascal et représenté par Ψ_w . Cependant, si la différence de potentiel hydrique entre les trois composantes du CSPA est facilement appréhendable, qu'en est-il des mouvements d'eau au sein de la plante ?

Pour que l'absorption de l'eau par les racines soit optimale, il est impératif qu'elles soient en contact direct avec le sol. Ce contact est maximisé par la croissance de racines primaires et secondaires qui explorent un volume de sol qui est de dimension au moins aussi importante que le volume de la plante en surface. De plus, les poils absorbants ou racinaires participent fortement à augmenter cette surface. Les poils absorbants sont des cellules spécialisées de l'épiderme des racines. Du fait de leur excroissance et de leur paroi fine et hydrophile, ils sont le point d'entrée préférentiel de l'eau dans la racine. A partir du moment où l'eau entre en contact avec les cellules de l'épiderme de la racine, il existe trois voies pour alimenter les autres cellules : (1) une voie dite apoplastique, où l'eau circule à travers les parois cellulaires sans jamais passer à travers les membranes des cellules ; (2) une voie dite transmembranaire ou transcellulaire, où l'eau passe d'une cellule à l'autre en traversant les parois ; (3) une voie symplastique, où l'eau passe d'une cellule à l'autre sans sortir du cytoplasme *via* les plasmodesmes. Il semble que la voie principale de l'eau à travers les cellules de la racine soit par la voie apoplastique (Frensch et al., 1996; Steudle and Frensch, 1996). Finalement, l'ensemble des voies de circulation de l'eau converge vers les vaisseaux du xylème.

Chez les plantes de plus d'un mètre, plus de 99.5% de l'eau est transporté par le xylème (Taiz and Zeiger, 2002). En comparaison avec la circulation de l'eau dans la racine, le transport *via* le xylème est un processus relativement simple. En effet, le xylème est seulement constitué de parois de cellules mortes (les trachéides), donc l'eau n'a pas besoin de traverser la membrane des cellules, il en résultera une faible résistance dans le transport de l'eau. Une faible différence de potentiel entre le point de départ de l'eau (les racines) et le point d'arrivée (feuilles) permet la circulation de l'eau.

Une fois que l'eau a atteint les cellules de la feuille, *via* le xylème, elle les traverse pour se retrouver sous forme de vapeur d'eau dans les espaces intercellulaires. Le transfert de la vapeur d'eau entre les espaces intercellulaires et l'atmosphère est la transpiration. Elle est régie suivant deux facteurs

majeurs : la différence de concentration de la vapeur d'eau entre la feuille et l'atmosphère, et la résistance provoquée par les stomates. En effet, c'est par les stomates que l'eau terminera son cheminement vers l'atmosphère. Les stomates ont pour rôle de maîtriser les pertes d'eau de la plante, il a été estimé que seulement 5% des pertes d'eau se fait à travers la cuticule des feuilles [Figure I.7](#). En contrôlant la turgescence des deux cellules de garde qui forment le stomate, l'ostiole est plus ou moins ouvert et permet d'ajuster la quantité d'eau transpirée.

3.2.2 Caractérisation des flux d'eau au cours du cycle d'une culture grâce au modèle de culture CHN

Au cours du cycle d'une culture, les flux d'eau et les éventuels manques d'eau qui induisent un stress hydrique, sont probablement l'un des facteurs les plus complexes à caractériser. En 1998, Allen et al., (1998) ont proposé un guide complet pour rendre compte des flux d'eau à l'échelle d'une culture et suivre ceux-ci jours après jours. Ils définissent ainsi la notion de bilan hydrique. Si les notions d'évapotranspiration maximale et réelle sont toujours d'actualité, la méthode de calcul a pu être perfectionnée grâce au développement de modèles de culture qui vont prendre en compte un nombre de paramètres plus important et surtout suivre l'évolution de la culture de façon plus fine.

Ces modèles de culture sont utilisés pour décrire la croissance et le développement d'une plante en fonction des conditions pédoclimatiques et de l'itinéraire de culture. Deux principaux objectifs leur sont attribués : une utilisation comme outil d'aide à la décision, pour évaluer les conséquences d'une modification de l'itinéraire technique sur la culture et une utilisation plus orientée vers la recherche, où les modèles de culture permettent une meilleure compréhension des relations plante/environnement/conduite de culture (ici la conduite de la culture peut être la possibilité ou non d'irriguer).

Les premiers travaux d'implémentation de fonctions mathématiques décrivant la croissance d'une plante ont été réalisés par de Wit (1965). Depuis, l'ensemble des modèles de culture développés ces dernières années ont des niveaux d'intégration et de précision différents, qui dépendent principalement de leurs objectifs. Ils vont de la prédiction de l'apparition ou de la mort d'organes de la plante (prédiction des stades phénologiques de la plante), à la description de la qualité du grain (Martre et al., 2006), en passant par l'intégration d'effet de gènes ou de QTL affectant un processus physiologique (White and Hoogenboom 1996; Reymond et al 2003; Bogard et al., 2014) et par la modélisation des effets d'un stress biotique et abiotique (maladie, stress thermique, hydrique ou azoté).

Plusieurs modèles de culture ont été développés et rendus accessibles à l'ensemble de la communauté scientifique : Sirius (Jamieson et al., 1998), APSIM (Holzworth et al., 2014a), CHN (Le Bris et al., 2016), STICS (Brisson et al., 1998)... L'objectif ici n'est pas de faire une description fine des différentes fonctions mathématiques utilisées pour simuler la croissance du blé, mais de présenter le

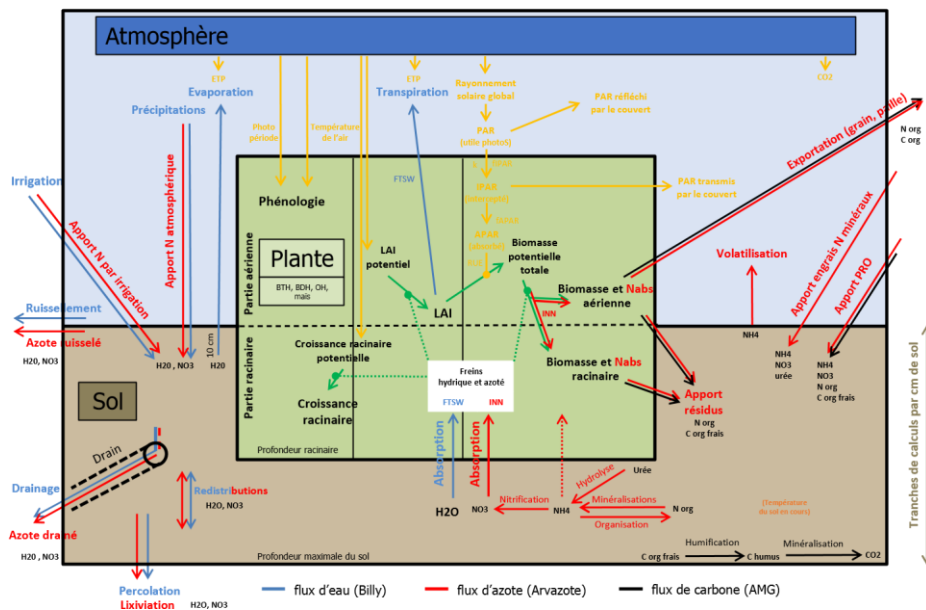


Figure I.10 : Synoptique générale du modèle CHN, d'après Le Bris et al., (2016)

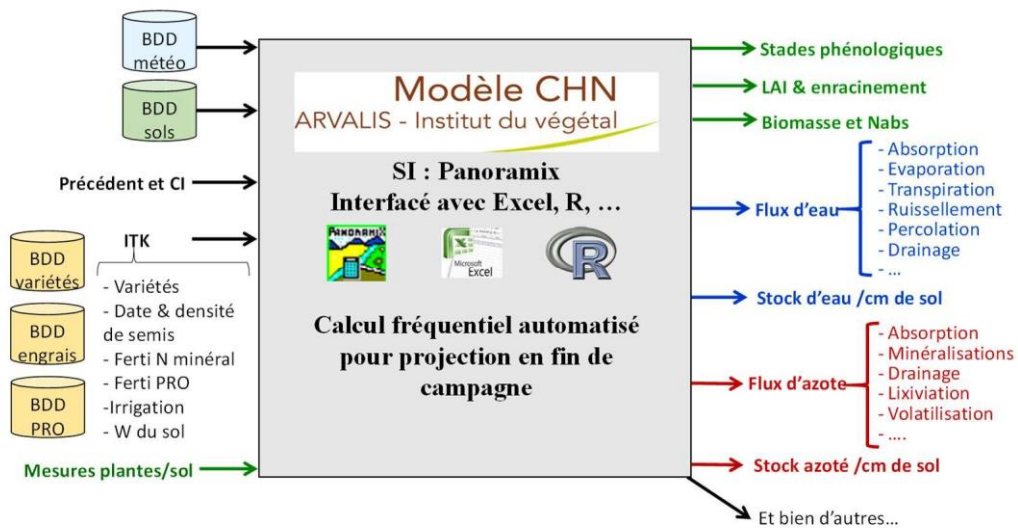


Figure I.11 : Variables d'entrées et de sorties du modèle CHN, d'après Le Bris et al., (2016)

principe de base des modèles de culture en s'appuyant sur CHN qui a été développé par Arvalis et qui a été utilisé dans les différents travaux présentés dans ce manuscrit.

3.2.2.1 Les variables d'entrées de CHN

CHN est un modèle de culture mécaniste qui prédit les flux journaliers de carbone (C), d'eau (H) et d'azote (N) modélisés dans les trois compartiments du CSPA (c'est-à-dire dans le sol, la plante et l'atmosphère). Ce modèle a été développé par Xavier Le Bris en priorité pour les céréales à pailles et le maïs (Le Bris et al., 2016) et il se base sur les principes physiologiques des céréales décrits par Gate (1995), Meynard (1997) et Chatelin et al. (2005). La synoptique du modèle est présentée dans la [Figure I.10](#). Les variables d'entrées, de sorties et les modèles d'établissement des flux d'azote et d'eau au sein du compartiment plante et du CSPA seront décrits par la suite.

Les variables d'entrées de ces modèles de culture sont réparties en 4 grandes classes ([Figure I.11](#)) : (1) les variables météorologiques, dont les plus classiques sont la température, les précipitations, le rayonnement et l'évaporation potentielle de Penman-Monteith (ETP) ; (2) les variables décrivant les caractéristiques du sol et des horizons pédologiques (profondeur, granulométrie...), et son état en cours de campagne (humidité, reliquat d'azote); (3) les variables liées aux interventions culturales, par exemple fertilisation, irrigation, travail du sol ; (4) les variables liées à l'espèce et/ou la variété, par exemple gamme de précocité, architecture aérienne ou encore des mesures prises en cours de campagne telles que la date de stade, les mesures de biomasse ou les composantes de rendement.

Chaque modèle de culture va intégrer de façon plus ou moins poussée des connaissances (formalismes, nombre de paramètre) qui permet de modéliser ou non des phénomènes que l'on souhaite simuler. Par exemple, dans CHN, on peut forcer le modèle pour que les stades de développement prédits soient égaux aux stades de développement observés.

3.2.2.2 Les variables calculées par CHN

Pour bien comprendre le calcul du suivi des flux d'eau dans CHN, il est important d'aborder des notions de développement de la plante dans des conditions optimales et limitantes. Ces suivis de flux vont passer par la détermination de différentes variables d'état de la plante comme l'évolution de la surface des feuilles ou l'évolution de la biomasse.

3.2.2.2.1 Le LAI

L'indice de surface foliaire (LAI : pour leaf area index) est le rapport entre la surface d'une face des feuilles sur une surface de sol (Chen and Black, 1992). Cet indice peut varier de 0 à plus de 6. Il est égal à 0 s'il n'y a pas de feuille sur une surface donnée, et il est supérieur à 6 si sur 1cm² de sol, il y a six feuilles qui se superposent et recouvrent l'ensemble de la surface du sol.

Le modèle d'évolution temporelle du LAI potentiel est exprimé en fonction du temps thermique et combine un modèle logistique pour la phase de croissance et un modèle exponentiel négatif pour la sénescence :

$$\text{Eq 1}_{[\text{CHN}]} : \quad LAI_{pot}(ST) = LAI_{Max} \times \frac{1}{1+e^{-A.(ST-ST_1)}} - e^{B.(ST-ST_2)}$$

avec LAI_{Max} surface maximale sans sénescence (qui est égal à 6.7), A la vitesse maximale de croissance, ST somme de température, ST_1 et ST_2 sont des paramètres homogènes à une somme de température, ST_1 contrôle le moment de croissance maximale, ST_2 le moment de décroissance maximale, B la vitesse de sénescence, ST_2 somme de température à la récolte.

A partir du LAI potentiel et en faisant intervenir les facteurs de stress azotés ou hydriques susceptibles d'influencer le développement de la plante, le LAI réel est défini par :

$$\text{Eq 2}_{[\text{CHN}]} : \quad LAI(j) = LAI(j-1) + \Delta LAI_{pot} \times \min(\text{FacteurH}, \text{FacteurN})$$

avec $\Delta LAI_{pot} = LAI_{pot}(j) - LAI_{pot}(j-1)$, FacteurH et FacteurN qui sont définis dans le paragraphe [1.3.2.2.3.](#)

3.2.2.2.2 La biomasse totale

La biomasse totale potentielle est quant à elle déterminée par jour suivant le formalisme classique de Monteith, (1977) qui définit la biomasse potentielle (en g) suivant trois principaux facteurs : la capacité de la plante à convertir l'énergie lumineuse en biomasse, le rayonnement solaire (ou global) et le LAI que nous avons défini précédemment.

$$\text{Eq 3}_{[\text{CHN}]} : \quad \text{Biomasse}_{pot} = RUE \times APAR$$

avec RUE l'efficacité de conversion du rayonnement en biomasse aérienne sèche utilisée dans le modèle de Kumar-Monteith (g/MJ). Pour des plantes en C3, comme le blé, ce coefficient peut prendre des valeurs entre 2 et 2.5 (Gate, 1995), dans CHN, la valeur 2.5 a été sélectionnée pour modéliser la RUE du blé tendre et du blé dure. Seule une partie du rayonnement global (RG) est interceptée par les plantes. Cette fraction, comprise entre 400 et 700nm, est appelée le PAR (Photosynthetically Active Radiation). Elle varie bien évidemment en fonction de l'heure de la journée et de l'époque de l'année. Cependant, Varlet-Grancher et al., (1982) l'ont considérée comme étant une constante égale à 0.48, donc on peut écrire la relation $PAR = 0.48 \times RG$. Enfin, la relation entre le PAR et le LAI peut être décrite grâce à l'introduction de la fraction absorbée et interceptée de la lumière (FaPAR et FiPAR). Cette fraction absorbée, fAPAR (ou Ea) qui est égale à $0.95 \times fIPAR$. $fIPAR$ est la fraction interceptée et elle dépend du LAI et du coefficient d'extinction K . Le coefficient d'extinction dépend des propriétés architecturales du couvert en lien avec sa capacité à intercepter du rayonnement (port de

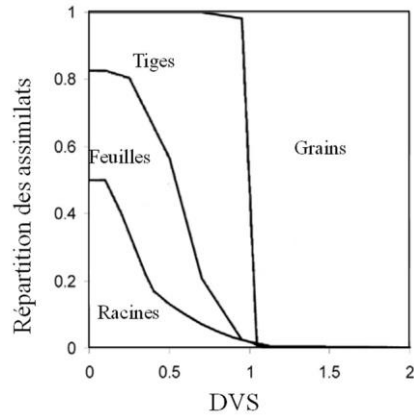


Figure I.12 : Répartition classique des assimilats envers les différents organes pendant le cycle des céréales, par Savary and Willocquet, (2014)
DVS est une échelle de temps avec levée = 0, floraison = 1 et maturité = 2.

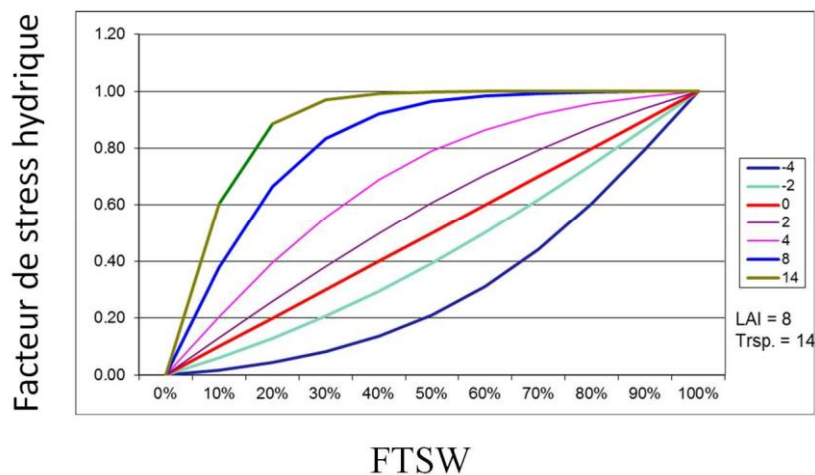


Figure I.13 : Représentation des différentes formes des relations entre Facteur de stress hydrique et l'eau dans le sol d'après les équations présentés par Sinclair (1986)

feuilles érectophiles ou planophiles) (Gallo and Daughtry, 1986; Baret and Guyot, 1991; Gallo et al., 1993). A noter que c'est le LAI réel et non le LAI potentiel qui est utilisé dans le calcul de fIPAR.

Comme pour le LAI, la biomasse aérienne réelle est calculée en fonction de la biomasse potentielle corrigée d'un facteur stress azoté ou hydrique suivant la formule :

$$\text{Eq 4}_{[\text{CHN}]} : \quad \text{Biomasse}(j) = \text{Biomasse}(j - 1) + \Delta \text{Biomasse}_{\text{pot}} \times \min(\text{FacteurH}, \text{FacteurN})$$

avec $\Delta \text{Biomasse}_{\text{pot}} = \text{Biomasse}_{\text{pot}}(j) - \text{Biomasse}_{\text{pot}}(j - 1)$, FacteurH et FacteurN qui sont définis dans le paragraphe [1.3.2.2.3](#).

A partir de la biomasse totale, CHN calcule la biomasse aérienne et racinaire. Cette répartition varie considérablement tout au long du développement de la plante ([Figure I.12](#)). Dans les modèles de cultures APSIM et STICS la répartition de biomasse totale en biomasse aérienne et racinaire débute à 50-50 et diminue au cours du temps. La profondeur d'enracinement est fonction de la somme des températures de l'aire et modifiée par un facteur de stress hydrique ou azoté.

3.2.2.2.3 Les fonctions de stress

Comme nous avons vu précédemment, les facteurs de stress azotés ou hydriques permettent de pénaliser ou non le LAI et la biomasse réelle par rapport au potentiel. Ces facteurs varient de 0 à 1, si le facteur de stress est égal à 1 alors la plante n'est pas limitée dans sa croissance, et au contraire, si le facteur de stress est égal à 0 alors la plante ne se développe plus. Les facteurs de stress hydriques et azotés sont calculés de façon indépendante et seul le facteur le plus limitant participe à diminuer le développement de la plante, suivant la loi des facteurs limitants de Liebig (1850).

Le niveau de stress hydrique pour la plante est exprimé comme une fonction de la teneur en eau du sol suivant les travaux de Sinclair, (1986) ([Figure I.13](#)). Cette approche met en relation le facteur stress hydrique de la plante en fonction de la fraction d'eau transpirable dans le sol (FTSW) suivant la formule :

$$\text{Eq 5}_{[\text{CHN}]} : \quad \text{FaHydrique} = F(\text{FTSW}, k)$$

Avec F la fonction développée par Sinclair, (1986), cette fonction de stress hydrique est spécifique au LAI, à la biomasse, à la transpiration et à l'enracinement (profondeur du front racinaire) au travers d'un paramètre k ajusté spécifiquement pour chacune de ces 3 variables. FTSW la fraction d'eau transpirable dans le sol et k un facteur déterminé.

Le niveau de stress azoté quant à lui est déterminé par la formule suivante :

$$\text{Eq 6}_{[\text{CHN}]} : \quad \text{FaAzote} = (\text{INN} - \text{INN}_{\text{min}}) / (1 - \text{INN}_{\text{min}})^p$$

où INN (Indice de nutrition azotée) et INN_{min} sont déterminés par espèce.

Dans le modèle CHN, aucun facteur de stress de température n'est pris en compte, en dehors du fait que les degrés jours sont accumulés seulement à partir d'un certain seuil (0°C pour le blé). Aucun facteur n'est pris en compte quant à la présence de températures inférieures à 0 ou bien de températures échaudantes (températures élevées stressantes) comme c'est le cas dans d'autres modèles de culture comme SIRIUS ou APSIM.

3.2.2.2.4 Les flux d'eau

L'état de la réserve en eau du sol implémenté dans CHN est calculé sur un pas de temps journalier. Pour initialiser l'état de la réserve en eau au semis, le calcul du bilan hydrique est réalisé non seulement sur l'ensemble de la culture mais aussi en amont de celle-ci. Les flux d'eau font l'objet de plusieurs échanges entre les différents compartiments du CSPA.

Dans le sol, le bilan hydrique est calculé par tranche d'un centimètre et une redistribution de l'eau par capillarité est réalisée pour chaque tranche toutes les nuits. Il fait aussi intervenir la notion de fraction d'eau transpirable du sol précédemment présenté. L'état de la réserve en eau du sol au jour j (ou bilan hydrique) est calculé grâce aux simples addition et soustraction des quantités d'eau qui entrent et sortent du système. Les apports d'eau résultent du cumul des pluies et des irrigations. Le niveau de pluie par jour est limité à un plafond et les pluies sont transformées en pluies efficaces par un coefficient d'infiltration. Le plafond et le coefficient dépendent tous deux du type de sol. En ce qui concerne les sorties d'eau, elles font intervenir le ruissellement, la percolation en profondeur, le drainage et l'évapotranspiration. Le ruissellement est la différence entre la pluie et la pluie efficace. De plus, il est considéré qu'une partie de l'eau ruisselle (80%), notamment entre capacité au champ et saturation en eau. Concernant la percolation, il y a redistribution de l'eau par tranche de 1cm selon le modèle suivant : si l'humidité de la tranche est supérieure à l'HCC, alors l'eau s'achemine dans la tranche suivante. Si les champs sont drainés, alors un module est pris en compte dans le modèle. Enfin, l'évapotranspiration est séparée en deux dans CHN, un module d'évaporation et un module de transpiration.

Le calcul de l'évaporation est réalisé sur une profondeur de 10cm sans compter la redistribution de l'humidité. L'évaporation potentielle de Penman-Monteith (ETP) est issue des données météorologiques et suivant les recommandations de la FAO, elle est transformée par un facteur 1.05 pour les sols nus. Les jours de pluie, l'évaporation potentielle est considérée égale à l'évaporation potentielle quelle que soit la réserve hydrique. Dans tous les autres cas, elle est multipliée par la conductivité et la contribution des différentes tranches de sol.

La transpiration quant à elle répond à la notion de l'offre et de la demande. L'offre est modélisée par la quantité d'eau présente dans le sol au cours du cycle de la plante. La demande est quant à elle modélisée par le calcul de la transpiration maximale réalisée par la plante, selon les conditions climatiques et le stade de la culture, avec la formule :

$$\text{Eq 7}_{[\text{CHN}]} : \quad \text{Transpiration}_{\text{Max}} = 1.2 \times \text{ETP} \times \text{Ea}$$

Avec *ETP* l'évaporation potentielle de Penman-Monteith et *Ea* l'efficacité d'absorption du couvert (*fAPAR*) définie *via* le modèle classique faisant intervenir le LAI et la sénescence (Lecoeur et al. 2002). L'extraction de l'eau par les racines (= transpiration) est modulée par la césibilité de l'eau au niveau des tranches de sol. Si l'offre est inférieure à la demande, alors la plante ne transpire qu'une partie de son potentiel, il y a donc une différence entre la transpiration maximum et la transpiration réelle :

$$\text{Eq 8}_{[\text{CHN}]} : \quad \text{Transpiration}_{\text{réel}} = \text{Transpiration}_{\text{Max}} \times \text{FaHydrique}$$

La combinaison de l'évaporation et de la transpiration va permettre de définir l'évapotranspiration maximale et réelle de la culture. Ces notions ont été introduites par Allen et al., (1998). Tout ceci va rendre possible la réalisation du bilan hydrique tout au long du cycle de la culture et au final permettre de construire un scénario pour identifier potentiellement les fréquences et l'intensité des stress hydriques.

3.2.3 Impact du stress hydrique sur la physiologie de la plante

Comme nous l'avons vu précédemment, c'est l'ensemble de l'accumulation de la biomasse pendant la période végétative de la plante qui va participer à l'établissement du rendement final. Ainsi, chaque événement de stress qui diminuerait le potentiel de biomasse de la plante impactera le rendement. Plusieurs études ont reporté une diminution du nombre de cellules et de leurs tailles en condition de stress hydrique (Westgate and Boyer, 1985; Tardieu, 2006), donc les organes et le cumul de la biomasse de la plante sont diminués par des manques d'eau. De plus, dans des conditions de stress hydrique la plante va plus rapidement initier sa phase de sénescence et augmenter la vitesse de celle-ci (Christopher et al., 2016), ce qui conduit à une diminution du cycle donc une diminution de la biomasse et une diminution de la durée de remplissage des graines.

Comme pour le stress thermique, tous les organes de la plante peuvent être impactés par un manque d'eau. Cependant, s'il y a peu de chance qu'un stress thermique impacte la mise en place des talles dans un contexte de culture du blé en région tempérée, il est bien plus probable qu'un stress hydrique survienne à ces périodes, même de façon ponctuelle. Ainsi, tous les organes les plus sensibles de la mise en place du rendement peuvent être touchés de façon précoce. Par exemple, au cours de la méiose du blé, un stress hydrique peut avoir un impact significatif sur le développement des gamétophytes (Fischer and Maurer, 1978; Sheoran and Saini, 1996). Cependant, il est difficile de clairement classer les organes par rapport à leur sensibilité au stress. Par exemple, Moss and Downey, (1971) ont montré que l'impact le plus important de la sécheresse sur le nombre de grains venait d'une diminution de la fertilité des gamètes mâles, plus que d'une diminution de la fertilité des gamètes femelles. D'autres ont indiqué que la sécheresse a eu des impacts plus importants sur le développement

des sacs-embryons du maïs et du blé que sur le pollen, mais le contraire a été observé avec le stress thermique (Westgate and Boyer, 1985; Prasad et al., 2011). Enfin, dans leurs études Saini and Aspinall, (1981) décrivent un stress hydrique juste avant la floraison qui serait responsable de 40% des avortements des grains, donc des grains qui sont bien fécondés mais ne sont pas assez bien alimentés pour continuer leurs phases de division et de remplissage.

Le remplissage du grain est aussi impacté par un manque d'eau. Comme nous l'avons vu précédemment, le fait de raccourcir le cycle en accélérant la phase de sénescence et donc la phase de remplissage va induire une mise en réserve moins importante et donc une diminution de la taille des grains. Cette réduction peut venir aussi bien d'une réduction de la durée de remplissage (Brooks et al., 1982) que de la vitesse (Brocklehurst et al., 1978). De plus, un stress hydrique pendant le remplissage des grains induit une diminution du nombre de cellules de l'endosperme et donc diminue la capacité du grain à accumuler des réserves (Nicolas et al., 1985; Ober et al., 1991). L'état hydrique des grains, pendant la phase de remplissage, semble être moins affecté que les autres organes (feuilles et organes floraux) (Barlow et al., 1983; Ober et al., 1991). Pourtant, cela n'empêche pas que sous un stress hydrique les grains sont aussi impactés et des grains en condition de stress vont passer plus rapidement de la phase de remplissage à la phase de dessiccation, ce qui entraîne un déclin rapide du potentiel en solutés (Ψ_s) et du potentiel hydrique ((Ψ_w)) (Westgate and Boyer, 1985). Cette perte d'eau rapide pourrait inhiber les événements métaboliques impliqués dans le remplissage et dans la composition protéique. En effet, dans leur étude (Zhao et al., 2009) ont montré que la quantité et la composition en protéines sont impactées par un événement de stress hydrique.

Enfin, il ne faut pas oublier l'impact du stress hydrique sur le développement des racines. En effet, en plus d'être la seule source d'acquisition de l'eau et des minéraux, les racines peuvent être impactées par une sécheresse du sol, ce qui ne fait qu'accentuer le phénomène de sécheresse des parties aériennes. L'impact d'événements de sécheresses peut être multiple. Bien évidemment et au même titre que pour la biomasse aérienne, il y a une réduction de la multiplication cellulaire et de la croissance des cellules en cas de manque d'eau. De plus, lors d'une sécheresse sévère, souvent sur des sols argileux, il est fréquent d'observer des fentes de dessiccation et/ou le compactage du sol qui peuvent freiner la croissance des racines, voir les endommager (Lucas et al., 2000). Les propriétés du sol vont donc changer en fonction de l'intensité de la sécheresse et la multiplication d'espaces remplis d'air va diminuer la surface de contact entre les racines et l'eau. Il est aussi possible qu'un assèchement en surface rende l'enracinement des racines nodales plus difficile. En comparant des variétés de blé dans un essai à deux modalités irriguée et non irriguée, Barraclough & Weir 1988 ont montré que la croissance des racines était réduite dans des conditions de stress hydrique, mais que l'absorption d'eau était la même jusqu'à la floraison. Il est donc possible que les plantes réduisent leur croissance racinaire sans pour autant impacter la nutrition en eau des parties aériennes.

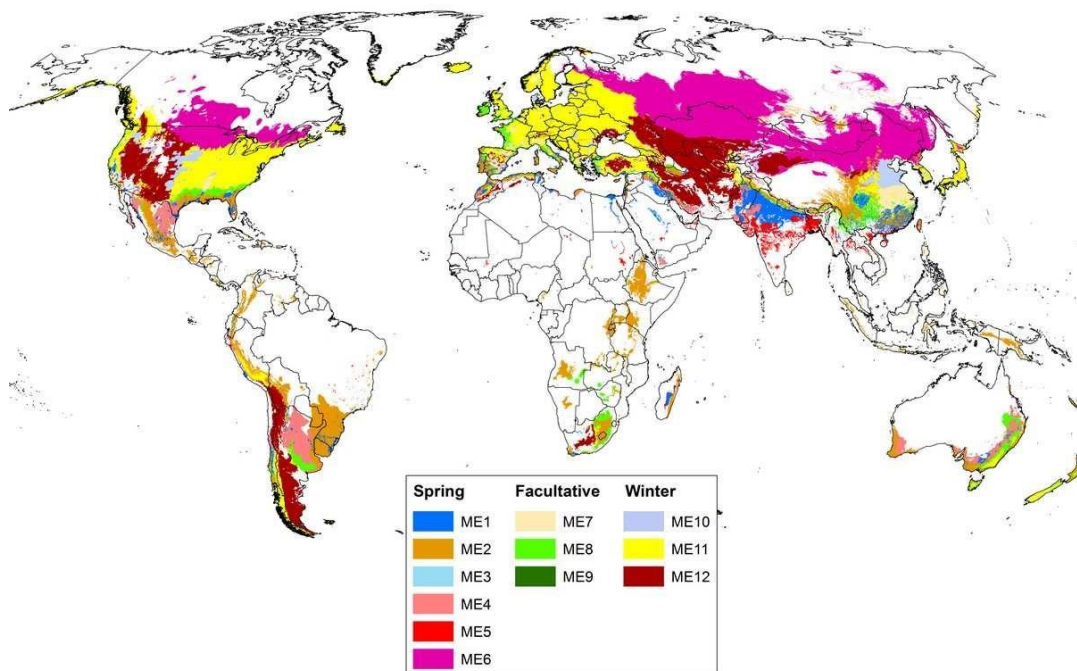


Figure I.14 : Répartition des productions et des méga-environnements définis par le CIMMYT d'après Sonder, (2018)

4 Les différences d'adaptation du blé en réponse aux stress hydriques et thermiques

Le blé est aujourd'hui cultivé sur un large panel de conditions pédoclimatiques, donc potentiellement de stress hydriques et thermiques différents (Braun et al., 2010). Ces deux stress ont évidemment tendance à s'appliquer de façon concomitante dans les régions semi-arides et chaudes comme le bassin méditerranéen, l'Argentine, le Mexique, l'Australie et l'Afrique du sud. Pour cibler le type de germoplasme adapté aux différentes régions du monde, Rajaram et al., (1995) ont été les premiers à introduire la notion de méga-environnement. A chaque méga-environnement est associé un idéotype variétal, c'est-à-dire une gamme de variétés réunissant des caractéristiques morphologiques et physiologiques qui leur permettent d'exploiter au maximum les ressources qui sont à leur disposition sans s'exposer à des niveaux de stress trop pénalisants pour la mise en place du rendement.

Par la suite, le CIMMYT a compilé et identifié douze méga-environnements (ME) qui réunissent les grands types de conditions pédoclimatiques et de conduites de culture à travers le monde entier ([Figure I.14](#)). Au sein de ces douze ME, cinq sont directement associés à des conditions de culture sous stress hydriques et/ou thermiques : le ME4 représente 21.6 Mha, c'est un ME où les précipitations sont faibles puisqu'en moyenne, il y a moins de 500mm de pluie dans la saison ; le ME5, où les températures sont hautes, couvre 7.2 Mha ; les ME6-9-12 sont, quant à eux, des régions où les stress hydriques et thermiques sont prépondérants et représentent plus de 25% des surfaces de blé cultivées (Sonder, 2018). Même si les ME avec stress hydriques et/ou thermiques définis par le CIMMYT ne sont pas présents en Europe (où de façon très ponctuelle avec l'Espagne et l'Italie), des évènements de stress peuvent toucher ces régions et comme le germoplasme n'est pas, ou très peu, adapté à ce type de stress, l'impact sur le rendement n'en est pas moindre (Brisson et al., 2010).

Dans cette partie, nous allons passer en revue les différents outils et méthodes qui sont à disposition des sélectionneurs pour adapter leur matériel végétal aux environnements où la fréquence de stress hydriques et thermiques est importante. Dans un premier temps, nous allons présenter les traits physiologiques et morphologiques de l'adaptation de variétés, en soulignant l'existence de variabilité génétique pour la sélection et les outils de phénotypages qui sont à disposition. Dans un second temps, nous nous attarderons sur les méthodes statistiques qui sont communément utilisées pour permettre aux sélectionneurs d'identifier leur matériel dans un contexte d'essais multi-environnementaux.

4.1 Les concepts de traits physiologiques associés à l'adaptation des stress hydriques et thermiques

Avant de commencer à présenter comment le blé résiste à ces stress ainsi que les traits physiologiques permettant au blé de tolérer des évènements de stress hydriques ou thermiques, il semble nécessaire de clarifier la terminologie qui s'applique à ces différents processus. Depuis Levitt en 1980, la communauté scientifique s'est accordée à définir comme résistance tous les mécanismes et stratégies

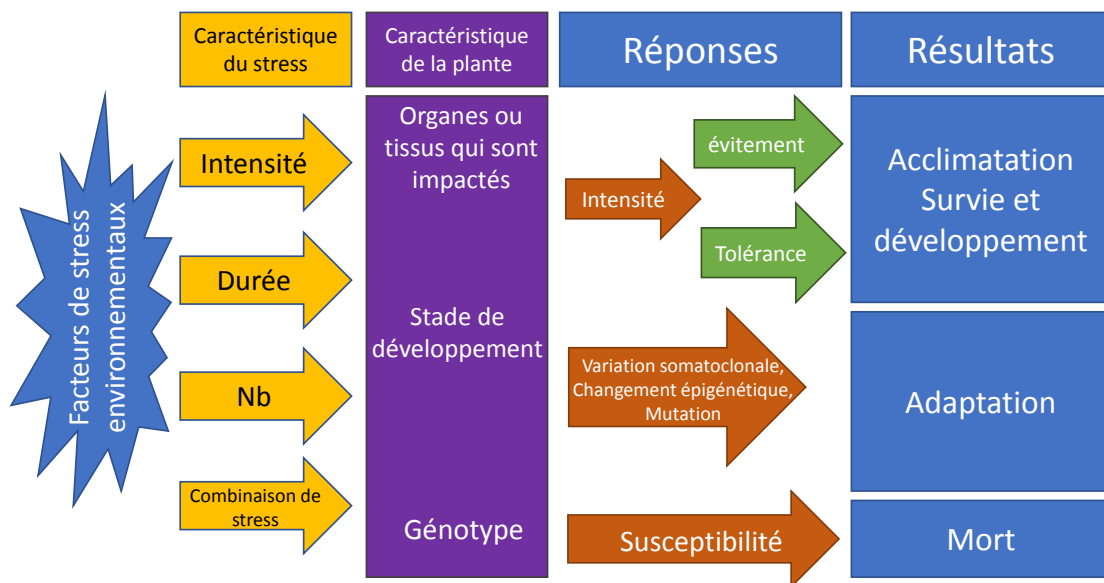


Figure I.15 : Impact d'un stress environnemental sur les plantes, d'après Gaspar et al., (2002)

qui permettent à la plante de survivre dans des conditions stressantes. Ces stratégies peuvent être séparées en trois classes bien distinctes : (1) la première classe est une stratégie dite d'échappement ("escape" en anglais), où les plantes ne sont pas atteintes par le stress car elles ont pu compléter les phases critiques de leur développement avant qu'elles soient impactées par ce stress. Certains auteurs ne considèrent pas cette stratégie comme une résistance à un stress étant donné que les plantes n'ont pas été confrontées à celui-ci. Cependant, à une échelle globale, c'est la stratégie qui est la plus utilisée en agriculture car elle est en lien direct avec l'adaptation des plantes (ou variétés) à la condition climatique de la région d'exploitation. Cependant, comme nous avons vu précédemment, raccourcir le cycle de la plante est aussi synonyme de réduction du rendement. Donc si aucun événement de stress n'arrive en fin de cycle, il y a un manque à gagner pour l'agriculteur puisque la plante aurait encore pu accumuler de la biomasse. (2) la deuxième stratégie est dite d'évitement ("avoidance" en anglais). Ici les plantes vont réduire leurs activités métaboliques grâce à différents processus (*eg.* la fermeture des stomates). En ralentissant son métabolisme, la plante va en quelque sorte économiser de l'énergie et des réserves d'eau dans le sol (si c'est un stress hydrique) pour les valoriser lorsque les conditions redeviendront plus propices. Cette stratégie est efficace seulement s'il y a un retour à des conditions dites « normales » (sans stress). (3) la dernière stratégie est dite de tolérance. Dans ce cas les plantes vont maintenir un métabolisme élevé sous des stress médians et ne réduisent leur métabolisme que sous un stress fort. Cette stratégie est efficace dans un cas de stress terminal, c'est-à-dire que les conditions de culture ne reviendront pas optimales et donc la plante doit exploiter un maximum les ressources de l'environnement pour finir son cycle. Les stratégies et les mécanismes physiologiques qui leur sont liés sont présentés dans la publication de Gaspar et al., (2002), dont la [Figure I.15](#) provient.

Chacune de ces stratégies est plus ou moins adaptée à des conditions de culture et climatiques propres à chaque région. C'est pourquoi les traits physiologiques qui sont responsables de la mise en place de ces stratégies doivent être sélectionnés en fonction de l'environnement qui est ciblé par les sélectionneurs.

4.1.1 Modèle de mise en place du rendement en condition stressante présenté par Passioura, (1977)

La modélisation de l'établissement du rendement en condition de stress hydriques et thermiques, a été présentée la première fois par Passioura, (1977). Cette représentation théorique, largement réutilisée par la suite (Reynolds et al., 2009a et 2009b), a permis de présenter les traits physiologiques de base de l'adaptation des variétés à des environnements stressants. Ce modèle découpe le rendement en condition de stress hydriques en trois composantes :

$$YLD = WU \times WUE \times HI$$

$$\text{YLD} = \text{WU} \times \text{WUE} \times \text{HI} \text{ (stress hydrique)}$$

$$\text{YLD} = \text{LI} \times \text{WUE} \times \text{HI} \text{ (stress thermique)}$$

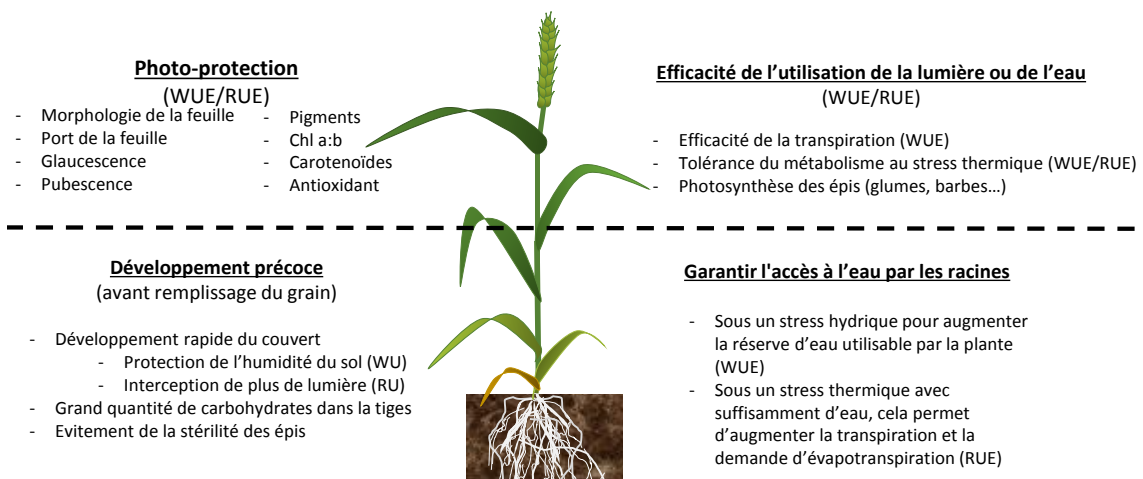


Figure I.16 : Modèle conceptuel représentant les traits physiologiques associés à l'adaptation des plantes à un environnement stressé hydriquement et/ou thermiquement, d'après Reynolds et al. (2017)

Où YLD est le rendement final de la plante, WU est la quantité d'eau absorbée par la plante, WUE est l'efficacité de l'utilisation de l'eau et HI l'index de récolte (Reynolds et al., 2009b).

De la même façon, il est possible de décliner la construction du rendement en condition de stress thermiques, sans stress hydrique suivant trois composantes :

$$YLD = LI \times RUE \times HI$$

Où les termes YLD et HI sont les mêmes que dans l'équation précédente, LI définit la quantité d'interception de la lumière et RUE l'efficacité de l'utilisation des radiations lumineuses. Comme nous l'avons vu précédemment, les modèles de culture ont décliné ces formules pour permettre d'intégrer plus de paramètres climatiques. Cependant, l'objectif de cette partie est de voir comment la sélection peut essayer de manipuler chacune des trois composantes de chaque équation pour augmenter la résistance des variétés à des conditions de stress hydriques et thermiques.

4.1.2 Description des traits les plus connus permettant une adaptation aux stress hydriques et thermiques

La [Figure I.16](#) qui a été adaptée des travaux de Reynolds et al., (2009b) présente les traits physiologiques qui peuvent permettre à la plante de résister à un stress hydrique ou thermique. En complément, le [Tableau I.3](#), qui a été adapté de la publication de Tardieu, (2012), liste les traits physiologiques d'adaptation des variétés aux différents scénarios de stress. En effet, étant donné qu'un trait physiologique peut être favorable dans un scénario bien précis et défavorable dans un autre, il est très utile pour le sélectionneur de bien définir les traits choisis en lien avec l'environnement cible. La publication de Tardieu, (2012) présente d'ailleurs un titre des plus explicite : « *Tout trait ou allèle lié à ce trait peut conférer une tolérance à la sécheresse : il faut juste définir le bon scénario* ». Notez ici que la notion de "tolérance" regroupe les deux termes évitement et tolérance, que nous avons défini précédemment.

Le but de cette partie n'est pas de faire une description complète de chacun des traits qui sont utilisés ou utilisables par les sélectionneurs. Cependant, nous avons présenté les traits les plus communément décrits dans la littérature.

4.1.2.1 Water Use

Garantir un accès à l'eau par les racines permet à la plante d'élever la composante WU pour augmenter la biomasse. Comme nous l'avons vu précédemment, il existe différents types d'architecture racinaire qui vont être plus ou moins adaptés aux conditions climatiques. Par exemple, dans certains environnements, il est courant qu'en fin de cycle, l'eau ne soit accessible qu'en profondeur, alors que dans d'autres environnements, où les pluies sont présentes tout au long du cycle de la plante mais de façon sporadique, c'est l'eau de surface qui doit être valorisée rapidement par la plante. Dans le premier type d'environnement, il est donc préférable que les racines explorent le sol un maximum en

profondeur, alors que dans le second type d'environnement, la multiplication des racines dans les couches supérieures est plus efficace (Lucas et al., 2000).

En plus de l'architecture racinaire, la modulation de la pression osmotique peut jouer un rôle prépondérant dans l'absorption de l'eau en condition de stress. En effet, à partir du moment où le sol atteint un niveau de potentiel hydrique égal au potentiel de la plante, celle-ci doit ajuster son potentiel osmotique pour continuer à absorber de l'eau. C'est à ce moment-là que la réserve en eau passe de facilement utilisable à difficilement utilisable par la plante. L'ajustement osmotique des racines est l'un des processus de tolérance au stress hydrique les plus importants. C'est en accumulant des solutés dans les racines que la plante va augmenter son potentiel hydrique et donc conserver la capacité à absorber de l'eau (Hellebusi, 1976). Dans leur publication Reynolds et al., (2006) présentent l'ajustement osmotique comme un trait difficilement classable en tant que trait de tolérance à la sécheresse ou d'évitement de la sécheresse. En effet, suivant Munns and Weir, (1981), l'ajustement osmotique est une stratégie de tolérance du stress hydrique, puisque le coût énergétique de mise en place des solutés dans les cellules de la racine l'emporte sur les avantages liés à cette stratégie, sauf en cas de stress fort qui nécessite la survie de la plante. Alors que, Morgan, (1991) présente lui l'ajustement osmotique comme une stratégie d'évitement de la sécheresse, étant donné que les cellules continuent de croître.

Pour les sélectionneurs, et de façon générale, les racines ne sont pas la partie de la plante la plus accessible et ne sont donc pas faciles à phénotyper. Cependant, plusieurs études ont montré une variabilité génétique pour l'architecture racinaire et des méthodes de phénotypage haut débit ont été développées (eg. Wasson et al., 2012; Richard et al., 2018). De plus, Reynolds et Trethowan, (2007) ont montré que l'enracinement des racines en profondeur pouvait être approché en mesurant la température de canopée. En effet, en comparant les températures de canopée d'un génotype à l'autre, Olivares-Villegas et al., (2007) et Pinto et al., (2010) ont montré que les génotypes qui ont une température faible sont aussi les génotypes qui ont une bonne performance dans des environnements secs et chauds. S'il est entendu que cette température est liée à une augmentation de l'évapotranspiration, Reynolds et Trethowan, (2007) ont aussi montré que dans des conditions de stress thermiques et sans stress hydrique, les génotypes qui ont une température de canopée basse sont ceux qui transpirent le plus, donc ceux qui ont une meilleure exploitation de la quantité d'eau extraite par les racines.

4.1.2.2 Water and Radiation Use Efficiency, WUE/RUE

Comme nous avons vu précédemment, l'efficacité d'utilisation des rayonnements et l'efficacité d'utilisation de l'eau représentent respectivement l'efficacité avec laquelle une culture transforme l'énergie lumineuse en biomasse et l'efficacité avec laquelle une culture transforme l'eau qui s'est évaporée en biomasse. La quantification du rendement maximal réalisable par unité de ressource

Table I.3 : Tableau récapitulatif de l'effet des traits sous différents stress hydriques, d'après Tardieu (2012)

Trait	Abondance de gènes connus/allèles affectant les traits	Variable à phénotyper	Effet positif	Effet négatif	Scénario où l'effet positif est maximisé	Scénario où l'effet négatif est maximisé
Réduire le cycle de culture	***	Durée du cycle	Echappement : fin du cycle ou des périodes du cycle les plus sensibles avant occurrence du stress	Faible accumulation de biomasse via la photosynthèse	Année très sèche	Fin de cycle avec des conditions favorables
Protection des cellules contres un stress	***	Aspect et biomasse	Controversé, probablement un impact mineur, sauf dans des cas de stress sévères	Controversé et variable	Stress sévère	?
Evitement avec fermeture des stomates	*	Aspect, biomasse, échange gazeux et température du couvert	Garde l'humidité du sol, ↓ gradients hydraulique	↓ la photosynthèse et ↑ température des feuilles	Stress sévère terminal	Fin de cycle avec des conditions favorables
Evitement avec réduction de la surface des feuilles	**	Aspect et biomasse	Garde l'humidité du sol, ↓ gradients hydraulique	↓ photosynthèse	Stress sévère terminal	Fin de cycle avec des conditions favorables
Water Use Efficiency	**	$\Delta^{13}\text{C}$, ratio de biomasse/transpiration		↓ photosynthèse	Stress sévère terminal	Fin de cycle avec des conditions favorables
Maintien de la photosynthèse et de la conductance stomatique	*	Echange gazeux et température du couvert	↑ biomasse	↑ risques de stress en fin de cycle	Stress intermédiaire en milieu de cycle	Stress sévère terminal
Maintien de la croissance végétative	**	NDVI et proxidtection	↑ biomasse	↑ risques de stress en fin de cycle	Stress intermédiaire en milieu de cycle	Stress sévère terminal
↑ croissance des racines	**	Quantification ADN, imagerie	↑ quantité d'eau absorbée	Compétition pour le C: ↑ risques de stress en fin de cycle	Eau disponible dans les couches profondes	Sol peu profond
Architecture racinaire sans changement de biomasse	?	Rhizotrons	↑ quantité d'eau absorbée	↓ nutrition en minéraux	Eau disponible dans les couches profondes ou ponctuellement dans la saison	
Réduction de l'avortement des grains	*	Observation directe du nombre de grains	↑ du rendement	↓ qualité	Stress pendant la floraison et pas terminal après	Stress sévère

disponible est donc essentielle pour aider les sélectionneurs à identifier les traits physiologiques qui permettraient d'augmenter ce potentiel.

Farquhar and Richards, (1984) ont montré que l'efficacité d'utilisation de l'eau des variétés de blé été corrélée à la composition isotopique du carbone. C'est notamment le rapport $^{13}\text{C}/^{12}\text{C}$ qui est utilisé comme proxy de l'efficacité de la transpiration et de la photosynthèse (Reynolds et al., 2006). Le principe de cette méthode repose sur l'ouverture/fermeture des stomates lorsque les plantes subissent un stress hydrique. La mesure de ce rapport sur les feuilles des génotypes qui n'ont pas subi de stress hydrique estime de façon intrinsèque le WUE. Ce trait a d'ailleurs été utilisé comme critère de sélection dans des programmes de sélection australien, sous des stress postfloraison, et il en résulte des inscriptions de variétés avec de grandes efficacités d'utilisation de l'eau (Rebetzke et al., 2002). Dans le projet FSOV Drought-2012, Deswarte et al., (2017) ont étudié la variation de ce paramètre sur un large réseau d'essais de lignées élites françaises. Leurs résultats montrent que ce paramètre permet bien de caractériser les conditions de croissance des plantes (plantes stressées et non-stressées) et aussi de différencier les variétés entre elles. Cependant, dans leur condition de réseau d'expérimentation, il est difficile d'associer la tolérance des variétés à ce paramètre. En effet s'il existe bien une corrélation avec le rendement, il s'avère que les plantes les plus performantes sont les plantes dont l'efficacité de l'utilisation de l'eau est la plus faible. En effet, toute la difficulté dans un réseau multi-environnemental et de définir clairement les types de stress et les traits qui peuvent être favorables à la plante. Comme nous l'avons vu précédemment Tardieu, (2012) a fait une revue de ces traits et de leurs avantages/inconvénients en fonction des scénarios de stress. Il résume l'ensemble des effets dans un tableau ([Table I.3](#)) et présente effectivement l'efficacité d'évaporation de l'eau comme un trait favorable dans certains scénarios de stress et défavorable dans d'autres.

En plus de l'efficacité de la transpiration, la physiologie, la morphologie et le port des feuilles jouent un rôle très important dans de nombreux processus de tolérance. En effet, lorsque les plantes subissent un stress hydrique ou thermique, elles sont exposées à une forte radiation lumineuse, notamment les UV-B, qui peuvent provoquer des dommages aux cellules (Alexieva et al., 2001). Ainsi, la pigmentation des feuilles, la glaucescence, la pubescence, le port des feuilles, la capacité à l'enroulement peuvent jouer un rôle dans la limitation de l'élévation de température et de l'impact de radiations fortes (Reynolds and Trethowan, 2007). Par exemple, (Johnson and Thornley, 1984; Richards et al., 1986) ont montré que la cire sur la cuticule des plantes permettait un refroidissement du couvert végétal par une réflexion des radiations lumineuses et permettait un gain de rendement dans des conditions de stress hydriques et thermiques terminaux. De même que la morphologie des feuilles, la morphologie des épis peut permettre aux plantes une meilleure tolérance à ces stress. Par exemple, al-Shalalkeh and Duwayri, (1986) ont montré que les barbes, grâce à leur activité photosynthétique, pouvaient participer au remplissage du grain.

4.1.2.3 L'indice de récolte

L'indice récolte est le rapport entre le rendement en grain et la biomasse sèche aérienne. Il représente donc la capacité de la plante à remobiliser l'ensemble de la biomasse accumulée tout au long de la saison dans le grain sous forme de protéines. Comme nous avons vu précédemment, en cas de réduction de l'activité photosynthétique, due à un stress, le remplissage du grain vient essentiellement de la remobilisation des réserves *via* les tiges. Une différence variétale pour l'accumulation de sucres solubles avant un stress et la capacité de la plante à remobiliser ces sucres a été identifiée et associée à l'un des mécanismes de tolérance aux stress hydriques ou thermiques (Blum, 1997; Tahir and Nakata, 2005). Dans ces nombreuses études, les auteurs montrent que la robustesse des tiges est liée à la quantité de sucres solubles dans celle-ci. Phénotyper la robustesse des tiges est donc potentiellement un moyen indirect de sélectionner les plantes avec ce caractère.

4.1.2.4 Le Stay-green, un trait favorable en condition de stress hydriques et thermiques

Comme nous avons vu précédemment, la sénescence peut être initiée à la fois par un stress hydrique et un stress thermique. Les plantes présentant un phénotype *stay-green*, c'est-à-dire la capacité à maintenir leur photosynthèse et donc retarder leur sénescence après la floraison (Thomas and Howarth, 2000), ont largement été associées avec une tolérance aux stress (Olivares-Villegas et al., 2007; Pinto et al., 2010; Bouffier, 2014; Christopher et al., 2016). (Górny and Garczyński, 2005) ont reporté une corrélation positive entre le maintien de la surface verte de la feuille drapeau, l'efficacité de l'utilisation de l'eau et l'indice de récolte. Evidemment, ce type de caractéristique a un effet positif dans des scénarios de stress hydriques et thermiques pendant la floraison et/ou post-floraison (Farooq et al., 2011).

Différentes techniques peuvent être utilisées par les sélectionneurs pour discriminer la capacité des plantes à retarder leur sénescence. La première est le suivi visuel de la sénescence avec des notes allant de 0 à 10 (avec 0 totalement vert et 10 totalement sénéscent) : même si elle prend du temps et n'est pas très précise (notamment en début de sénescence), cette technique est suffisante pour discriminer les plantes. D'autres méthodes plus précises utilisent la mesure de la chlorophylle dans la feuille drapeau grâce à l'utilisation d'un SPAD (par exemple Borrell et al., 2000 et Christopher et al., 2008) : cette technique est bien plus précise mais néanmoins pas très facile à implémenter sur de nombreuses variétés, comme dans un programme de sélection. La dernière méthode consiste à mesurer l'indice de végétation par différence normalisée (NDVI). Il est construit à partir des ondes de réflexion rouge et proche infrarouge. Cette mesure peut être réalisée grâce à l'utilisation de GreenSeeker qui permet de mesurer rapidement le NDVI d'un couvert végétal ou même avec l'utilisation d'un drone comme vecteur pour mesurer le NDVI à l'échelle d'un champ.

Peu importe la technique utilisée pour le suivi de la sénescence, il est important de multiplier le nombre de mesures. Dans la littérature, de nombreux modèles permettant d'estimer des paramètres de sénescence ont été utilisés. Ils vont souvent de la simple régression linéaire à des modèles plus

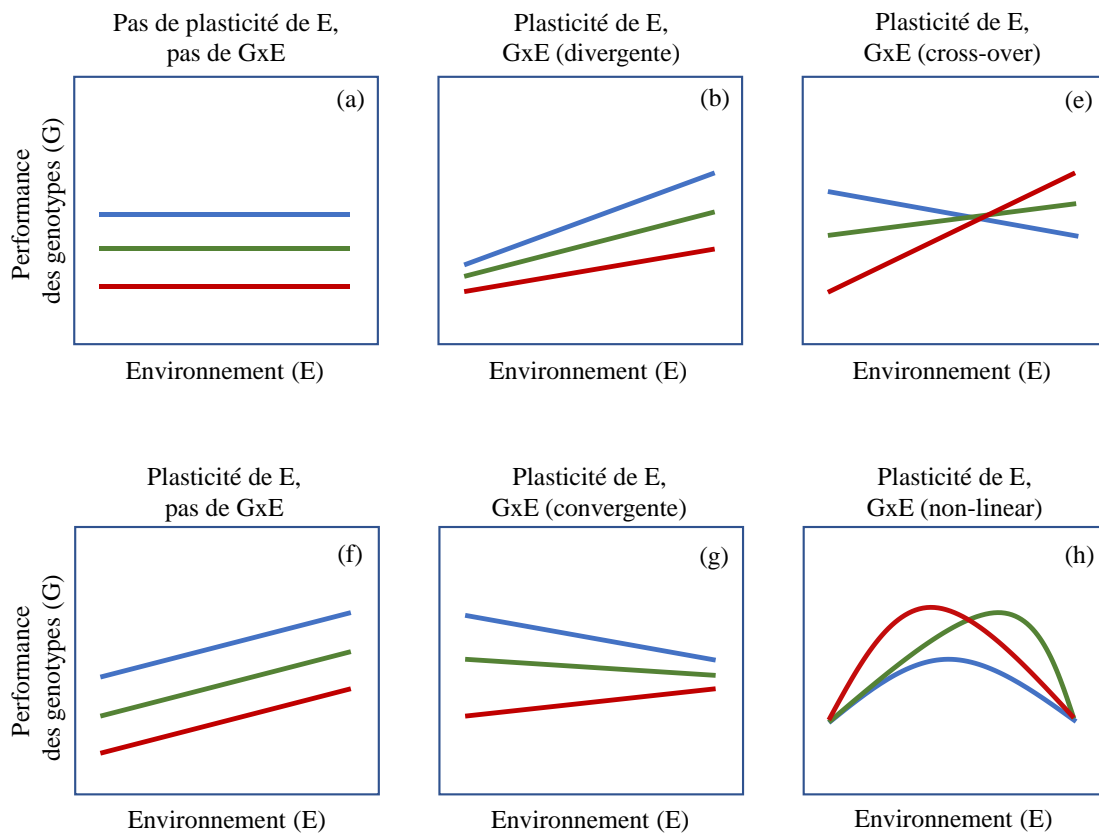


Figure I.17 : Illustration de l'effet génotypique (G), de l'effet de l'environnement (plasticité de E) et de l'interaction Génotype x Environnement (GxE), d'après van Eeuwijk et al., (2016)

complexes comme présentés dans l'évolution du LAI. Par exemple, Christopher et al., (2014) proposent de modéliser la sénescence par une fonction représentant une sigmoïde et d'utiliser certaines propriétés de cette sigmoïde pour estimer les paramètres de *stay-green* tel que le début de sénescence, le point d'inflexion, la vitesse... D'autres études comme (Genard et al., 1999) modélisent la sénescence avec une fonction représentant une sigmoïde, mais ils implémentent le modèle en rajoutant une phase de sénescence lente en début de sénescence et une phase de sénescence rapide dans un second temps. Le choix de la modélisation est souvent dû au nombre de points de cinétique et de la précision de chacun d'entre eux.

4.2 Méthode d'analyse de réseau d'essai pour sélectionner les variétés en lien avec un stress

Pour explorer un maximum de scénarios de stress, les sélectionneurs mettent en place un réseau d'expérimentation plus ou moins large en fonction de leurs moyens. Ce réseau permet de tester les variétés dans un maximum d'environnements différents mais représentatifs des environnements ciblés par le marché. Cependant, nous avons vu en cherchant à identifier les traits physiologiques permettant de sélectionner ou d'adapter de nouvelles variétés à des environnements stressants, que chaque trait pouvait aussi bien avoir un effet positif qu'un effet négatif en fonction du scénario de stress (Tardieu, 2012). Ainsi, pour estimer la performance globale des variétés ou en réponse à un stress, les sélectionneurs sont confrontés aux problèmes de l'interaction génotype-environnement (GxE). Dans cette partie, nous allons présenter les méthodes statistiques les plus utilisées pour prendre en compte cette interaction et extraire le maximum d'informations à partir d'un réseau d'expérimentation.

4.2.1 Le problème des interactions génotype x environnement (GxE)

Déterminer quelle est la meilleure variété, c'est-à-dire la plus stable et/ou la plus performante, dans une large gamme d'environnements peut être une chose difficile, que ce soit du point de vue du sélectionneur ou pour conseiller aux agriculteurs une variété et la conduite culturale la plus adaptée pour maximiser son potentiel. Rapidement, il est possible de séparer les variétés en deux classes : les variétés qui sont performantes dans de nombreux types d'environnements et les variétés qui sont spécifiquement adaptées à des conditions pédoclimatiques particulières (Malosetti et al., 2004). Cette plasticité phénotypique peut être décomposée en trois composantes, un effet de l'environnement (E), un effet du génotype (G) et l'interaction entre les deux (GxE) (Jones and Mather, 1958). La [Figure I.17](#) présente les différents changements de performance phénotypique en fonction de la valeur des différentes composantes E, G et GxE dans différents scénarios. Dans la [Figure I.17a](#), on peut voir que la performance des variétés n'est pas dépendante de l'effet de l'environnement, mais seulement de l'effet génotypique. Dans la [Figure I.17f](#), l'effet de l'environnement change la performance des variétés, mais les différences de performance entre variétés restent les mêmes dans les différents environnements. Les [Figures I.17b](#) et [I.17g](#) présentent en plus l'effet de l'interaction GxE. Ces figures

montrent que chaque génotype ne réagit pas de la même façon en fonction de la qualité de l'environnement. Par exemple dans la [Figure I.17b](#), il est possible de voir que la variance phénotypique est différente en fonction de la valeur des environnements. Il y a présence d'une interaction GxE dite divergente. Parmi ces quatre exemples, choisir la variété la plus performante pour un agriculteur ou un sélectionneur n'est pas très difficile. En effet, peu importe l'effet de l'environnement ou du GxE sur les variétés, il n'y a pas de changement de rang entre les variétés. Il suffit donc de connaître la performance de chacune d'entre-elles dans un seul environnement pour s'assurer de sélectionner la plus performante dans n'importe quel environnement. Malheureusement, ces exemples ne sont pas les plus fréquents chez les plantes. En effet, lorsque les environnements sont assez différents, alors le classement des génotypes peut varier en fonction de la qualité des environnements de façon linéaire ou non, [Figure I.17c](#) et [I.17h](#) (van Eeuwijk et al., 2016). Dans ces deux situations courantes chez les plantes il s'avère compliqué de faire de la sélection variétale ou de la préconisation auprès des agriculteurs de façon globale.

4.2.2 L'importance du GxE en expérimentation variétale

Dans un contexte d'essais variétaux multi-environnementaux (MET, pour « multi-environmental trial ») le rendement en grain est souvent la cible principale. Malheureusement, c'est aussi une variable sensible à l'interaction GxE. En effet, étant donné qu'il est la dernière composante à être mise en place, il accumule toutes les compensations qu'il y a eu pendant le développement de la plante. Nous avons vu que différentes stratégies pour éviter un stress pouvaient être mises en place par la plante. Par exemple, l'ouverture des stomates va augmenter la photosynthèse et diminuer la température des feuilles, ce qui est un caractère favorable en cas de stress thermique mais peut être handicapant en cas de stress hydrique terminal. C'est pourquoi, une plante qui va adopter cette stratégie sera plus performante que les autres en cas de stress thermiques, mais moins performante en cas de stress hydriques. Il y a donc changement de classement entre ces variétés dans ces environnements, c'est-à-dire de l'interaction GxE (Tardieu, 2012). Disséquer et étudier l'interaction GxE est donc d'une grande importance en expérimentation variétale ainsi que pour les sélectionneurs (Finlay and Wilkinson, 1963; Cooper et al., 1995; Lacaze et al., 2009; van Eeuwijk et al., 2010). Cela permet de comprendre l'impact de l'environnement sur la plante et de définir quels types de génotypes (idéotypes) sont les plus adaptés à chaque type d'environnement (ici pour le stress thermique et/ou le stress hydrique terminal).

En expérimentation variétale, la source de variation sur le rendement due à l'effet d'interaction GxE est souvent plus importante que celle de l'effet génotype, mais moins que celle de l'effet de l'environnement ($E > GxE > G$). Chenu, (2015) a reporté, grâce à de nombreuses études sur le blé, que la variation sur le rendement due à l'effet de l'environnement était comprise entre 48 et 93%, contre 4-35% pour le GxE et seulement 1-16% pour l'effet génétique (G). Cependant, cet ordre n'est pas

	Env1	Env2	Env3	Env4	Env5	Env6	Env7	Env8	Env9
Geno1	103.26	97.37	99.08	67.49	63.76	96.27	66.24	79.74	61.08
Geno2	85.30	94.38	78.21	65.20	95.68	75.52	96.86	87.29	92.46
Geno3	101.83	98.74	64.42	88.32	82.10	62.12	96.54	94.97	76.73
Geno4	72.69	76.94	71.01	65.99	98.09	62.85	98.45	85.14	52.98
Geno5	97.78	90.23	62.36	76.27	58.67	96.84	72.35	55.61	51.69
Geno6	109.78	95.58	77.25	76.27	94.59	68.99	66.52	54.16	74.42
Geno7	81.48	66.90	79.32	100.75	105.22	90.55	90.51	60.80	98.29
Geno8	99.89	91.96	104.10	91.71	95.52	88.57	76.05	74.04	54.98
Geno9	99.54	66.85	64.24	75.75	72.70	95.28	94.72	69.75	70.28
Geno10	78.80	104.35	74.92	111.11	97.78	85.41	80.54	92.67	70.59
Geno11	110.49	80.25	100.03	96.10	81.44	99.34	68.46	96.71	64.87
Geno12	79.09	72.48	85.19	99.91	107.93	77.56	97.70	85.93	105.42
Geno13	76.83	84.29	109.37	88.11	70.46	87.25	86.44	107.47	78.26
Geno14	107.68	97.70	100.20	105.35	80.95	83.46	86.99	95.03	80.43
Geno15	116.21	95.97	97.19	79.83	82.42	70.77	110.67	107.55	102.66
Geno16	97.48	100.54	73.78	112.50	105.42	109.93	96.22	109.31	104.92
Geno17	71.06	106.50	114.60	67.99	90.73	101.33	81.47	102.48	104.06

↓

Décomposition Classique du Phénotype

$$P_{ij} = \mu + E_j + G_i + G \times E_{ij} + \varepsilon_{ij}$$

↓

	Env1	Env2	Env3	Env4	Env5	Env6	Env7	Env8	Env9	Effet Geno
Geno1	14.43	12.63	18.09	-14.25	-18.87	15.59	-12.95	-1.33	-13.24	-4.74
Geno2	-0.36	8.72	-7.45	-20.65	10.03	-10.13	11.20	1.63	6.81	-0.67
Geno3	22.97	17.87	-16.45	7.46	1.24	-18.75	-24.33	14.11	-4.13	-5.46
Geno4	-3.42	0.82	-3.11	-11.03	21.98	-13.27	22.33	9.02	-23.32	-10.20
Geno5	24.25	16.70	-11.17	2.74	-14.87	23.31	-1.18	-17.93	-21.84	-12.79
Geno6	30.05	15.85	-2.48	-3.46	14.86	-10.74	-13.21	-25.57	-5.31	-6.59
Geno7	-4.43	-19.07	1.41	14.24	1.99	4.64	4.00	-25.11	12.38	-0.41
Geno8	13.65	5	17.35	7.46	9.67	2.32	-10.20	1.43	-31.27	-0.07
Geno9	19.66	-13	-15.15	1.68	-4.13	15.46	-8.84	-4.33	-9.60	-6.44
Geno10	-8.92	16.65	-11.01	2.30	4.23	-2.31	-7.18	4.19	-17.12	1.39
Geno11	20.75	-9.50	10.29	6.35	1.69	9.60	-21.28	6.97	-24.89	3.42
Geno12	-10.49	-17.10	-4.39	10.33	13.36	-12.02	8.12	-3.65	15.84	3.26
Geno13	-10.78	-3.32	21.76	0.50	-17.15	-0.36	-1.17	19.87	-9.35	1.28
Geno14	14.18	4.80	7.30	12.25	-12.85	-9.44	-5.92	2.13	-12.47	6.58
Geno15	20.29	0.06	1.27	-16.09	-13.50	-25.15	14.75	11.63	6.74	9.60
Geno16	-3.64	-0.58	-27.34	11.38	4.30	8.81	-4.90	8.19	3.81	14.80
Geno17	-22.30	13.14	21.24	-25.36	-2.63	7.38	-11.89	9.12	10.70	7.04
Effet Env	7.24	3.15	-0.60	0.25	1.05	-0.91	-2.40	-0.52	-7.27	86.32

↓

Figure I.18 : Illustration du calcul de l'effet génotypique (G), de l'environnement (E) et de l'interaction Génomotype x Environnement (GxE)

toujours observé, par exemple, dans les études de (Bernardo, (2001); Varshney et al., (2011); Sadras et al., (2012)) l'ordre est inversé, avec $G > G \times E > E$.

Décomposer et comprendre la réponse phénotypique des variétés dans un large réseau d'expérimentations est très complexe. De nombreux outils et méthodes d'analyse ont été développés dans ce sens et sont présentés dans la revue de (van Eeuwijk et al., 2016).

4.2.3 Comment analyser et modéliser le GxE

Les premières modélisations de l'interaction GxE ont été réalisées grâce à un modèle linéaire additif relativement simple suivant la formule $P_{ij} = \mu + E_j + G_i + G \times E_{ij} + \varepsilon_{ij}$ (Jones and Mather, 1958), avec P_{ij} qui est la réponse phénotypique de la variété i dans l'environnement j , μ est la moyenne de l'ensemble des données, E_j est l'effet de l'environnement j , G_i est l'effet du génotype i et enfin une résiduelle (des effets qui ne peuvent pas être expliqués par le modèle). Cette résiduelle peut être décomposée en effet de l'interaction du génotype i dans l'environnement j ($G \times E_{ij}$) et une résiduelle ε_{ij} qui est souvent associée à l'erreur expérimentale. La [Figure I.18](#) présente le tableau à double entrée de l'ANOVA qui peut être utilisée pour estimer les effets et calculer les variances de ce modèle. Dans les cas où les génotypes sont répétés plusieurs fois dans chaque environnement, il est possible de modéliser le GxE et la résiduelle de façon indépendante.

Il existe aussi des approches multidimensionnelles de l'analyse de la réponse des génotypes qui sont plus flexibles et efficaces que les modèles additifs. Le modèle AMMI (pour Additive Main effects and Multiplicative Interaction) développé par Gauch, (1988 et 1992) ou encore le modèle GGE (Yan and Kang, 2002) sont les deux plus utilisés. Ils présentent deux intérêts majeurs : la décomposition multidimensionnelle du GxE et une possibilité de représenter graphiquement cette décomposition pour une meilleure interprétation. Le modèle AMMI décompose la résiduelle $G \times E_{ij} + \varepsilon_{ij}$ en valeurs singulières (SVD) environnementales et génotypiques comme $G \times E_{ij} + \varepsilon_{ij} = \sum_{k=1}^K \lambda_k a_{ik} b_{jk} + \varepsilon'_{ij}$, avec λ_k les valeurs singulières de l'axe k , a_{ik} est la valeur propre du génotype i sur l'axe k , b_{jk} est la valeur propre de l'environnement j sur l'axe k , l'erreur expérimental ε'_{ij} . La décomposition de la résiduelle en valeur propre permet de trouver les meilleurs discriminants génétiques et environnementaux qui expliquent la résiduelle. Le nombre d'axes de K peut être déterminé de différentes manières (Cornelius, 199 ; Bro et al., 200 ; Josse and Husson, 2012).

La modélisation GGE est une autre approche multi-dimensionnelle très répandue dans la littérature, elle est proche du modèle AMMI. L'idée est de rassembler les effets génotypes et GxE pour former un effet GGE. Les modèles GGE sont souvent appréciés des sélectionneurs grâce à leurs représentations graphiques sous forme de bi-plot. Ce type de graphique permet d'identifier rapidement les génotypes ou les environnements qui participent le plus à l'interaction GxE. La distance entre le génotype de

cordonnées (a_{1i} , a_{2i}) et la projection des axes environnements est assimilée à la stabilité du génotype (Yan et Kang, 2002).

4.2.4 Caractérisation de l'environnement pour expliquer le GxE

Distinguer l'effet génotypique des effets d'interaction GxE ne suffit cependant pas aux sélectionneurs pour choisir quel type de variété est adapté ou pas à certains environnements. Dans ce contexte, souvent de façon intuitive, les sélectionneurs ont identifié des environnements cibles dans lesquels leurs variétés pourront être cultivées. Ces groupes d'environnements sont communément appelés TPE (pour Target Population of Environments), cf [Figure I.14](#). Par exemple, le GEVES a séparé la France en deux grands TPE, une zone nord et une zone sud. Pour leur inscription, les variétés peuvent postuler au catalogue nord et/ou sud. Cela n'est sans doute pas suffisant et les sélectionneurs distinguent plus de zones.

De façon générale, il est intéressant de grouper les environnements pour diminuer la part d'interaction GxE (Braun and Rajaram, 1996; Cooper and Fox 1996; Lacaze and Roumet 2004; Chenu et al. (2011); van Eeuwijk et al., 2016). Une bonne façon d'illustrer cette observation est de se référer à la [Figure I.17e](#). En effet, si l'on compartimente les environnements en deux au niveau de l'intersection, alors on obtient deux sous-classes : une sous-classe avec une interaction GxE divergente et une sous-classe avec une interaction GxE convergente. Dans chacune des deux sous-classes, il reste de l'interaction, mais au sein de chacune d'elles, le classement entre les variétés reste globalement le même. Différentes méthodes de groupement d'environnements ont été proposées. L'une des plus utilisées consiste à regrouper les essais en fonction de leur année et localisation géographique, elle permet de séparer les grands ensembles (année sèche/pluvieuse, localisation nord/sud). Cependant, il est compliqué d'expliquer biologiquement l'effet de ce regroupement sur le GxE. En effet, au sein d'un réseau d'essais il y a souvent des déséquilibres entre les années et la localisation. Enfin, un classement des environnements basé sur le rendement des performances de variétés témoins (variétés qui ont été observées dans l'ensemble du réseau d'essais) a largement été utilisé (Brancourt-Hulmel, 1999; Cooper and Fox, 1996; Desclaux, 1996). Faire ce type de classement est très intéressant car le rendement intègre l'ensemble des facteurs pédo-climatiques qui ont impacté la culture tout au long de son cycle. De plus, de nombreuses études segmentent le rendement en plusieurs composantes (nombre de grains – poids des grains).

Dans l'ensemble, ces types d'analyses sont très performantes pour décomposer le phénotype en composantes (E, G, GxE...), mais il peut être difficile de conclure sur les raisons biologiques de la présence d'interaction GxE. Dans ce sens, les approches de régression factorielle présentées par (Denis, 1988; Van Eeuwijk et al., 1996) permettent d'expliquer une part des interactions au regard de facteurs environnementaux comme les précipitations, les températures, la radiation... Dans ces modèles, le GxE peut être expliqué par n'importe quelles variables environnementales, sur lesquelles

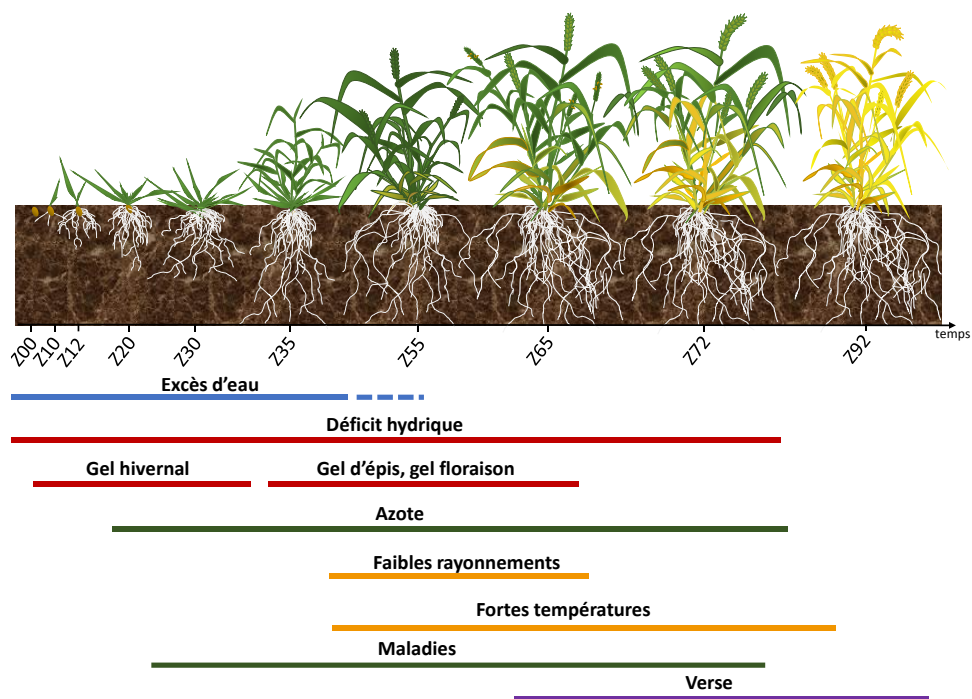


Figure I.19 : Les différents stades de développement du blé et l'initiation des différentes composantes du rendement

il est possible d'émettre des hypothèses quant au comportement physiologique des variétés (Malosetti et al., 2013). Le principe de ces modèles est de décomposer le $G \times E_{ij}$ grâce à une régression linéaire de l'effet d'une variable environnementale en fonction des génotypes suivant la formule $G \times E_{ij} = \beta_i z_j + \delta_{ij}$, avec β_i qui est la pente de la régression spécifique du génotype i (sensibilité de la variété au facteur environnemental), z_j est la valeur du facteur environnemental (ou covariable environnementale) dans l'environnement j , et δ_{ij} la résiduelle de cette régression. Donc une pente importante (valeur importante de β_i) indique une sensibilité importante de la variété (positive ou négative) au facteur environnemental, alors qu'une pente proche de zéro est synonyme d'une stabilité de la variété. Il est donc possible dans ce cadre de conclure sur des hypothèses biologiques et/ou physiologiques de la réponse de la variété à un facteur environnemental (van Eeuwijk et al., 2016). De plus, le terme $G \times E_{ij}$, qui ne permet pas de prédire la performance de la variété dans un nouvel environnement, est remplacé par l'index $\beta_i z_j$ qui peut être modifié et offrir la possibilité d'estimer la performance d'une variété dans un nouvel environnement caractérisé par une valeur de z_j . Par exemple, si le facteur de stress est un déficit hydrique, alors en connaissant la pente du génotype face à ce facteur il est possible de prédire la performance de la variété en condition de déficit plus ou moins fort.

La régression factorielle est donc très intéressante puisqu'elle permet d'introduire une caractérisation environnementale dans le modèle. Cependant, à l'échelle d'un réseau d'expérimentations il y a souvent plus d'un facteur limitant. Il est donc intéressant d'introduire plusieurs covariables environnementales dans les modèles de régression factorielle. Un ensemble de covariables environnementales (EC) peut être introduit dans un modèle de type régression factorielle avec $G \times E_{ij} = \sum_{k \in EC} \beta_{ki} z_{kj} + \delta_{ij}$ (Van Eeuwijk et al., 1996; Brancourt-Hulmel et al., 2000). Le nombre maximal de facteurs environnementaux qui peut être introduit dans ce type de modèle est égal au nombre d'environnements moins un (J-1). Si le nombre de covariables environnementales est supérieur à J-1 alors certaines méthodes de pénalisation ou de réduction de dimensions doivent être utilisées. Par exemple, il existe les méthodes de PLS (Partial Least Square) (Aastveit and Martens, 1986; Vargas et al., 1999, 2015) ou encore des méthodes de régression à rang réduit (Denis, 1988; van Eeuwijk, 1992). Les exemples classiques de covariables environnementales sont les données pédoclimatiques définies à chacun des stades sensibles de la plante (Meynard and Sebillotte, 1994; Lecomte, 2005). La [Figure I.19](#) présente le type de covariables environnementales qui peuvent être calculées tout au long du cycle du blé. Cependant, la sélection d'un sous set de covariables non redondantes qui décrirait parfaitement les interactions GxE n'est pas une chose facile. Plusieurs études ont été réalisées dans ce sens (Meynard et Sebillotte, 1994; Lecomte, 2005; Bouffier, 2014) et peu importe la méthodologie statistique qui est utilisée, il convient d'intégrer une vision physiologique et agronomique pour ne pas faire de non-sens. Dans cette optique, l'utilisation de modèle de culture semble être une bonne façon de calculer et identifier quels types de covariables sont les plus aptes à

bien décrire la culture dans le temps (Chapman, 2008; Chenu et al., 2011, 2013; Heslot et al., 2014; Ly et al., 2018).

4.2.5 L'utilisation des modèles mixtes pour modéliser les interactions GxE

Dans sa revue sur les méthodes d'analyse des interactions GxE van Eeuwijk et al., (2016) ont rendu compte de l'évolution des méthodes utilisées dans 175 articles avec un impact important et publiés depuis les années 70. Il s'avère que l'utilisation de modèles mixtes pour l'analyse du GxE a connu un grand saut ces dernières années. Un modèle mixte est un modèle qui contient à la fois des facteurs fixes et des facteurs aléatoires (Verbeke and Molenberghs, 2011). Les facteurs fixes ont différents niveaux qui sont définis une fois pour toutes. Les effets fixes interviennent dans la moyenne du modèle. Les facteurs aléatoires présentent, quant à eux, un grand nombre de niveaux. Les observations réalisées correspondent à un nombre restreint de ces niveaux, pris aléatoirement. Les niveaux de facteurs observés vont être modélisés de façon aléatoire, ils suivent une loi normale définie par une moyenne nulle (souvent définie par les effets fixes) et une ou plusieurs variances. L'estimation des variances est, la plupart du temps, réalisée grâce à la méthode de maximum de vraisemblance restreint (en anglais REML pour Restricted Maximum Likelihood). Le choix de définir un facteur comme étant fixe ou aléatoire doit être guidé par l'objectif de la modélisation. Par exemple, l'effet d'un traitement fongique sur une culture est souvent considéré comme fixe, car on veut savoir si la présence ou l'absence de traitement sur la culture a un effet sur le rendement. Si l'on veut connaître la part de variance génétique (facteur dont les différents niveaux sont les variétés) dans la variance phénotypique, on préférera définir le facteur comme aléatoire. De plus, si les niveaux de facteur ne sont pas très équilibrés, par exemple, l'observation de chaque génotype dans chaque essai d'un réseau d'expérimentation, alors on préférera définir le facteur comme aléatoire. Enfin, les modèles mixtes sont très utilisés en expérimentation variétale car ils permettent une analyse fine des différents dispositifs expérimentaux. Par exemple, l'analyse de design augmentés, qui ont été largement utilisés dans cette thèse, permet de prendre en compte les effets terrain d'un sous-set de parcelles grâce à l'utilisation de génotypes témoins (ou répétés). Puisque tous les génotypes ne sont pas présents dans tous les sous-blocs, il est plus correct de mettre cette variable en aléatoire (Scott and Milliken, 1993).

Dans un réseau d'essais multi-environnemental (MET), les dispositifs expérimentaux sont souvent différents entre les expériences. L'analyse de ce type de données peut être réalisée en deux étapes (Möhring and Piepho, 2009). La première étape est l'analyse des essais un par un. Dans cette étape, le dispositif expérimental et les effets terrain de chaque essai du MET seront analysés finement pour essayer d'extraire avec le plus de précision possible les effets génétiques ainsi qu'une résiduelle relatant la précision de l'essai (Rodríguez-Álvarez et al., 2016). La seconde étape va être dédiée à l'analyse de l'interaction GxE. Grâce au modèle mixte, les valeurs génétiques et résiduelles de la première étape vont pouvoir être utilisées pour prendre en compte la précision de chaque essai au sein de l'analyse du MET (Möhring and Piepho, 2009; van Eeuwijk et al., 2016).

En plus du fait de modéliser les variances des facteurs aléatoires, les modèles mixtes permettent de modéliser différentes structures de covariances (Boer et al., 2007). Par exemple, dans un réseau d'essais multi-environnementaux, l'hypothèse de variance génétique et de covariances équivalentes dans chaque environnement n'est pas très réaliste (Zheng et al., 2010). Dans le cas d'une analyse du GxE le modèle mixte peut être défini de la façon suivante : $y_{ij} = \mu_j + G_{ij} + \varepsilon_{ij}$, avec μ_j l'effet fixe des environnements de 1 à J, et G_{ij} les effets aléatoires des environnements spécifiques du génotype i dans l'environnement j . G_{ij} suit une distribution normale multivariée avec $\{G_{ij}\} \sim MVN(0, \Sigma)$ (Smith et al., 2005), avec Σ la matrice de variance-covariance définie par $\Sigma = \Sigma^G \otimes \Sigma^E$ où Σ^G est la matrice de variance/covariance entre les génotypes souvent calculée avec des informations de marquage moléculaire et Σ^E est la matrice de variance/covariance entre les génotypes dans différents environnements, \otimes est le produit de Kronecker. Le nombre de paramètres à estimer dans la matrice Σ^E dépend des relations génotypiques entre les environnements. Dans le cas le plus simple, cette matrice peut être résumée à une seule variance pour l'ensemble des environnements, et dans le cas le plus complexe on peut introduire une variance par environnement et une covariance pour chaque couple d'environnements (Boer et al., 2007). Le choix du meilleur modèle peut être fait en comparant des critères d'informations comme l'AIC (Akaike Information Criterion) ou le BIC (Bayesian Information Criterion).

5 Les composantes de variabilité génétique en réponse aux stress et méthode de détection

La variabilité génétique est donc la base de l'adaptation des plantes à de nouveaux environnements. Sans variabilité génétique pour un caractère, il est impossible de faire de la sélection. Comme nous l'avons vu précédemment, plusieurs études ont identifié une variabilité génétique pour la plupart des traits utiles à l'adaptation des variétés en conditions de stress hydriques et/ou thermiques. S'il existe une variabilité génétique pour les traits physiologiques c'est qu'il existe des variations dans le génome des variétés qui leur permettent d'exprimer ces caractères. Le [Tableau I.3](#) présente d'ailleurs l'abondance de gènes connus et/ou d'effets alléliques (ou locus de caractère quantitatifs (QTL)) affectant les traits physiologiques de tolérance aux stress hydriques et thermiques.

Pour les sélectionneurs, la connaissance et l'exploitation de ces variations génétiques sont essentielles pour adapter leur matériel et améliorer leur germoplasme avec de nouveaux traits. Les récentes avancées dans le marquage moléculaire des plantes ont permis de mettre à disposition des sélectionneurs une résolution et un volume d'exploration des génomes des variétés sans précédent. Grâce à ces techniques, plusieurs QTL ont pu être identifiés et utilisés comme une aide à la sélection. Le suivi de ces QTL dans le matériel de sélection est connu sous le nom de sélection assistée par marqueurs. En effet, le principe de base de la sélection dite "directe" est d'identifier un trait d'intérêt dans du matériel pour l'introduire dans du matériel élite par rétrocroisements successifs avec un

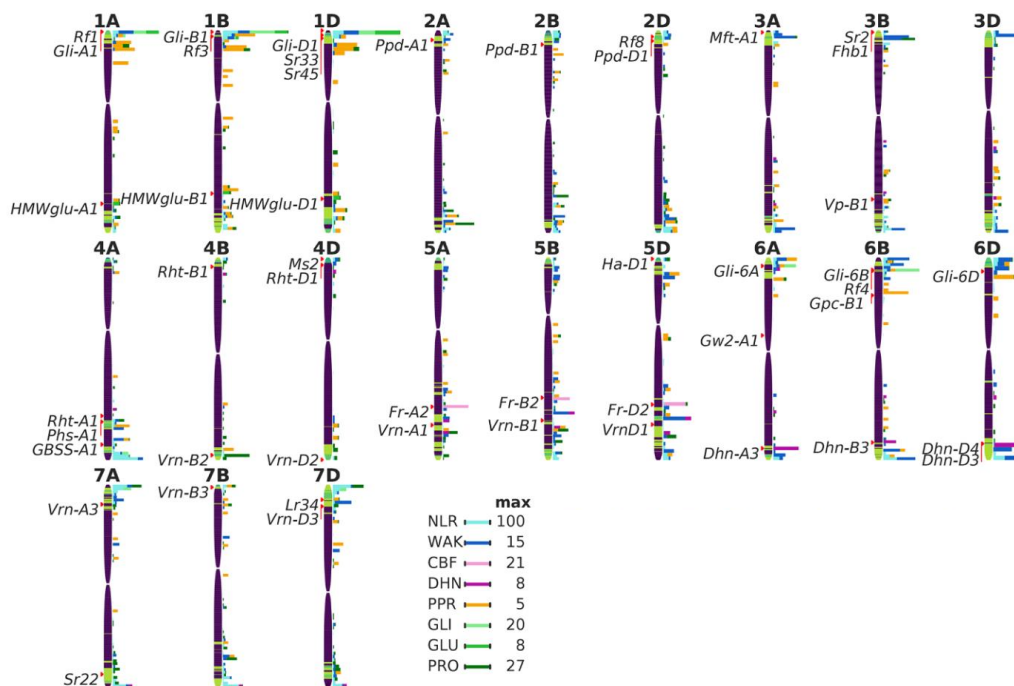


Figure I.20 : Positions des gènes majeurs sur le génome du blé, d'après IWGSC et al (2018)

Définition des gènes et des familles de gènes : Dhn: Dehydrin Fhb: Fusariumheadblight Fr: Frost-tolerance GBSS: Granule-BoundStarchSynthase Gli: Gliadine Gpc-B1: Grainproteincontent Gw2-A1: Grainweight Ha-D1: Hardness HMWglu: Hightmolecularweight-subunitofglutenine Lr34: Leafrustresistance Mft: Motheroffloweringtranscription Ms2: Malesterility Phs-A1: pre-harvestsproutingresistance Ppd-A1: Photoperiod Rf1: Fertilityrestoration Rht-A1: Reducedheight Sr22: geneconditioningresistancetostemrust Vp-B1: Viviparoustranscriptor Vrn-B2: Vernalizationrequirement NLR: nucleotide-bindingleucine-richrepeat WAK: Wallassociatedkinase CBF: C-repeatbindingfactor PPR: PentatricopeptideRepeatProtein Gli: Gliadine GLU: Glutenine

receveur. Dans ce cadre, l'identification de QTL et le suivi de ce QTL au fil des croisements, grâce aux marqueurs moléculaires, va permettre de réduire la part du génome introgressé dans la variété élite qui ne serait pas impliqué dans le caractère d'intérêt. En plus de sélectionner en introduisant de nouveaux QTL, il est tout aussi indispensable d'identifier et de contre-sélectionner des QTL qui seraient pénalisant pour la plante.

5.1 Les gènes majeurs chez le blé

De nombreux gènes majeurs bien décrits dans la littérature impactent la physiologie, la résistance aux maladies, la qualité du grain etc... La position de ces gènes et/ou famille de gènes est présentée dans la [Figure I.20](#). Dans cette partie, nous avons relevé les gènes majeurs impactant la précocité du blé et sa hauteur. Ces gènes sont largement connus pour avoir de multiples impacts sur le rendement.

La capacité des plantes à échapper à un stress, en réduisant la durée de leur cycle, est un des caractères d'adaptation les plus importants chez le blé. Trois classes de gènes sont responsables de la précocité de floraison chez le blé : (1) les gènes de précocité proprement dits *Eps* : la variabilité allélique à ces gènes sont responsables d'un décalage de floraison de quelques jours seulement et contrairement aux autres gènes, ils ne dépendent pas d'effets de l'environnement (Snape et al., 2001) ; (2) les gènes de vernalisation *Vrn*, déterminent la durée minimum d'exposition des variétés au froid pour induire la transition florale, ces gènes vont séparer les céréales en une catégorie printemps et une catégorie hiver ; (3) les gènes de sensibilité à la photo-période *Ppd* déterminent les besoins des variétés en durée de luminosité (jour-long) pour fleurir, ces gènes divisent les céréales en deux catégories : les variétés sensibles à la photopériode (variétés tardives) et les variétés insensibles à la photopériode (variétés précoces). Ces deux dernières familles de gènes sont souvent associées à des variations physiologiques, agronomiques et biochimiques chez le blé. Cependant, il est souvent difficile de séparer l'effet du gène sur le trait qui est étudié de l'effet du gène sur la précocité, donc potentiellement à un effet d'échappement. De nombreuses populations d'expérimentation, populations bi-parentales ou multi-parentales (di-allèle, MAGIC, NAM ...), sont généralement développées en faisant bien attention à ne pas avoir de ségrégation pour ces gènes.

Les gènes de réduction de la taille des pailles (*Rht*) ont été largement associés à ce que l'on a appelé la « révolution verte » et la sélection de variétés dites naines ou semi-naines. D'après Borojevic et Borojevic, (2005) l'origine de la plupart des gènes *Rht* (*Rht1*, *Rht2* et *Rht8*) proviendrait de variétés japonaises, soit en arrivant directement dans des variétés Italiennes (*Rht8*), soit en passant par des variétés de blé Américain (*Rht1* et *Rht2*). Comme les gènes de précocité, les gènes *Rht* sont associés à de nombreux effets sur la morphologie et la physiologie des plantes. En effet, en réduisant la taille des pailles, les gènes *Rht* vont augmenter l'indice de récolte des blés, ce qui va accroître le rendement. Si l'indice de récolte est augmenté, il devrait aussi avoir un impact positif sur le rendement en conditions

de stress hydriques. Cependant, comme le font remarquer Reynolds et al., (2006), il n'y a pas d'association claire dans la littérature d'un tel effet.

Ces gènes sont déjà bien connus et suivis par les sélectionneurs grâce à l'utilisation de marqueurs diagnostique. Pour cela, de nombreuses études et méthodes, dont la détection de QTL, ont dû être réalisées pour identifier, dans un premier temps, les régions génomiques de ces gènes puis les gènes eux-mêmes.

5.2 La détection de nouveaux gènes et/ou allèles avec la GWAS

La génétique d'association (« association mapping », « linkage disequilibrium mapping » ou genome-wide association study (GWAS) en anglais) a été développée dans un premier temps pour étudier les maladies génétiques chez l'homme. Un grand nombre d'études a été mené dans les années 1985-2000. Cette technique permet d'évaluer la significativité d'une corrélation statistique entre les allèles à un locus donné et des valeurs de phénotype. Cependant, il a fallu attendre les années 2000 pour voir se développer cette technique chez les plantes. En effet, la présence de structuration complexe dans un panel de plantes (Hey and Machado, 2003) a tendance à fausser l'analyse par de fausses associations de marqueurs avec le phénotype. Il a donc fallu attendre le développement de techniques de marquage moléculaire recouvrant le génome entier et le développement de méthodes statistiques permettant d'estimer la structuration d'une population pour voir le nombre de ces études augmenter. L'un des intérêts de cette approche est d'utiliser des panels de variétés déjà existants, comme les collections de lignées ou le matériel des sélectionneurs. Ceci permet d'économiser le temps nécessaire à la création de populations F2 ainsi que la fixation du matériel par back-cross, haploïdes doublés ou encore la création de lignées recombinantes, lignées quasi-isogéniques. De plus, avec un panel, il est possible d'étudier plusieurs caractères alors que pour une population bi-parentale, les parents sont généralement choisis pour avoir des valeurs extrêmes pour le caractère étudié. Un panel permet donc une diversité de gènes plus large et plus représentative peut être exploitée. Enfin, du fait d'un nombre plus grand de recombinaisons, les panels offrent une meilleure résolution que les populations bi-parentales.

Dans la plupart des études de GWAS réalisées aujourd'hui, les marqueurs moléculaires les plus utilisés sont les SNP (Single Nucleotide Polymorphisme), grâce notamment, au développement de puces qui permettent de caractériser un grand nombre de SNP sur un grand nombre d'individus. La détection des SNP est d'ailleurs souvent représentée en Manhattan plots ($-\log(P\text{-value})$) en fonction des coordonnées des SNP) où l'on ne se focalise que sur un point significatif (quantitatif trait nucléotide, QTN). Mais avant d'arriver à cette étape, les analyses de génétique d'association vont être conduites suivant quatre étapes principales : (1) La description et l'analyse des données phénotypiques et génotypiques utilisées. (2) Le calcul du déséquilibre de liaison. (3) L'analyse et l'impact de la structuration des panels sur le phénotype avec des modèles statistiques appropriés. (4) L'analyse des

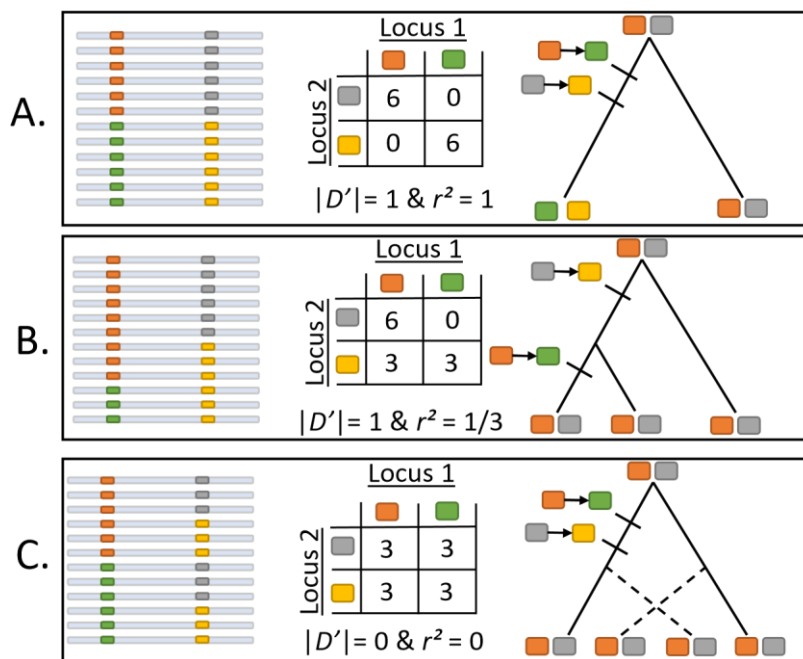


Figure I.21 : Différents scénarios de LD, d'après Flint-Garcia, (2003)

Cette figure a été adaptée de Flint-Garcia, (2003). A) Le LD est complet : mutation de deux loci liés physiquement et il n'y a pas eu de recombinaison. B) Les loci sont en équilibre de liaison : les recombinaisons ont supprimé le LD causé par des mutations qui se sont produites à des moments différents et il n'y a pas eu de recombinaison entre les loci ($D'=1$ et $r^2 < 1$). C) Les loci sont en équilibre de liaison : les recombinaisons ont supprimé le LD causé par des mutations qui se sont produites à des moments différents.

marqueurs identifiés comme associés aux caractères phénotypiques donnés. Le déséquilibre de liaison (LD).

Le déséquilibre de liaison sera noté 'LD' pour 'Linkage Disequilibrium' en anglais. Le LD décrit une association statistique entre deux allèles à deux locus dans une population. Dans une population panmictique idéale, c'est-à-dire une population de grande taille où il n'y a ni mutation, ni migration, ni sélection, les allèles sont indépendamment distribués, donc en équilibre. Le LD fait part de l'écart par rapport à cet équilibre. Dans une population non panmictique, le LD va avoir tendance à augmenter. En effet, le LD existe lorsque la probabilité d'observer un couple d'allèles n'est pas égale au produit des probabilités d'observer ces allèles individuellement. La [Figure I.21](#) présente, pour exemple, deux paires de loci avec les allèles A et a au premier locus et les allèles B et b au second, avec les fréquences alléliques respectives ω_A , ω_a , ω_B et ω_b . Les fréquences haplotypiques qui en résultent sont ω_{AB} , ω_{Ab} , ω_{aB} et ω_{ab} . Le LD va mesurer la différence entre les fréquences haplotypiques attendues et celles que l'on observe (Flint-Garcia et al., 2003), avec comme base le calcul de D comme :

$$D_{ab} = (\omega_{AB} - \omega_A \omega_B)$$

Pour cette mesure, les deux méthodes statistiques les plus utilisées sont le r^2 et le D' . Le r^2 est le carré du coefficient de corrélation des doses alléliques entre deux locus, il est défini par :

$$r^2 = \frac{(D_{ab})^2}{\omega_A \omega_a \omega_B \omega_b}$$

Le r^2 varie de 0 (en équilibre de liaison) à 1 (déséquilibre maximal).

Le D' est une échelle basée sur la fréquence allélique observée, il varie entre 0 et 1 dans le cas où les fréquences alléliques diffèrent entre locus, il est défini par :

$$|D'| = \begin{cases} \frac{(D_{ab})^2}{\min(\omega_A \omega_b, \omega_a \omega_B)}, & \text{pour } D_{ab} < 0 \\ \frac{(D_{ab})^2}{\min(\omega_A \omega_B, \omega_a \omega_b)}, & \text{pour } D_{ab} > 0 \end{cases}$$

Les estimateurs r^2 et D' reflètent différents aspects du LD. Flint-Garcia et al., (2003) présentent trois scénarios de LD. La [Figure I.21a](#) montre un cas où il y a LD total, c'est-à-dire que les deux polymorphismes sont complètement corrélés. Ce cas se présente lorsque deux événements de mutation se sont produits au même moment et où aucune recombinaison n'a eu lieu entre les deux mutations. L'histoire génétique entre les recombinaisons et les mutations est la même pour ce site, c'est pourquoi r^2 et D' seront tous deux égaux à 1. La [Figure I.21b](#) présente un exemple où la corrélation entre les allèles n'est pas complète, mais il n'y a pas eu d'évènement de recombinaison, dans ce cas $D' = 1$ et $r^2 < 1$. Le LD vient des mutations qui se sont produites à des moments différents et où il n'y a pas eu de

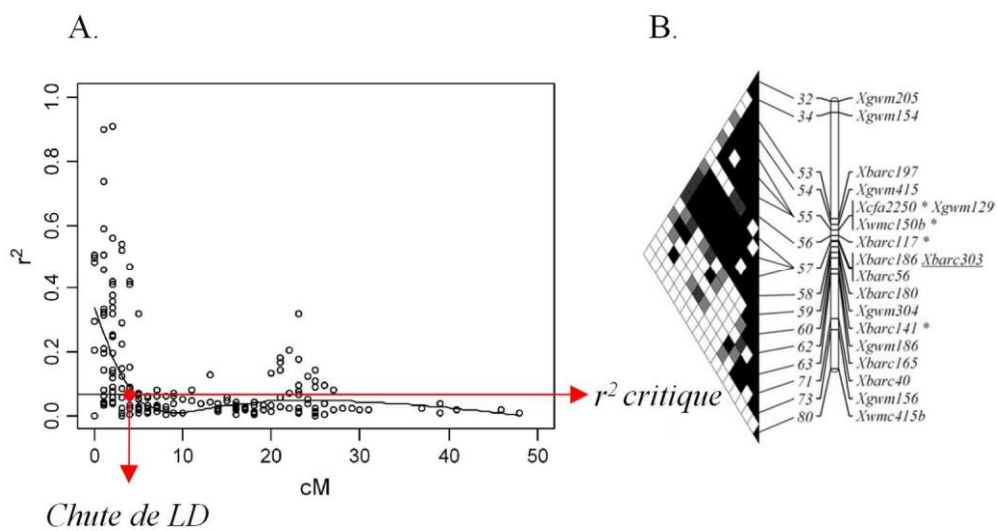


Figure I.22 : Représentation possible du LD le long du chromosomes 5A, d'après Breseghello, (2005)

A. Relation entre la valeur de LD (r^2) et la distance génétique entre chaque couple de marqueurs moléculaires. Le r^2 critique est la valeur de r^2 au-dessus duquel les marqueurs sont significativement en déséquilibre de liaison. B. Le triangle présente les valeurs de r^2 pour chaque couple de marqueur en fonction de la distance génétique représenté sur la droite du triangle, plus les carrés sont sombres et plus la valeur de LD est haute.

recombinaison. Enfin, la dernière [Figure I.21c](#) montre un exemple où il y a équilibre de liaison. Dans le cas où les différents sites seraient proches génétiquement (cM), il est possible que l'équilibre de liaison vienne d'un événement de recombinaison entre les deux sites. Donc l'histoire des recombinaisons diffère pour les haplotypes, mais les mutations se sont produites au même moment, c'est pourquoi $D' = 0$ et $r^2 = 0$. Ainsi, le calcul de r^2 va nous permettre d'estimer des recombinaisons et des mutations, alors que D' ne prendra en compte que les recombinaisons.

Le LD peut être visualisé graphiquement de deux façons différentes ([Figure I.22](#)). La première présente un nuage de point des valeurs de r^2 ou D' en fonction de la distance génétique ou physique entre deux locus. La seconde présente les valeurs du déséquilibre de liaison entre chaque couple de marqueurs le long d'un chromosome (Brescaglio, 2005). Cependant, il faut faire attention si les populations sont de petite taille et/ou si les fréquences alléliques sont faibles, car il y a une estimation moins précise du r^2 et D' . Dans le cadre d'une étude d'association, l'estimateur r^2 est le plus adapté pour caractériser un déséquilibre de liaison. En effet, cet estimateur permet d'avoir une idée de la résolution de détection de polymorphismes causaux, de la proximité entre le locus détecté et le phénotype étudié et du nombre de marqueurs à utiliser pour avoir une couverture complète du génome.

Enfin, en GWAS, la chute de LD, telle qu'elle est représentée dans la [Figure I.22](#), est souvent utilisée pour définir un intervalle autour du QTN dans lequel les marqueurs sont en déséquilibre de liaison et qui peuvent être utilisés pour définir la région génomique qui est liée avec la mutation causale (Cormier et al., 2014).

5.2.1 La structuration des panels

L'un des problèmes les plus communs dans les analyses d'associations est la détection de faux positifs, c'est-à-dire que l'hypothèse H_0 (pas d'association entre marqueur et caractère) est rejetée alors qu'il n'y a pas d'association entre le marqueur moléculaire et le trait (erreur de Type I). En GWAS, le nombre d'erreur de Type I est impacté par deux facteurs : le nombre de tests (ceci sera discuté dans la partie suivante) et la présence d'une structuration dans le panel testé (Flint-Garcia et al., 2003). Cette structuration est commune à chaque organisme, elle vient naturellement du fait des origines géographiques, de la sélection naturelle ou artificielle, etc... Cette structuration doit être prise en compte dans les modèles statistiques pour que chaque association détectée entre marqueur moléculaire et trait soit le fait d'une vraie liaison entre le marqueur moléculaire et la mutation causale. Dans cette partie, nous allons présenter différentes approches statistiques qui ont été et sont utilisées pour minimiser le risque de faux positifs tout en conservant la puissance du test d'association.

5.2.1.1 Approche par groupe génétique (matrice Q)

Pritchard et al., (2000) ont publié deux articles, où ils montrent comment les résultats obtenus au préalable à l'aide du logiciel STRUCTURE, issu de leur premier article, peuvent être intégrés dans les tests d'association effectués sur des échantillons structurés. Cette approche, qui a été rapidement

utilisée pour l'étude de populations végétales, permet d'estimer *a priori* l'appartenance des individus aux différents groupes génétiques présents dans une population à partir de marqueurs génétiques neutres et indépendants. Le logiciel STRUCTURE repose sur une approche bayésienne avec laquelle on cherche à classer N individus en K populations. En sortie, le logiciel retourne la probabilité que l'individu i appartienne à la population k . Le nombre K doit être fixé par l'expérimentateur, cependant, Evanno et al., (2005) ont proposé un test complémentaire pour déterminer le nombre de groupes optimaux. Ce test est basé sur les changements de ratio du log de la probabilité obtenue grâce au logiciel STRUCTURE avec des valeurs de K successives. L'intérêt de cette approche est qu'aucune hypothèse n'est faite sur le nombre et la composition des groupes. Ainsi, le choix du niveau de structuration à prendre en compte comme covariable dans l'analyse sera laissée à l'utilisateur. Pritchard et al., (2000) ont publié deux articles, dans lesquels ils montrent comment les résultats obtenus au préalable à l'aide du logiciel STRUCTURE, issu de leur premier article, peuvent être intégrés dans les tests d'association effectués sur des échantillons structurés.

5.2.1.2 Approche par l'Analyse des Composantes Principales (ACP) (matrice P)

L'utilisation d'ACP a été proposée afin de pallier à la lourdeur des calculs nécessaires au logiciel STRUCTURE. Ainsi, une approche par ACP est proposée pour estimer la structuration des groupes. L'analyse peut être réalisée à l'aide de marqueurs neutres et indépendants. Les axes représentant le maximum de variabilité des variétés sont utilisés pour faire une première ANOVA. Les valeurs résiduelles de cette ANOVA sont ensuite utilisées pour rechercher les associations significatives. Cependant, comme pour les matrices Q (matrices de structure), le choix du nombre d'axes pris en compte influencera la puissance du test. En effet, plus il y a d'axes et plus le risque est grand de diminuer la puissance du test (Mezmouk et al., 2011). La sélection du nombre d'axes peut être réalisée grâce à un test présenté par Patterson et al., (2006).

5.2.1.3 Approche par l'estimation des relations de parenté (matrice K)

Toujours à partir de marqueurs neutres, il a été montré qu'il est possible d'estimer des matrices d'apparentement. La matrice K repose sur l'estimation des liens de parenté entre individus. L'apparentement peut être défini comme la probabilité qu'ont deux individus de posséder, a un locus pris au hasard, une copie d'un allèle issu d'un ancêtre commun. Des valeurs négatives signifient que la paire d'individus testés est moins proche qu'une paire d'individus prise au hasard dans le panel. De plus, une alternative à la matrice K est la matrice K^* , qui est définie comme la proportion d'haplotypes partagés (en enlevant les singletons) par chaque paire d'individus (K. Zhao et al., 2007). Yu et al., (2006) proposent d'intégrer cette matrice K dans un modèle mixte, en tant que matrice de variance/covariance entre génotypes, pour détecter les effets des marqueurs moléculaires.

Pour éviter une contamination proximale de l'apparentement, c'est-à-dire une diminution de la puissance du test en prenant en compte l'effet du marqueur et/ou de la région à travers la matrice K,

Listgarten et al., (2013) ont publié le logiciel FaST-LMM-Select qui va calculer une matrice K en excluant l'ensemble des marqueurs qui se trouve dans la même région que le marqueur testé. Pour simplifier les calculs, Rincent et al., (2014) proposent de tester le marqueur en prenant une matrice K calculée avec tous les chromosomes sauf celui sur lequel le marqueurs testé est positionné.

5.2.1.4 Choix du meilleur modèle à prendre en compte

Toutes les approches présentées précédemment ont un intérêt, cependant elles sont plus ou moins appropriées suivant l'espèce, le niveau de structuration de la population et les liens de parenté entre les individus, s'ils sont connus. (K. Zhao et al., 2007) ont comparé différents modèles (Q, K, P, Q+P et P+K), et ont montré que les modèle P+K permettait des performances similaires au modèle Q+K proposé par Yu et al., (2006). L'avantage étant que la matrice P est plus facile à obtenir. D'après Raman et al., (2010) le modèle Q est plus performant que les modèles K et K+Q chez le blé. Aujourd'hui, la grande majorité des études de GWAS sur les plantes corrige la structuration des panels de diversité en utilisant au minimum l'apparement grâce à une matrice K. En effet, si une structuration est présente dans le panel, elle est prise en compte par la matrice de K. Cependant, certains panels très structurés nécessitent toujours de prendre en compte en plus une matrice Q ou P dans les modèles, par exemple, dans les panels qui mélangent plusieurs types d'orge 6 rangs et 2 rangs.

5.2.2 Choix du niveau de significativité

La seconde raison de l'augmentation d'erreur de Type I vient du fait que les tests d'association sont réalisés sur un grand nombre de marqueurs, il faut donc définir un seuil de significativité pour limiter le nombre de faux-positifs. Si un niveau de significativité de 5% est utilisé dans un test d'association utilisant 100,000 marqueurs, alors on s'attend à obtenir en moyenne 5,000 faux positifs. Pour ajuster le niveau de significativité, afin de le rendre plus discriminant, plusieurs méthodes sont employées. La méthode de Bonferroni (Dunn, 1964) permet d'obtenir un seuil de significativité ajusté au nombre de tests réalisés, avec :

$$\text{Test Bonferroni} = \alpha / N_{\text{tests}}$$

où α est le niveau de significativité attendu pour un seul test et N_{tests} , le nombre de tests total réalisés. Ce seuil est très strict et ne prend pas en compte la non-indépendance des tests. Pour éviter cela, Gao et al., (2009) proposent par exemple de ne prendre en compte que le nombre de tests indépendants pour estimer le seuil.

Une autre alternative est d'estimer la proportion de faux positifs parmi les associations significatives. Cette méthode est appelée FDR (False Discovery Rate) (Benjamini et Hochberg, 1995), avec :

$$FDR = m\alpha / n$$

où α est le niveau de significativité attendu pour un seul test, n est le nombre de QTL détectés avec un niveau de significativité supérieur au seuil initial et m est le nombre de marqueurs testés.

5.3 Dissection des composantes génétiques de l'interaction génétique par environnement

La dissection des déterminants génétiques nécessite dans tous les cas un phénotypage de qualité et ciblé sur les traits d'intérêt, qu'ils soient phénologiques, morphologiques, physiologiques et/ou agronomiques. Dans ce contexte, il existe deux stratégies de phénotypage qui sont complémentaires : le phénotypage en conditions contrôlées et le phénotypage au champ dans un réseau multi-environnemental.

Le phénotypage en conditions contrôlées va concentrer les efforts vers un type de traits dans un type de conditions. Il est souvent réalisé pour appréhender des traits qui sont difficilement mesurables au champ, par exemple, la croissance des racines (utilisation de rhizotron, Postic, (2016)), la tolérance à un stress sans aucune interaction avec d'autres stress (ce qui est souvent le cas dans les études au champ) ou encore la tolérance au stress thermique.

Le phénotypage au champ, quant à lui, est plus proche des activités de sélection. De la même façon, il est possible de se focaliser sur un trait particulier ou un stress. Cependant, dans ce type d'expérimentation de nombreux paramètres entrent en jeu dans la mise en place du trait ciblé, notamment du fait des interactions avec l'environnement. Il en résulte qu'un QTL peut avoir un effet positif sur un trait dans certaines conditions et pas d'effet, voire même un effet inversé, dans d'autres conditions.

La sensibilité de l'expression des QTL en fonction des environnements est définie comme de l'interaction QTLx E . La prospection des interactions QTLx E a été étudiée dans le cadre de population biparentales chez de nombreuses espèces et chez le blé (Malosetti et al., 2004; Boer et al., 2007; Tardieu, 2013; Bouffier, 2014). De la même façon que pour l'étude du Gx E , les interactions QTLx E mises en relation avec des variations climatiques (covariables environnementales) permettent une interprétation biologique de l'effet d'un QTL. Si beaucoup d'études sont basées sur des populations biparentales et peu sont réalisées avec des panels de diversité, c'est en partie du fait de la prise en compte de l'effet polygénique (fond génétique modélisé par la matrice K) en fonction des environnements. En effet, la prise en compte de cet effet dans le cas d'un travail sur panel de diversité augmente largement les temps de calculs. Cependant, Millet et al., (2016) ont proposé une approche en quatre temps pour estimer l'effet allélique de QTL dans un réseau multi-environnemental et corrélérer ces QTL à une variation de température et stress hydrique. La première étape est de modéliser les matrices de variance/covariance avec un modèle n'intégrant pas les marqueurs. Ensuite, un scan de GWAS est réalisé par environnement et en multi-environnement (en intégrant les matrices de variance/covariance estimées dans la première étape) dans le but d'identifier les SNP qui ont au moins

Table I.4 : Récapitulatif de l'ensemble des publications où des QTL sont associés à des phénotypes dans des environnements optimaux, en stress hydrique et/ou thermique

	Références	ID	Espèces	Croisement ou panel	Structure Génétique
Bouffier, (2014)	Araki et al., (1999)	1	T. aestivum L.	CS x CS (Kanto107 4A)	RICL
	Beecher et al., (2012)	2	T. aestivum L.	Halberd x Cutter	RIL
	Bennett et al., (2011)	3	T. aestivum L.	Kukri x Rac875	DH
	Bennett et al., (2012a)	4	T. aestivum L.	Kukri x Rac875	DH
	Bennett et al., (2012b)	5	T. aestivum L.	Kukri x Rac875	DH
	Cadalen et al., (1998)	6	T. aestivum L.	Courtot (Rht-B1b; Rht-D1b) x CS	DH
	Ciuca and Petcu (2009)	7	T. aestivum L.	-	-
	Czyczyło-Mysza et al., (2011)	8	T. aestivum L.	CS x SQ1	DH
	Dia et al., (2008)	9	T. turgidum	Jannah Khetifa x Cham1	RIL
	Do et al., (2007)	10	T. aestivum L.	7C x Seri M82	RIL
	Edae et al., (2014)	11	T. aestivum L.	AMP WAMII	-
	Hanocq et al., (2004)	12	T. aestivum L.	Renan x Récital	RIL
	Huang et al., (2004)	13	T. aestivum L.	Flair x XX86	BC2F3
	Kirigwi et al., (2007)	14	T. aestivum L.	Dharwar Dry x Sitta	RIL
	Kuchel et al., (2007)	15	T. aestivum L.	Trident x Molineux	DH
	Kumari et al., (2007)	16	T. aestivum L.	WL711 x PH132	RIL
	Kumar et al., (2010)	17	T. aestivum L.	-	-
	Liu et al., (2013)	18	T. aestivum L.	Hanxuan 10 x Lumai 14	DH
	Mason et al., (2013)	19	T. aestivum L.	-	-
	Mathews et al., (2008)	20	T. aestivum L.	Seri M82 x Babax	RIL
	Molero et al., (2014)	21	T. aestivum L.	AMP WAMII	-
	Morgan (1991)	22	T. aestivum L.	-	F2
	Morgan and Tan (1996)	23	T. aestivum L.	-	-
	Olivares-Villegas et al., (2008)	24	T. aestivum L.	Seri M82 x Babax	RIL
	Peleg et al., (2009)	25	T. turgidum	Langdon x Emmer	RIL
	Pinto et al., (2010)	26	T. aestivum L.	Seri M82 x Babax	-
	Quarrie et al., (1994)	27	T. aestivum L.	Ciano 67 x CS	RICL
	Quarrie et al., (2005)	28	T. aestivum L.	CS x SQ1	DH
	Rebetzke et al., (2001)	29	T. aestivum L.	Cranbrook x Halberd	DH
	Rebetzke et al., (2007)	30	T. aestivum L.	-	DH
	Rebetzke et al., (2008a)	31	T. aestivum L.	CD87 x Katepwa	-
	Rebetzke et al., (2008b)	32	T. aestivum L.	-	-
	Rebetzke et al., (2013)	33	T. aestivum L.	CD87 x Katepwa	DH
	Snape et al., (2001)	34	T. aestivum L.	-	-
	Verma et al., (2004)	35	T. aestivum L.	Beaver x Soisson	DH
	Vijayalakshmi et al., (2010)	36	T. aestivum L.	-	-
	Yang et al., (2007)	37	T. aestivum L.	Hanxuan 10 x Lumai 17	DH
Tricker et al., (2018)	Acuna-Galino et al., (2015)	38	T. aestivum L.	Méta-analysis	RIL
	Graziani et al., (2014)	39	T. aestivum L.	Kofa x Svevo	RIL
	Merchuk-Ovnat et al., (2016)	40	T. aestivum L.	Bar Nir x Zahir	NIL
	Ogbonnaya et al., (2017)	41	T. aestivum L.	GWAS spring wheat CIMMYT	
	Pandey et al., (2014)	42	T. aestivum L.	WH730 x RAL4014	DH
	Shirdelmoghanloo et al., (2016)	43	T. aestivum L.	Drysdale x Waagan	DH
	Tahmasebi et al., (2017)	44	T. aestivum L.	SeriM82 x Babax	RIL
	Talukner et al., (2014)	45	T. aestivum L.	Ventnor x Karl 92	RIL
Depuis 2014	Ahmad et al., (2014)	46	T. aestivum L.	GWAS	
	Christopher et al., (2018)	47	T. aestivum L.	SeriM82 x Hartog	DH
	Gahlaur et al., (2017)	48	T. aestivum L.	Excalibur x Kukri	DH
	Li et al., (2015)	49	T. aestivum L.	Jingdong 8 x Aikang 58	RIL
	Maphosa et al., (2014)	50	T. aestivum L.	Drysdale x Gladius	RIL
	Mwadzingeni et al., (2017)	51	T. aestivum L.	GWAS spring wheat CIMMYT	
	Shahinnia et al., (2016)	52	T. aestivum L.	RAC875 x Kukri	DH
	Sharma et al., (2016)	53	T. aestivum L.	K 7903 x RAJ 4014	RIL
	Shukla et al., (2016)	54	T. aestivum L.	WL711 x C306	RIL
	Tadesse et al., (2018)	55	T. aestivum L.	GWAS spring wheat CIMMYT	
	Xu et al., (2017)	56	T. aestivum L.	Chuan 35050 x Shannong 483	

un effet dans l'un des environnements. Dans la troisième étape, un modèle multi-locus est réalisé avec les SNP sélectionnés à l'étape précédente et une sélection pas à pas est réalisée pour identifier le set de SNP optimal. Enfin, les effets des SNP qui ont été extraits du modèle multi-locus sont confrontés à l'évolution de variables environnementales pour quantifier leur réponse à l'environnement.

5.4 Les QTL de tolérance à la sécheresse

En complément des travaux de Bouffier, (2014) et de Tricker et al., (2018), une synthèse de QTL liés à des traits phénotypiques dans des conditions de stress hydriques et/ou thermiques a été réalisée. Cette synthèse rassemble 56 publications scientifiques ([Table I.4](#)) faisant référence à des détectons de QTL sur des traits phénotypiques qui ont été présentés précédemment. Les traits phénotypiques ont été rassemblés en trois classes : les traits reliés aux organes de types puits ; les traits reliés aux organes de types sources ; les traits qui font la liaison entre puits et sources. Il est à noter que de nombreuses publications de QTL en lien avec la tolérance à un stress hydrique et/ou thermique ont été identifiés grâce à des populations bi-parentales, très peu sur des panels de diversité, et encore moins, voire pas du tout sur du matériel européen. C'est pourquoi notre étude semble être l'une des premières dans ce cadre.

L'[Annexe I.1](#) présente sur quels chromosomes se trouvent les QTL qui ont été identifiés comme étant associés au rendement en grains et ses composantes, mais aussi avec la biomasse, phénologie et les traits physiologiques liés à la résistance au stress hydrique et thermique (température de canopée, taille des racines, stay-green). Peu importe le trait ou l'environnement, des QTL ont été identifiés sur l'ensemble du génome. Un résumé rapide nous montre que les chromosomes 2B, 3B, 7A et 1B présentent le plus de QTL, alors que les chromosomes 2D, 5D, 3D et 6D sont les chromosomes qui ont le moins de QTL. Dans l'ensemble, il est intéressant de noter que le génome D présente deux fois moins de QTL que les génomes A et B. En effet, le génome DD présente moins de polymorphisme que les autres génomes, donc moins de QTL peuvent être identifiés. Cette différence a été largement décrite dans la bibliographie du fait de l'hybridation tardive entre le blé tétraploïde (génome BB AA) et *Aegilops tauschii* (génome DD), cf [Partie 1.1](#). Plusieurs programmes pour enrichir la diversité du génome D ont d'ailleurs été lancés. On peut citer le programme FSOV Ae Tauschii 2006, dans lequel des hybridations entre blé dur et *Ae tauschii* ont été réalisées pour créer des blés dits synthétiques, ce qui a permis d'identifier plusieurs gènes de résistance aux maladies.

Comme attendu, de nombreux QTL ont été identifiés sur les chromosomes portant les gènes majeurs *Rht*, *Ppd* et *Vrn*, c'est-à-dire les chromosomes des groupes 4, 2 et 5, respectivement. La cartographie de ces gènes est présentée dans la [Figure I.20](#). Si Bouffier, (2014) a reporté que le chromosome 4A était le chromosome le plus fréquemment identifié dans la littérature comme associé au rendement en grains et à ses composantes, il ne semble pas que ce soit le cas dans les 20 publications qui ont été rajoutées à cette synthèse. En effet, c'est le chromosome 1B qui ressort le plus, suivi du 2B et du 7A.

Une analyse plus large, type Meta-QTL, pourrait être réalisée, comme par exemple Acuña-Galindo et al., (2015). Cependant, ces types d'analyses sont souvent réalisées sur des populations bi-parentales avec des QTL dont la taille est supérieure à plusieurs cM et qui sont identifiés sur des cartes génétiques différentes, ce qui rend le travail extrêmement complexe et fastidieux. La publication de la séquence de référence du blé, faite par le consortium international, va certainement faciliter ce travail avec la possibilité de s'appuyer sur la même référence pour la position des QTL. De plus, l'annotation des gènes peut permettre une prospection des gènes candidats sous les QTL identifiés, par exemple, dans la publication de l'IWGSC et al., (2018) une famille de 25 gènes liés à la tolérance à la sécheresse a été localisée sur les chromosomes 6A, 6B et 6D.

Chapitre II : Le panel BREEWHEAT

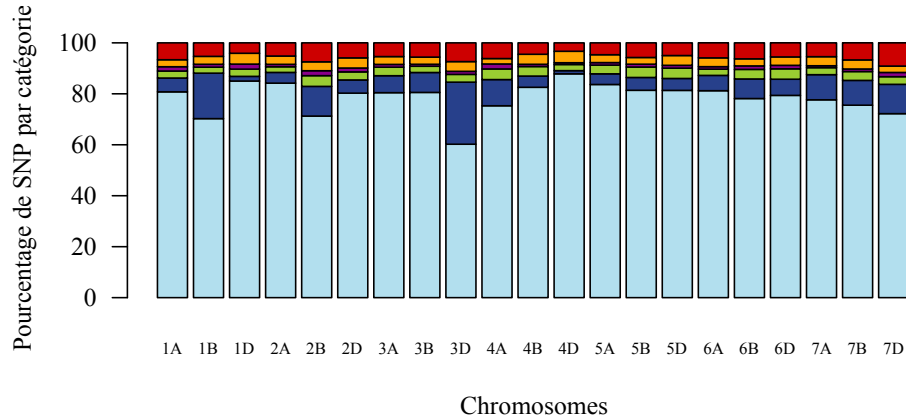


Figure II.1 : Répartition de différents clusters de catégories de qualité des SNP de TaBW280K spécifique au panel élite

Chaque couleur représente le pourcentage du nombre de SNP, par chromosome, appartenant aux différentes classes de qualités identiques à celles présentées par (Rimbert et al., 2018). Les SNP PolyHighResolution sont en bleu clair, les OTV sont en bleu foncé, les MonoHighResolution sont en vert, les NoMinorHom sont en magenta, les CallRateBelowThreshold sont en orange et les autres sont en rouge.

Table II.1 : Répartition et recouvrement des marqueurs sur la carte physique (Appels et al., 2018) et génétique (Rimbert et al., 2018)

	NOMBRE DE SNP	TAILLE EN MBP	TAILLE EN CM	PIC
GENOME A	85937	4932.18	1142.57	0.28
1A	11597	592.84	158.90	0.25
2A	14145	780.51	126.40	0.27
3A	12513	750.63	163.77	0.28
4A	9088	744.08	186.36	0.22
5A	13676	709.75	194.02	0.32
6A	11127	617.92	128.79	0.30
7A	13791	736.44	184.33	0.30
GENOME B	90504	5176.86	993.82	0.29
1B	12548	688.62	133.21	0.27
2B	12126	801.21	140.66	0.28
3B	16981	829.35	166.13	0.28
4B	9957	673.24	125.98	0.35
5B	14123	713.00	178.42	0.28
6B	12802	720.87	109.73	0.27
7B	11967	750.57	139.69	0.29
GENOME D	29895	3948.50	1375.05	0.31
1D	5495	495.17	176.92	0.39
2D	5486	651.34	217.12	0.31
3D	3363	615.21	230.94	0.24
4D	2492	509.14	139.01	0.32
5D	4041	565.73	233.38	0.34
6D	4448	473.47	179.54	0.32
7D	4570	638.44	198.13	0.24
TOTAL	206336	14057.53	3511.44	0.29

Chapitre II : Le panel BREEDWHEAT comme support

Le matériel biologique qui a été utilisé dans ces travaux est le même pour l'ensemble des parties. Un total de 220 variétés élites européennes ont été assemblées en panel de diversité dans le cadre du WP2 du projet BREEDWHEAT, il est souvent appelé simplement panel élite. La plupart de ces variétés ont été inscrites au catalogue français par différentes entreprises d'obtention variétale entre 1974 et 2012. Ce panel comporte donc une diversité génétique représentative du matériel élite français récent. Cependant, certaines variétés ont été sélectionnées en dehors de la France (Italie, Angleterre, Hollande, Allemagne), mais sont agronomiquement adaptées à nos conditions de culture. Au sein de ce panel, l'écart-type de la date d'épiaison est d'un peu plus de six jours. Comme nous le verrons par la suite, une grande partie de cet écart de date d'épiaison est due à la présence de la forme fonctionnelle, ou non, du gène Ppd-D1, dont la fréquence au sein de notre panel est quasiment à l'équilibre. Ceci a un impact significatif sur les interprétations des résultats, notamment ceux issus d'expérimentations au champ ou un stress peut impacter un génotype précoce ou tardif à des stades de développements différents. Il est donc impératif de prendre en compte cet effet d'échappement dans l'ensemble de nos analyses.

L'ensemble du matériel a été génotypé grâce à la puce TaBW280K, qui, comme son nom l'indique, contient plus de 280K marqueurs SNP répartis sur l'ensemble du génome de façon plus ou moins homogène. Cette puce a largement été décrite dans la publication de Rimbart et al., (2018). Comme il a été souligné précédemment, plusieurs points importants sont à étudier avant de réaliser des analyses de génétique d'association dont : (1) L'analyse des données moléculaires ; (2) l'analyse du déséquilibre de liaison ; (3) l'analyse de la structuration.

1 Analyse des données moléculaires

Les SNP utilisés dans ces travaux ont été groupés en six classes de qualités différentes ([Figure II.1](#)). Ces classes sont équivalentes aux classes définies par Rimbart et al., (2018). Il apparaît que la distribution des classes au sein du panel élite est équivalente à la distribution au sein de la population bi-parentale (Renan-Recital), qui a été utilisée dans les travaux de Rimbart et al., (2018). Dans l'ensemble des travaux qui sont présentés dans cette thèse, nous avons décidé d'utiliser seulement les SNP avec une qualité notée « Polymorphic High Resolution ». Cette catégorie regroupe 206,336 SNP dans le jeu de données du panel élite. La distribution des SNP par chromosome et par génome, ainsi que le recouvrement et le polymorphisme sont présentés dans le [Tableau II.1](#). Si les nombres de SNP identifiés sur les génomes A et B semblent être équivalents, le génome D montre, comme attendu, une densité de marqueurs moléculaires bien moindre.

La variation pour chaque locus a été mesurée en utilisant l'indice de polymorphisme PIC pour « Polymorphism Index Content » calculé suivant la formule $PIC = \sum_i^n p_i^2$, avec p_i la fréquence de

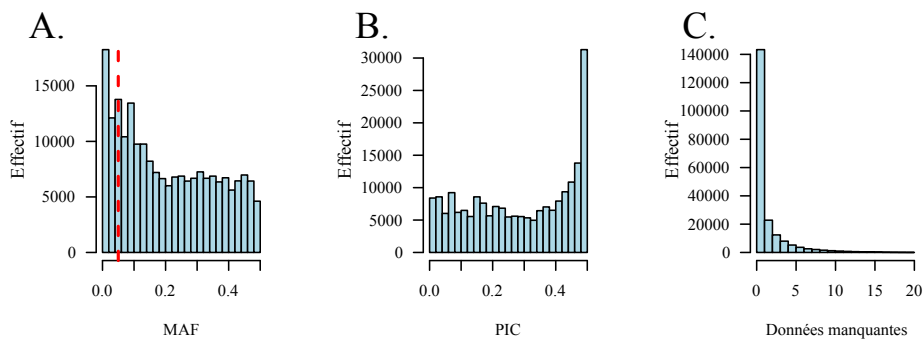


Figure II.2 : Représentation des polymorphismes aux marqueurs et données manquantes pour l'ensemble des SNP 'PolyHighResolution' du panel élite
 A. Présentation de la distribution des MAF (Minor Allele Frequency). B. Présentation de la distribution de l'indice de polymorphisme (PIC). C. Présentation de la distribution du nombre de données manquantes par SNP

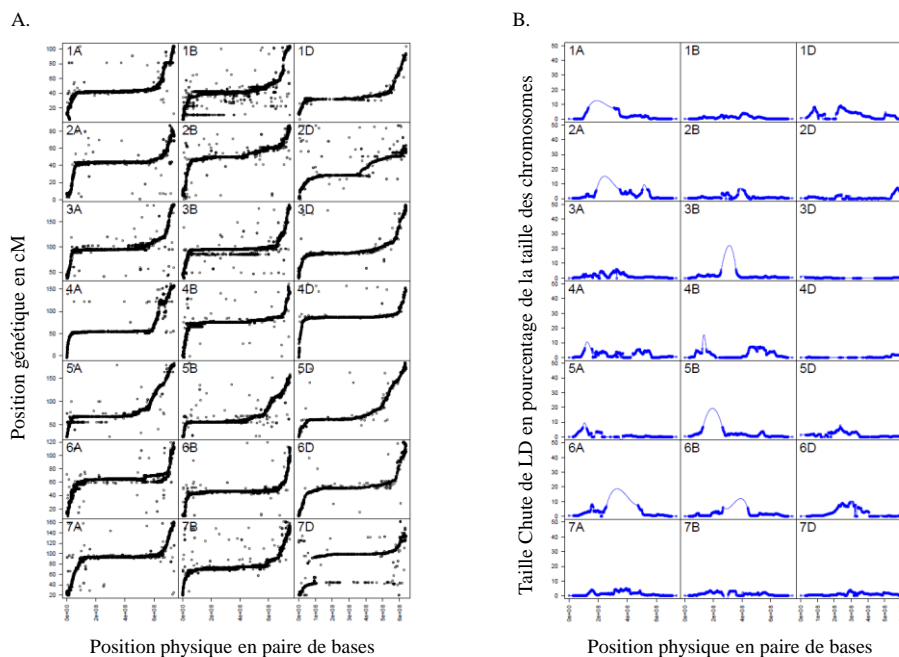


Figure II.3 : Visualisation des recombinaisons le long des chromosomes
 A. Représentation de la dispersion des positions génétiques en fonction des positions physiques (dit cartes de Marey) les positions génétiques sont issues d'une carte consensus entre la cartographie génétique de Renan/Récital (Rimbert et al. 2018) et une carte génétique de Limagrain (data non publiée). B. Représentation de l'évolution de la chute de LD le long du chromosome. La chute du LD a été calculée pour chaque fenêtre de 5 Mb et par pas de 1Mb, la valeur de LD critique a été estimée à 0.23 pour le panel élite.

l'allèle i et n le nombre d'allèles au marqueur. Cette variation a été moyennée pour chaque chromosome et pour chaque génome. Au sein des 220 génotypes du panel élite, le PIC s'élève à 0.29. Sans surprise cet indice est plus faible que dans la population bi-parentale Renan-Recital, mais reste en accord avec les résultats qui ont pu être identifiés sur des panels de blé similaires (Le Couviour et al., 2011). Les distributions représentant le polymorphisme et les qualités des données (présence de données manquantes) sont présentées dans la [Figure II.2](#). La distribution des MAF (*Minor Allele Frequency*) et des PIC est en accord avec ce que nous pouvons attendre d'un panel de diversité et, de plus, le nombre relativement faible de données manquantes par SNP semble être largement acceptable. Cependant, pour faciliter les calculs et la comparaison des associations entre des SNP qui n'auraient pas le même nombre de données, nous avons utilisé le logiciel Beagle 4.1 pour minimiser l'impact des données manquantes (Browning et Browning, 2016). Beagle cherche à identifier des segments de descendance (haplotype) pour imputer les données manquantes.

Enfin, le [Tableau II.1](#) montre le recouvrement et la répartition des SNP sur l'ensemble du génome. De plus, la [Figure II.3](#) montre la relation entre les positions physiques et génétiques. La forme typique de ces courbes représente le taux de recombinaison le long des chromosomes. Il apparaît clairement que le taux de recombinaison dans les régions centromériques est très faible. Ces grands blocs péri-centromériques ont pour conséquence une faible résolution pour la détection de QTL. Il est donc particulièrement difficile d'identifier les gènes causaux lorsqu'ils sont positionnés dans ces régions, par une approche génétique seule, du moins.

Pour conclure, il apparaît que les SNP qui ont été utilisés dans ces travaux permettent d'avoir une couverture uniforme et équivalente, voire supérieure, aux études réalisées jusqu'à maintenant sur le blé tendre.

2 Etude du déséquilibre de liaison pour définir des QTL

La plupart des études utilisent la chute de LD pour étendre les QTN en QTL (Tian et al., 2011; Zhao et al., 2011; Le Gouis et al., 2012). L'étendue du LD est définie par l'intersection entre le seuil de LD critique (c'est-à-dire la valeur de LD à partir de laquelle deux SNP sont considérés comme significativement liés) et la chute des valeurs de LD en fonction de la distance entre les SNP ([Figure II.3](#)) (Brescaglio, 2005). Cette étendue moyenne de LD peut être très différente en fonction du matériel travaillé, par exemple, elle va être élevée dans une population multi-parentale alors qu'elle va être faible dans un panel de diversité. Cependant, le LD est loin d'être constant le long des chromosomes, comme on peut le voir sur les courbes en S de de la [Figure II.3](#) ou encore grâce à l'étude de Choulet et al., (2014) qui montre que le taux de recombinaison méiotique (composante du LD et des positions génétiques) n'est pas constant le long des chromosomes et où l'on peut même observer des points chauds de recombinaisons.

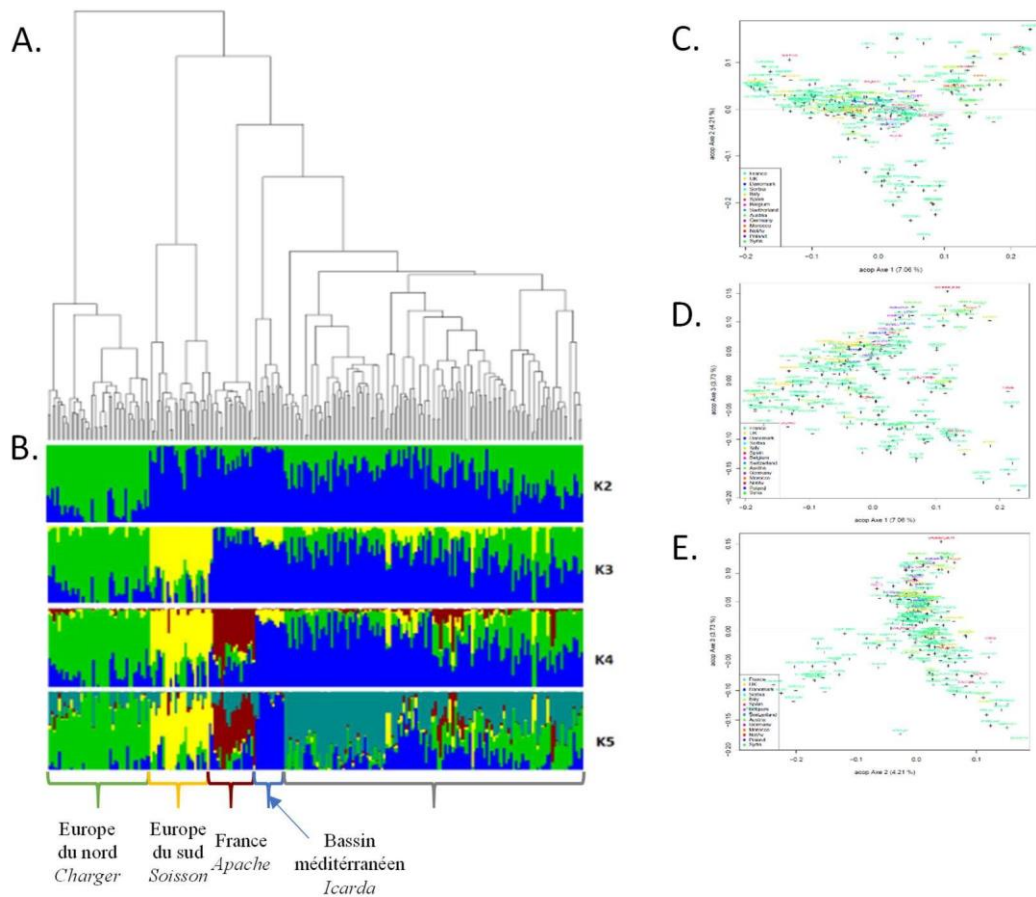


Figure II.4 : Structuration et apparentement dans le panel élite

- A. Dendrogramme réalisé sur 1-Kinship ; B. répartition des génotypes en différents groupes de structure suivant le logiciel STRUCTURE ; C, D, E dispersion du panel suivant 3 axes d'ACP

Dans cette thèse, nous avons défini les bornes des QTL en suivant la méthode présentée par Cormier et al., (2014). Dans cette optique, la [Figure II.3](#) présente l'évolution de la chute du LD le long des chromosomes. Pour cela nous avons calculé et modélisé (grâce à une fonction de régression locale) la chute du LD sur des fenêtres chevauchantes de 5Mb et par pas de 1Mb. Dans le reste de ces travaux, les courbes de chute de LD seront utilisées pour définir les bornes des QTL. Cette méthode présentée par Cormier et al., (2014) permet d'augmenter la chance de détection de QTL encadrant le gène causal. Cependant, dans certaines régions, notamment les régions centromériques, il est impossible de calculer la chute du LD car l'ensemble des SNP sont en LD total. La fonction de régression locale permettra d'estimer l'étendu du LD dans ces zones, mais définira des QTL bien trop grand pour espérer identifier le gène causal.

3 Analyse de la structuration et de l'apparentement du panel élite

Dès qu'une variété est inscrite au catalogue français ou européen par une entreprise de sélection, les entreprises concurrentes ont la possibilité d'utiliser librement ces lignées dans leur propre schéma de sélection ce qui contribue à une faible structuration au sein des lignées élites européennes. Il existe néanmoins une structuration qui peut être expliquée, soit par des effets fondateurs (exemples de lignées commerciales à succès utilisées massivement par une partie des obtenteurs et pas par d'autres), soit par une stratification du matériel en fonction des régions géographiques (Le Couviour et al., 2011; Cormier, 2015). Ces deux dernières études ont été réalisées sur un panel comprenant plus de 80% de variétés en commun avec le panel élite. Nous n'avons donc pas poussé les analyses visant à étudier la structuration de notre panel dans cette thèse, car les conclusions sont sensiblement les mêmes. Cependant, quelques analyses de base ont été réalisées et sont présentées dans la [Figure II.4](#). Ces analyses ont été réalisées *via* un sous-ensemble de plus de 20,000 SNP. Ce sous-ensemble comprend des SNP qui ont été triés sur la base du LD, ceci dans le but qu'il n'y ait pas trop de redondance et qu'il soit assez représentatif de l'ensemble des variations du génome. Ainsi, une approche par groupe génétique a été réalisée grâce au logiciel STRUCTURE (Pritchard et al., 2010). Une matrice d'apparentement (Kinship) a aussi été calculée, et enfin, une ACoP (Analyse en coordonnées principales) a été réalisée grâce à la Kinship. La méthode développée par Evanno et al., (2005), qui est utilisée pour estimer le nombre optimal de groupes suite aux résultats issus du logiciel STRUCTURE, nous a indiqué que quatre groupes semblent structurer le panel élite. C'est aussi le nombre d'axes d'ACP qui ont été identifiés suivant la méthode de Patterson et al., (2006). Enfin, l'observation du dendrogramme qui a été réalisé sur la Kinship montre qu'il est possible de structurer logiquement le panel en cinq groupes. L'ensemble de ces analyses nous laisse penser qu'il existe bien une structure au sein du panel élite. En s'appuyant sur l'avis d'experts (sélectionneurs) ayant une connaissance poussée des pédigrées des variétés du panel, il est possible d'attribuer aux différents groupes des origines géographiques différentes. Le panel serait donc structuré suivant les groupes Europe du nord, Europe du sud, France, Bassin méditerranéen et enfin le plus grand des groupes avec une importante

admixture (mélange de plusieurs groupes). Pour la suite des travaux, nous avons décidé d'écartier certains individus, notamment les variétés ICARDA, au regard de cette structuration et de leur phénotype (variétés précoces et peu performantes comparées aux autres variétés du panel).

Les analyses réalisées nous montrent que le panel élite présente une structuration. Cependant, nous avons décidé de ne pas prendre en compte directement cette structuration dans les modèles de GWAS. En effet, en comparant les résultats issus du logiciel STRUCTURE et le dendrogramme ([Figures II.4A et II.4B](#)) il apparaît que la Kinship résume de façon correcte la structuration du panel. Cette observation est assez commune, il apparaît en effet que la structuration des panels est la plupart du temps prise en compte par la Kinship, sauf dans des cas extrêmes. On peut mentionner l'exemple de l'orge où l'on distingue bien la séparation en orge à deux rangs et à six rangs. De plus, dans le chapitre IV, plusieurs modèles d'association ont été comparés grâce à des QQ-plots (modèle « naïf », modèle K, modèle K+ACP) et leur analyse montre qu'il n'y a effectivement pas de différence significative entre les modèle K et modèle K+ACP.

La caractérisation du panel élite face aux stress hydriques et thermiques

L'objectif principal de cette thèse était de caractériser le panel élite Européen pour sa résistance aux stress hydriques et/ou aux stress thermiques et si possible d'identifier des déterminants génétiques liés à la réponse des variétés à ces caractères. Dans un premier temps, pour remplir ces objectifs, nous avons décidé de nous focaliser sur un type de stress à la fois (chapitre III et IV). Dans un second temps, une prospection multi-stress a été réalisée pour essayer de faire une analyse globale des stress dans un réseau d'essais.

Les analyses et résultats du chapitre III sont présentés sous forme d'un article scientifique. Cet article a été soumis et accepté avec révisions majeures au journal "Theoretical and Applied Genetics". Les travaux de ce chapitre reposent sur une caractérisation du stress hydrique au champ. Cette caractérisation a été réalisée grâce à l'utilisation du modèle de culture CHN et repose principalement sur des données du réseau d'essais multi-environnemental issu des travaux du WP2 du projet PIA BREEDWHEAT et de données propres à Biogemma et à Arvalis. L'objectif principal de ce chapitre est d'explorer et d'identifier les différents types de scénarios de stress qui ont pu impacter les variétés dans notre réseau, afin de pouvoir classer les performances des variétés en fonction des types de stress hydriques. Une fois que les environnements ont été classés, une prospection large des QTL a été réalisée au sein de chaque scénario de stress. De plus, les scénarios de stress observés dans le réseau d'essais ont aussi été confrontés à l'occurrence des stress hydriques en France. Le but de ces analyses était d'identifier quels types de stress de notre réseau d'essais sont représentatifs des conditions de cultures françaises et surtout d'identifier quels types de QTL il pourrait être utile de suivre pour augmenter le progrès génétique en fonction du scénario de stress à l'échelle de la France.

Les analyses et résultats du chapitre IV sont aussi présentés sous forme d'article scientifique. Cet article sera certainement soumis au journal "Theoretical and Applied Genetics". Les travaux de ce chapitre reposent sur une caractérisation du stress thermique en condition contrôlée grâce aux données phénotypiques qui ont été produites dans le cadre du projet FSOV HeatWheat « *Analyse de la diversité génétique de la réponse au stress thermique via phénotypage fin et génétique d'association* ». L'objectif principal de ce chapitre est d'explorer les réponses des variétés à un stress thermique post-floraison. Le fait que le phénotypage ait été réalisé en conditions contrôlées nous a permis de découpler clairement le stress thermique du stress hydrique et de s'affranchir du problème d'échappement dû au fait que le panel présente des différences de précocité, ce qui n'est pas le cas dans les réseaux d'essais au champ.

Enfin, les analyses et résultats du chapitre V, auraient aussi dû être présentés sous forme d'article. Cependant les contraintes de temps ne nous ont pas permis de réaliser les analyses et l'écriture de cette partie sous cette forme. Le but de ce chapitre est d'explorer les réponses alléliques des QTL en fonction de covariables environnementales issues d'un sous-set du réseau d'essais qui a été sélectionné

grâce aux analyses réalisées dans le chapitre III. Cet article explore et discute de méthodes d'identification de QTL, de covariables et de réponses des QTL aux covariables dans un réseau multi-environnemental. Etant donné que les deux premiers chapitres de cette thèse sont focalisés sur des types de stress particuliers et que la limite de cette méthode est clairement de ne pas faire fi d'autres types de stress ou bien de se focaliser sur un scénario de stress très spécifique. Il nous semblait important d'ouvrir le champ des possibles avec une prospection plus large et d'essayer de valider les QTL des deux premiers chapitres dans une démarche sans *a priori*.

Chapitre III : Tolérance aux stress hydriques

Chapitre III: Using environmental clustering to identify specific drought resistant QTLs in bread wheat (*T. aestivum* L.)

Gaëtan Touzy^{1,2} – Renaud Rincint³ – Matthieu Bogard⁴ – Stephane Lafarge² – Pierre Dubreuil² – Agathe Mini³ – Jean-Charles Deswarte⁵ – Katia Beauchêne⁶ – Jacques Le Gouis³ – Sébastien Praud^{2*}

¹ Arvalis-Institut du végétal, Biopole Clermont Limagne, 63360 Saint-Beauzire, France.

² Biogemma, Centre de recherche de Chappes, Route d'Ennezat CS90216, 63720 Chappes, France.

³ INRA, UCA UMR 1095, Génétique, Diversité et Ecophysiologie des Céréales, 24 Avenue des Landais, 63177 Aubière Cedex, France.

⁴ Arvalis-Institut du végétal, 6 Chemin de la côte vieille, 31450 Baziège, France.

⁵ Arvalis-Institut du végétal, Route de Châteaufort, ZA des graviers, 91190 Villiers le Bacle, France.

⁶ Arvalis-Institut du végétal, 45 voie Romaine, Ouzouer Le Marché, 41240 Beauce La Romaine, France.

* : corresponding author : sebastien.praud@biogemma.com

Abstract

Drought is one of the main abiotic stresses limiting winter bread wheat growth and productivity around the world. The acquisition of new high-yielding and stress-tolerant varieties is therefore necessary and requires improved understanding of the physiological and genetic bases of drought resistance.

A panel of 210 elite European varieties was evaluated in 35 field trials. Grain yield and its components were scored in each trial. A crop model was then run with detailed climatic data and soil water status to assess the dynamics of drought stress in each environment. Varieties were registered from 1992 to 2011, allowing us to test timewise genetic progress. Finally, a genome-wide association study (GWAS) was carried out using genotyping data from a 280K SNP chip.

The crop model simulation allowed us to group the environments into four water stress scenarios: an optimal condition with no water stress, a post-anthesis drought stress, a moderate-anthesis drought stress and a high pre-anthesis drought stress. Compared to the optimal water condition, grain yield losses in the stressed conditions were 3.3%, 12.4% and 31.2%, respectively. This environmental clustering improved understanding of the effect of drought on grain yields and explained 20% of the GxE interaction. The greatest genetic progress was obtained in the optimal condition, mostly represented in France. The GWAS identified several QTLs, some of which were specific of the different drought stress patterns.

Our results make breeding for improved drought resistance to specific environmental scenarios easier and will facilitate genetic progress in future environments, *i.e.* drought stress environments.

1 INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most important annual crops in the world. Averaged over the last ten years, it has been cultivated on 219 million hectares with a production of 687 million tonnes per year. With 37 million tonnes per year, France is the leading European producer and ranks fifth worldwide (FAO, 2018). From the late 60s, a steady trend in the growth of wheat yields has been observed in France, mainly due to genetic progress and improvements in agricultural practices (Calderini and Slafer, 1998; Gervois et al., 2008; Brisson et al., 2010). However, progress in farm wheat yields has declined over the last two decades, while genetic improvements have been constant (Oury et al., 2012). This finding can be explained by yearly climatic variability, a factor usually unfavorable to cereal yields (Brisson et al., 2010; Oury et al., 2012). Drought stress is one of these climatic causes. It was estimated that yield may decrease by 0.15 to 0.20 tonnes.ha⁻¹ for each 10mm of water deficit (Gate 1995). Therefore, advances in wheat breeding to produce higher performing wheat varieties is the key to keep pace with increasing food demand, not only in optimal but also in stressed conditions. Deeper understanding of the physiological and genetic bases of drought resistance is crucial for maintaining and improving breeding-program efficiency.

When analyzing field experiment networks, yield variations can be split into three components: environmental (E), genetic variation (G), and their interaction (GxE) (Jones and Mather, 1958). GxE interactions are commonly observed in multi-environment trials (MET) and the variation due to GxE is often higher than G variation itself (Chenu, 2015). In this context, implementing an efficient breeding program is challenging, especially if the ranking of the best lines changes with the environments (e.g. Finlay and Wilkinson, 1963; Cooper et al., 1995; Lacaze et al., 2009; van Eeuwijk et al., 2010). To develop wheat varieties adapted to their market, breeders may decide to target a set of management practices and pedo-climatic conditions in which their varieties will be grown. Obviously, this target population of environments (TPE) can be grouped by geographical region (for example, new French variety candidates for the north or south of France). This facilitates the choice of exploiting either specific or local adaptations or selecting for broader adaptations, and possibly selecting the best lines for a given TPE (Braun and Rajaram, 1996). Although GxE interaction is still large within a given TPE, greater benefits from breeding may be obtained by further sub-dividing the TPE. Different grouping methods have already been proposed: (1) a yield-based characterization based on a set of probe-genotypes (although it integrates all biotic and abiotic stresses, this approach requires a large number of years and locations representative of the TPE) (Brancourt-Hulmel, 1999; Cooper and Fox, 1996; Desclaux, 1996); (2) a pedo-climatic characterization using environmental covariates that provide a broad spatial and temporal characterization, and which can be transposed to other locations not present in the MET (Denis, 1988; Van Eeuwijk et al., 1996; Lacaze and Roumet, 2004; Lecomte, 2005); and (3) a stress-specific characterization, for example, to identify the TPE. Muchow et al., (1996) used a crop model to monitor the water deficit index throughout the crop

season. Selecting one of these methods, or any combination of them, depends on the region, the mode of crop management and the aim of the study. In this study, we chose to classify our MET, as conducted by Muchow et al. (1996), to identify TPE using a crop model, by focusing only on drought stress.

During the wheat crop cycle, a drought stress can be defined by three different parameters to create a drought scenario: timing, intensity and frequency. To describe drought stress in wheat, Chenu et al. (2011) proposed to simulate the water-stress index along the crop cycle of the Australian crop production area using the APSIM crop model (Holzworth et al., 2014b). This index is the modeled ratio between water supply and water demand. Water supply indicates the degree of root water uptake, while water demand is the potential transpiration. Defining the drought stress scenario is the first step in selecting a tolerant trait. Unfortunately, a trait which could improve drought resistance in one scenario could have a negative effect in a different one, generating GxE interactions for grain yield (Tardieu, 2012), and in turn making breeding for drought resistance complicated. The main drought-resistance mechanisms include escape, avoidance and tolerance. Drought escape is a classic adaptive mechanism in plant development that enables the completion of the full life cycle, or at least the most critical part of it, before the drought event (Fischer and Maurer, 1978). In wheat, the genetic control of flowering time, by making use of major genes such as Ppd genes and Vrn (Yan et al., 2003; Bennett et al., 2012; Bentley et al., 2013), could be a useful tool to select lines depending on their climatic growing area. Regarding the other drought resistant types, since Monteith (1977), many studies have described physiological traits related to drought tolerance or avoidance mechanisms, including root traits, leaf traits, osmotic adjustment, water potential, and stomata control. These traits have been well described and used as indicators to evaluate drought resistance in plants and some of them have been used to improve genetic gain in drought prone environments (Bouffier, 2014; Christopher et al., 2016; Pinto and Reynolds, 2015; Rutkoski et al., 2016).

Another way to improve drought resistance in wheat is to identify quantitative trait loci (QTLs) and increase their frequency in a breeding germplasm. Numerous studies have already been conducted to decipher the genetic basis of complex traits driven by numerous QTLs. Most QTLs studies for drought resistance have been performed using a bi-parental spring wheat population. QTLs were reviewed in: Budak et al. (2015) ; Farooq et al. (2014) ; Gupta et al. (2017) ; Sheoran et al. (2015); Tricker et al. (2018). Genome-wide association studies (GWAS) have become popular in recent years, due to their direct applicability to existing material, high mapping resolution and the use of more diverse germplasms which are often more representative of the current elite breeding pool (Bressegello, 2005; Le Couviour et al., 2011; Cormier et al., 2014; Zanke et al., 2015). GWAS was used to identify genomic regions associated with grain yield and/or related traits in drought stress events (Gahlaut, 2012; Edae et al., 2014; Sukumaran et al., 2015; Ain et al., 2015; Mwadzingeni et al., 2017; Sukumaran et al., 2018) and a large number of studies were performed under irrigated vs rainfed trials.

Table III.1: Statistical indicators of experiment quality and the genetic variability of grain yield (t/ha) over the trial network

Statistic indicators were extracted from the first stage of the GxE analysis. ETs: Environment Types with OPT optimal condition, LWD late water deficit, MWD moderated water deficit and HWD high water deficit. Mean: general mean of environment (in t/ha) CVe: environmental coefficient variation. Res.: residual variance (in (t/ha)²). H_g² broad sense generalized heritability. AIC: Akaike information criterion.

Environments	ETs	Nb genotype	Mean	CVe	Variances			AIC
					Genetic	Res.	H _g ²	
Cha12RF	OPT	209	6.5	3.96	0.55	0.07	0.88	1177.64
Gre12IR	OPT	187	7.52	3.09	0.27	0.05	0.89	1722.38
Rec12RF	OPT	210	7.9	4.25	0.64	0.11	0.91	2162.6
Cap12RF	OPT	210	8.91	3.4	0.48	0.09	0.91	2087.02
Mon12RF	OPT	209	7.31	4.94	0.42	0.13	0.81	1648.25
Co12IR	OPT	209	8.47	6.81	0.42	0.33	0.58	1451.29
Co12RF	OPT	210	8.4	4.16	0.48	0.12	0.91	2966.48
Lou12RF	OPT	210	9.63	3.54	1.04	0.12	0.94	2295.63
Mau12RF	OPT	210	9.7	2.88	0.93	0.08	0.95	2187.51
Vra13RF	OPT	208	7.54	3.36	0.34	0.06	0.91	1862.53
Rea13RF	OPT	210	8.57	5.3	0.48	0.21	0.89	4145.26
Mon13RF	OPT	210	8.36	3.72	0.32	0.1	0.85	1984.42
Ver13RF	OPT	208	10.07	2.63	0.47	0.07	0.92	1960.13
And13RF	OPT	210	9.63	3.59	0.52	0.12	0.89	2198.92
All14RF	OPT	210	7.63	5.04	0.62	0.15	0.87	2097.79
Cha14IR	OPT	210	8.64	3.77	0.59	0.11	0.9	2209.28
Cha14RF	OPT	210	7.56	3.55	0.45	0.07	0.91	2120.04
Cle16RF	OPT	163	7.44	6.59	0.55	0.24	0.73	1767.73
Wal16IR	OPT	174	9.67	3.88	0.7	0.14	0.9	2047.56
Wal17IR	OPT	174	8.88	3.84	0.25	0.12	0.81	1856.29
Wal17RF	OPT	174	7.84	4.19	0.22	0.11	0.8	1817.3
Gre12RF	LWD	187	6.75	5.38	0.15	0.13	0.66	1945.45
Gre13RF	LWD	210	9	5.53	0.21	0.25	0.6	2357.1
Cle13RF	LWD	210	9.17	3.55	0.38	0.11	0.86	2150.41
Wal14IR	LWD	175	8.24	8.52	0.35	0.49	0.59	2311.52
Wal16RF	LWD	174	7.08	4.78	0.68	0.11	0.91	2099.07
Gre14IR	MWD	197	7.87	5.03	0.37	0.16	0.81	2149.09
Wal15IR	MWD	175	8.96	5.48	0.76	0.24	0.86	2240.98
Cle16RO	MWD	163	6.48	6.48	0.34	0.18	0.69	1595.83
Cas16IR	MWD	186	6.03	3.77	0.46	0.05	0.94	1727.5
Gre14RF	HWD	197	4.26	8.14	0.05	0.12	0.44	1814.23
Wal14RF	HWD	175	5.23	12.15	0.19	0.4	0.48	2202.48
Wal15RF	HWD	175	6.25	7.14	0.98	0.2	0.9	2276.87
Olm16RO	HWD	157	7.49	8.21	0.35	0.38	0.51	1729.33
Cas16RF	HWD	186	5.95	3.75	0.31	0.05	0.92	1621.39

To our knowledge, no study has explored the differences of water stress scenarios in a multi-environment trial of European winter wheats. Moreover, molecular markers for improving grain yield and grain yield prediction in these scenarios are lacking.

The aim of the research reported here was to characterize and understand the behavior of an elite French germplasm adapted to Western Europe under drought stress scenarios. More precisely, we analyzed a MET composed of 35 environments in which 210 registered winter wheat varieties were tested. Firstly, we defined the main drought scenarios present in the MET by modeling a water stress index along the crop cycle using a crop growth model and according to the method proposed by Chenu et al., (2011). To estimate the representativeness of the MET, we also simulated the water-deficit patterns undergone by wheat crops over the last 21 years in France as a function of various agronomic options. Then, we confirmed our hypothesis regarding the environmental and GxE effect on grain yield and its components. Lastly, we estimated genetic progress from 1992 to 2011 and tried to identify the genetic basis of drought resistance using genome wide association studies (GWAS). We report several first-ever QTLs specific to the different drought scenarios. We expect this work to contribute to future progress in breeding for drought resistance.

2 MATERIALS AND METHODS

2.1 Plant materials and field experiments

Two-hundred and ten elite European winter wheat lines, registered from 1974 to 2010, were assembled in a diversity panel ([Suppl. III.1](#)). Lines were evaluated in a MET of 35 environments as the combination of location, year and treatment. Three types of management were applied for ten of them: under rain-out shelters (RO), rainfed (RF) or irrigated (IR) ([Table III.1](#) & [Suppl. III.2A](#)) (Ly et al., 2018). Twenty-one experiments were conducted in France and four in Walla Walla, Washington state, USA. The Walla Walla station was chosen as a reference site for drought experiments because this location has a climate similar to that of the French Mediterranean region, with mild winters, hot summers, and scarce rainfalls during summer. By controlling the irrigation regime, it was then possible to manage a drought stress scenario. Another reason for testing European lines in Walla Walla was that one of the varieties most cultivated in the Walla Walla area (Norwest 553) was originally selected in France.

Classical agricultural practices were used for all the trials. Weeds, diseases and pests were controlled and Nitrogen, Potassium, Phosphate and Sulfur fertilizers were applied at sufficient levels to fulfil plant demand and make sure that yield potential would not be limited by any factor other than water availability ([Suppl. III.2B](#)).

2.2 Drought stress pattern characterization

Drought stress characterization is central in our study. A crop model was used to estimate the water stress undergone by the plants. Firstly, the MET was characterized to explain the grain yield and yield component variations. Secondly, a frequential analysis on historical French data was performed to describe the representativeness of this MET at the French scale.

WSI (supply/demand ratio for water) was calculated daily with the CHN crop model developed by Arvalis Institut du végétal (<https://www.english.arvalisinstitutduvegetal.fr/>). CHN is based on principles described in Gate (1995), Meynard (1997) and Chatelin et al. (2005), and presented by Le Bris et al., (2016). Briefly, CHN is a mechanistic crop model that estimates the daily flows of carbon (C), water (H) and nitrogen (N) through the soil-plant-atmosphere continuum. The soil compartment is connected to the Arvalis database which gathers over 400 different soil types built with more than 7,000 soil measurements from the worldwide soil data base of ISRIC (<https://www.isric.org/>). The atmosphere compartment is also connected to several stations of the METEO FRANCE network composed of more than 250 sites in France, with daily records over the last 25 years for at least minimum and maximum temperatures (°C), global radiation (J/cm²), Penman evapotranspiration (mm) and rainfall (mm). The plant compartment models leaf growth and the leaf area that intercepts radiation which is converted into biomass according to Monteith (1977). Model growth is dependent on nitrogen and water deficit, according to the functions of stress response developed by Sinclair (1986). Crop development is predicted based on phenology models validated using data covering 20 years and connected to a database containing the model parameter values for the 183 bread wheat varieties from our panel. For the other varieties, the parameter values of the standard Soisson variety were used.

2.2.1 WSI simulation and clustering of the multi-environment trial

The physical properties of the soil were characterized in each trial (% clay, % silt, % sand and % organic matter, etc.) and used to link each trial to the Arvalis soil database. On-site and under-10km away weather stations were used to run the simulations. The weather-related data and irrigation regimes are presented in [Suppl. III.2C and III.2D](#).

WSI was simulated daily from emergence to physiological maturity. For each environment, WSI values were calculated as the mean of the sub-set of varieties present in each environment of the MET. Ward's hierarchical clustering analysis was performed using the *pvc* function (Suzuki and Shimodaira, 2006) of the R statistical package (R Development Core Team, 2011). The inertia gain described in Murtagh and Legendre, (2014) was used to define the optimal number of groups. A multiscale bootstrap procedure (10,000 runs) provided the approximately unbiased P-values and the bootstrap probability values for each cluster in the hierarchical clustering. As described in Bouffier et al. (2014), WSI thresholds were determined to assess light, moderate and high drought stress levels

along the crop cycle ($WSI > 0.67$, $0.33 < WSI < 0.66$ and $WSI < 0.33$, respectively). Thus, we defined the four major environment types (ETs) encountered in our trial network, characterized by their water stress pattern in Optimal Conditions (OPT), a late water deficit (LWD), a medium water deficit (MWD) and a high water deficit (HWD).

2.2.2 WSI simulation and frequency analysis at the scale of France

WSI dynamics were simulated using a large grid of representative growing conditions encountered in France by wheat producers. This simulation set was built with two-hundred and fifty-eight sites representative of French wheat production areas ([Suppl. III.3A](#)), and with the crop seasons of 1997 to 2017. Daily meteorological data were available for each site. One optimal sowing date/variety combination was selected for each site based on the expertise of agronomists. We extended the optimal sowing date to ± 7 days to obtain three different sowing dates: early, optimal and late. Regarding the physical properties (depth, soil texture), three different representative soil types per region were selected, i.e. with low, medium and high available water capacity ([Suppl. III.3B](#)). Finally, a total of 48,762 simulations were carried out (21 seasons \times 258 sites \times 3 sowing dates \times 3 water capacities).

For each combination, WSI was simulated daily from emergence to physiological maturity and averaged over 100°C-day periods. Simulations were carried out with CHN assuming no biotic constraints (weed, pest and disease) and no nitrogen limitation. To do this, we declared 60 kg.ha⁻¹ residual soil nitrogen on 1st January of each year and then applied three nitrogen inputs of 50, 100 and 50 kg.ha⁻¹ on 15th February, 15th March and 25th April. Soil moisture was initialized every year on 1st August at 80% humidity in each soil layer.

Ward's hierarchical cluster analysis was performed to identify the main scenarios occurring in France. An average water deficit pattern was calculated to describe each cluster.

2.3 Phenotypic Traits

Spikes per unit area (SA), thousand kernel weight (TKW), Grain yield (GY), plant height (PH) and heading date (HD) were measured directly in each trial, except SA and PH, which are missing in one location, Mon12 and Olm16, respectively. The number of grains per spike (GPS) was calculated as $GY / (TKW \times SA)$. SA was measured by counting the number of spikes on two-meter rows in two different areas within each plot. GY was calculated using the dry weight of the harvested area and adjusted to 15% humidity. PH was measured after flowering from the ground to the top of the spike (without awns, if any) on two spots within each plot. HD was scored according to the Z55 stage (Zadoks et al., 1974), when half of the ear had emerged from the flag leaf on half of the ears.

One hundred and ten lines were present in all the environments, the rest of them were in at least fifteen environments. All the lines were present at least once in the different ETs. Missing varieties and

missing scoring resulted in 7.5% of missing data in the total dataset (210 varieties, 35 trials and 6 traits).

2.4 Statistical analyses

2.4.1 The two-Stage GxE analysis

Different experimental designs were used, depending on the local constraints. To deal with this complexity, we decided to use the GxE two-stage approach described by Frensham et al., (1997). In the first stage, we focused on applying the best linear model to each specific trial design. In each trial with two complete replicates, adjusted means and heritabilities were computed using 2D P-spline mixed models as implemented in the R package SpATS (Rodriguez-Alvarez et al., 2018). This spatial model adjusts for both global and local trends simultaneously. Genotypes were treated as fixed for the calculation of the best linear unbiased estimator (BLUEs) or as random for the estimation of heritabilities. For the remaining environments (6 out of the 35), ASREML-R (Butler et al., 2009) was used to fit the models best-adapted to each experimental design ([Suppl. III.4](#)).

In the second stage, the BLUEs were combined with the standard error of the adjusted mean (s.e.m.) of the i th genotype in the j th environment. The following part describes the implementation of the BLUEs and s.e.m. in the different mixed models used in the next section. The mixed model is given by:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where \mathbf{y} is the adjusted mean from the first stage, $\boldsymbol{\beta}$ and \mathbf{u} are fixed and random effects, \mathbf{X} and \mathbf{Z} are the corresponding design matrices, and \mathbf{e} is the residual error associated with adjusted means. The random effects are assumed to follow multivariate normal distributions with mean = 0 and variance $\text{var}(\mathbf{u}) = \mathbf{G}$ and $\text{var}(\mathbf{e}) = \mathbf{V}$. If the adjusted means are sorted by environments and then by genotypes within environments, \mathbf{V} takes a block-diagonal form:

$$\mathbf{V} = \bigoplus_{j=1}^n \mathbf{V}_j,$$

where \bigoplus is the direct sum operator, n is the number of environments, and \mathbf{V}_j is the estimated variance-covariance matrix as $\text{diag}[(\text{s.e.m.}_{ij})^2]$, conditioned on the environment. \mathbf{V} plays the role of a weighting matrix as described in Möhring and Piepho, (2009).

2.4.2 Statistical analyses of the MET

The variance components corresponding to the genotypes (G: $i = 1 \dots 210$), the environments (E: $j = 1 \dots 35$) and the environment types (ETs: $k = 1 \dots 4$) were modeled with the two following mixed-models:

[1] the classic compound symmetry model to test the $G_i \times E_j$

$$Y_{ij} = \mu + E_j + G_i + G_i \times E_j + \varepsilon_{ij}$$

[2] a compound symmetry model to test the $G_i \times ET_k$ term

$$Y_{ij} = \mu + E_j + G_i + G_i \times E_j + G_i \times ET_k + \varepsilon_{ij}$$

In both equations (1) and (2), Y_{ij} is the adjusted mean of the i th genotype in the j th environment, μ is the intercept, E_j is the random environment effect with $E_j \sim \mathcal{N}(0, \sigma_{E, nE}^2)$, G_i is the random genotype effect with $G_i \sim \mathcal{N}(0, \sigma_g^2 K)$ where K is the genomic relationship matrix calculated using VanRaden's (2008) formula, $G_i \times E_j$ is the random genotype-by-environment interaction effect with $G_i \times E_j \sim \mathcal{N}(0, \sigma_{g.e}^2 \cdot K \otimes I_{nE})$, $G_i \times ET_k$ is the random genotype-by-environment type interaction effect with $G_i \times ET_k \sim \mathcal{N}(0, \sigma_{g.et}^2 \cdot K \otimes I_{nET})$ and ε_{ij} is the residual trial error $\varepsilon_{ij} \sim \mathcal{N}(0, \sum_{nE} \otimes \mathbf{V})$, where \sum_{nE} represents the heteroscedastic residuals with an environment specific error variance, \mathbf{V} the weighting matrix defined previously, and the symbol \otimes denotes the kronecker product of the two matrices.

The variance components of random factors were tested one by one using the likelihood ratio test (LRT) (Kendall and Stuart, 1979), based on log-likelihood (Lmax) differences between the complete and the reduced model without the test factor. The variance components were extracted and the standard deviations expressed as a percentage of the general phenotypic mean (Gelman, 2005). More or less twice the standard deviation was used to represent the interval including 95% of the values around the mean. This range was also expressed as a percentage of the general mean.

2.4.3 Heritabilities

Generalized heritability (h^2_g) from the first stage of the MET analysis was calculated for each trial using the following formula developed by Cullis et al. (2006),

$$h^2_g = 1 - PEV / \sigma_g^2$$

where σ_g^2 is the genetic variance and PEV is the average pairwise prediction error variance of the best linear predictions (BLUPs) of genetic effects.

2.4.4 Genotypic value extraction from Environment Types

To test the relationship between traits, estimate the genetic progress and perform genome wide association studies (GWAS), we modeled genotypes and genotype by environment type as follows:

$$[4] Y_{ij} = \mu + G_i + G_i \times ET_k + \varepsilon_{ij}$$

Where Y_{ij} is the adjusted mean of the i th genotype in the j th environment, μ is the intercept, G_i and $G_i \times ET_k$ are used as fixed effects to avoid shrinkage issues and assume that $G_i \times E_j$ interactions that are not taken into account by $G_i \times ET_k$, are included in the residual, and ε_{ij} is the residual trial error

$\varepsilon_{ij} \sim \mathcal{N}(0, V)$, where V is the weighting matrix defined previously. The genotypic adjusted means for each environment type were then computed as:

$$[5] \overline{GET}_{ik} = \hat{\mu} + \hat{G}_i + G_i \times \overline{ET}_k$$

The performance of varieties at the French scale were estimated with a weighted arithmetic mean as $GET_i = \sum_1^K \overline{GET}_{ik} \times \omega_k$, where ω_k is the frequency of the k th ETs from the French WSI simulation.

2.5 Genetic progress

Because only five varieties were released between 1974 and 1992, they were not included in the following analysis, to avoid sampling bias. In all, 205 elite European lines were used to estimate the genetic progress from 1992 to 2011 on grain yield, yield components and plant height. Plant phenology is often correlated to the traits studied (Cormier et al., 2013), so in the first stage of the GxE analysis defined previously, we used a second order polynomial regression to correct for genotype heading dates in each environment ([Suppl. III.4](#)). The effect of the year of release (YR) was assessed by variance analyses (ANOVA) in a linear model:

$$[6] \overline{GET}_{ik} = \mu + ET_k + YR_i \times ET_k + \varepsilon_{ik}$$

Where \overline{GET}_{ik} is the adjusted mean of the i th genotype in the k th environment types defined in Model 5, μ is the intercept, ET_k is a fixed effect of the environment types, $YR_i \times ET_k$ is the interaction between the year of release of the i th genotype and the environment types ET_k , ε_{ik} is the residual error defined as $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_e^2)$.

2.6 GWAS analysis

All the lines were genotyped with the Affymetrix Axiom 280K SNPs array (Rimbert et al., 2018). Only the Polymorphic High-Resolution SNPs were used in this analysis. All the SNPs were physically mapped on the genome reference sequence RefSeq V1.0 (Appels et al., 2018). Heterozygous loci were considered as missing data. Monomorphic SNPs, SNPs with more than 10% missing data, and SNPs whose minor allele frequency (MAF) was below 5% were discarded. Missing data was imputed with Beagle 4.1 (Browning and Browning, 2016). GWAS was performed with FaST-LMM v2.07 (Lippert et al., 2011a). To sum up, 161,772 SNPs were tested on the 6 traits measured on 210 varieties in the four environment types. SNP-trait association studies were performed with the single locus mixed model (Yu et al., 2006):

$$Y_i = \mu + x\beta + g_i + \varepsilon_i$$

Where Y_i is the vector of phenotypic values (eq. 5), μ the overall mean, x is the vector of SNP scores, β is the additive effect of the SNP, and g_i and ε_i are the random polygenic and residual effects as $g_i \sim \mathcal{N}(0, \sigma_g^2 K_{chr})$ and $\varepsilon_i \sim \mathcal{N}(0, \sigma_e^2)$, respectively.

The SNP effects β were estimated by generalized least squares, and their significance ($H_0: \beta = 0$) was tested with an F-statistic. An SNP was considered significantly associated with the trait if the P-value was lower than 0.001. False discovery rate (FDR) adjusted q-values were also calculated using the R-package *qvalue* to estimate the variation of the rate of false positives depending on the trait. The variance explained by each SNP tested was estimated according to Cox and Snell (1989) as $R^2 = 1 - \exp\left(2 \times \frac{LL0-LL1}{n}\right)$, where n is the size of the panel, and LL0 and LL1 are the log-likelihood of the reduced and full models, respectively.

The random polygenic effect was modeled with a kinship matrix (K_{chr}) as a variance/covariance matrix. Pairwise kinship coefficients were computed according to the first method described by VanRaden (2008), with a sub-set of 27,680 SNPs obtained by removing highly correlated SNPs. In brief, we computed r^2 for each pair of SNPs located on the same chromosome. A dendrogram was built with this matrix of r^2 and one SNP was randomly selected from each cluster with a distance <0.1 (with distance= $1-r^2$). To avoid proximal contamination, SNPs were tested using a K_{chr} matrix computed with SNPs that were not located on the same chromosome as the SNP tested (Rincent et al., 2014).

The genetic structure was studied through a PCA on kinship and an admixture analysis with the STRUCTURE software (Pritchard et al., 2010) using the sub-set of 27,680 SNPs ([Suppl. III.5A](#)). Four different groups were identified in this panel. In line with Le Couviour et al., (2011) and Cormier et al. (2014), who performed genetic diversity analyses and GWAS on a similar panel, we found that adding significant PCA axes ([Suppl. III.5B](#)) as fixed effects in the GWAS model did not improve structure correction according to the QQ-plots ([Suppl. III.5C](#)). We therefore decided to keep only the polygenic random effect in the GWAS model.

QTL boundaries from GWAS were defined following the method described in Cormier et al., (2014). Briefly, SNP belonging to the same linkage disequilibrium (LD) cluster were defined as a group of quantitative trait nucleotides (QTN). The clustering was performed by averaging r^2 distance and the tree was cut at 1-“critical LD” (critical LD = 0.24 (Breseghello, 2005)). QTL boundaries were defined as the maximum and minimum map positions of SNP within the QTN extended by the LD decay specific to the genomic area of the QTL. QTL colocalizations were analyzed between all combinations of phenotypic trait and ETs to assess the impact of pleiotropy and genetic correlation. Only QTLs smaller than 10% of the chromosome length (95% of the QTLs) were used in this analysis. A QTL was defined as specific if there was no colocalization with another trait in any of the four ETs. Based on this colocalization network, specific QTLs and QTLs associated with GY and which did not colocalize with heading date and plant height were selected and their standardized effect was represented on a heatmap. The SNPs effects were standardized by dividing them by the standard error of their estimation.

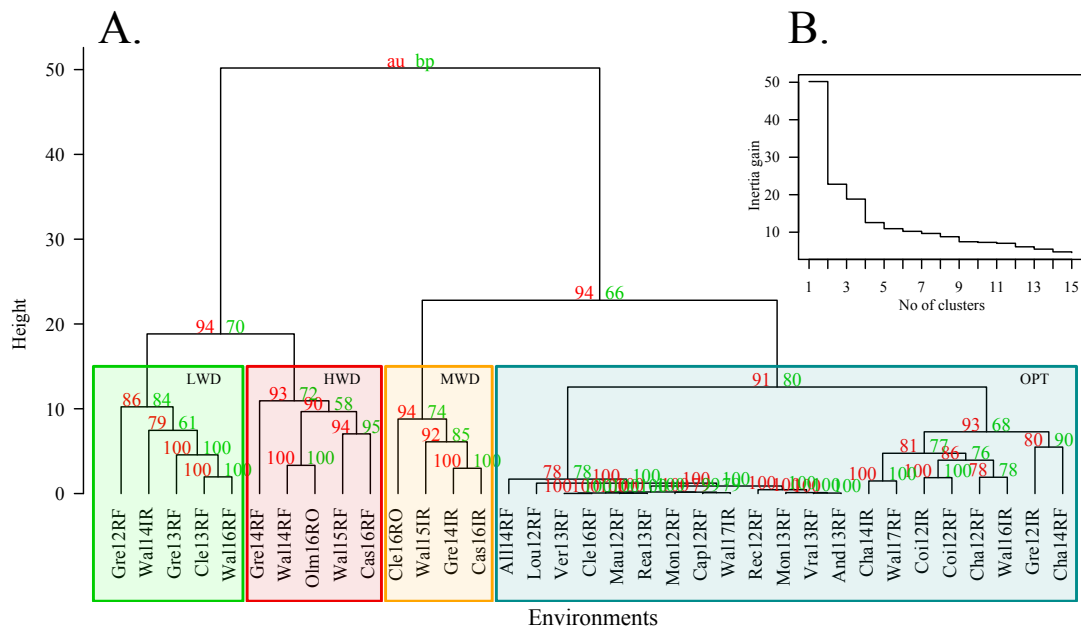


Figure III.1: Environment types dendrogram based on water stress index patterns

A. Ward's hierarchical clustering on the set of water stress indexes for each environment in the MET as the mean of the balance genotype. The uncertainty for each cluster in hierarchical clustering was calculated via multiscale bootstrap (10,000 run), it provides the approximately unbiased P-value in red and the bootstrap probability value in green. B. The inertia gain described in Murtagh and Legendre, (2014) was used to define the optimal number of groups. Environment types (ETs) with: in blue, optimal condition (OPT), in green, late water deficit (LWD), in orange, moderated water deficit (MWD) and in red, high water deficit (HWD).

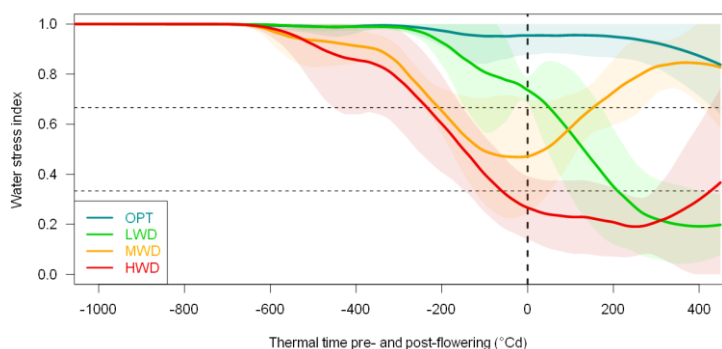


Figure III.2: Environment types water stress index throughout the crop cycle Water-stress index (WSI) throughout the crop cycle, represented as a function of thermal time relative to flowering times. WSI were averaged from environments within environment types, the shaded area refers to this average ± 1 standard deviation. Lines and area colors depend on the environment types: in blue, optimal condition (OPT); in green, late water deficit (LWD); in orange, moderated water deficit (MWD); and in red, high water deficit (HWD).

3 RESULTS

3.1 Environmental clustering identified four types of drought pattern in the MET

The 35 environments were clustered using a water stress index computed with a crop model. Following the inertia gain from the dendrogram of hierarchical clustering, we decided to define four groups of environments ([Figure III.1A and B](#)). As can be seen in [Figure III.2](#), each group is classified according to the timing and the intensity of the drought stress undergone by the crop: the largest group (21 environments) corresponded to optimal conditions (OPT) where the varieties were not substantially limited by water availability and there was no change in the WSI pattern from the sowing date to anthesis, although it decreased very slightly during grain filling. The second group (5 environments) was defined as a late water deficit stress (LWD) group, with the WSI pattern steadily decreasing from 400°Cd before anthesis and falling below 0.66 after anthesis. The third group (4 environments) showed moderate stress (MWD) around anthesis, relieved by rainfall events during grain filling. The fourth group (5 environments) showed a high water deficit (HWD) starting during the vegetative period and increasing throughout grain filling.

3.2 Pattern and frequencies of drought scenarios in France

Analysis of the frequency of drought scenarios occurring in France helped us to identify which MET water patterns were representative of French agricultural practices and pedoclimatic conditions. Ascendant hierarchical cluster analysis was carried out on the 48,762 simulations and the resulting dendrogram showed that different numbers of groups could be chosen ([Figure III.3](#)). Clearly, there were two main groups: one group with no or moderate water stress representing 45% of the simulated situations and one group with different levels of water stress representing 55% of the simulated situations ([Figure III.3A](#)). The inertia gain indicated that five groups provided the optimal clusterization ([Figure III.3B](#)). With five groups, it can be seen that the first cluster of environments with no or light water stress was split into two sub-groups: the largest represented 29% of the simulated situations and the average WSI pattern showed a light water stress at the end of the crop cycle; the other group represented 26% of the simulated situations, with an average WSI close to one during the whole crop cycle (no water stress). Stressed environments were split into three sub-groups, roughly corresponding to the three stress patterns found in the MET analyzed in this study. The largest of these stressed sub-groups represented 21% of the simulated situations corresponding to the LWD group. Each of the two other sub-groups represented 12% of the simulated situations and corresponded to the MWD and HWD groups ([Figure III.3C](#)).

3.3 Environment types helped to understand grain yield structure and GxE interactions

Mean grain yield in the 35 environments of the MET ranged from 4.26 to 10.07 tonnes.ha⁻¹ in Gre12RF and Ver13RF, respectively. The environmental coefficient of variation (CVe) indicates that

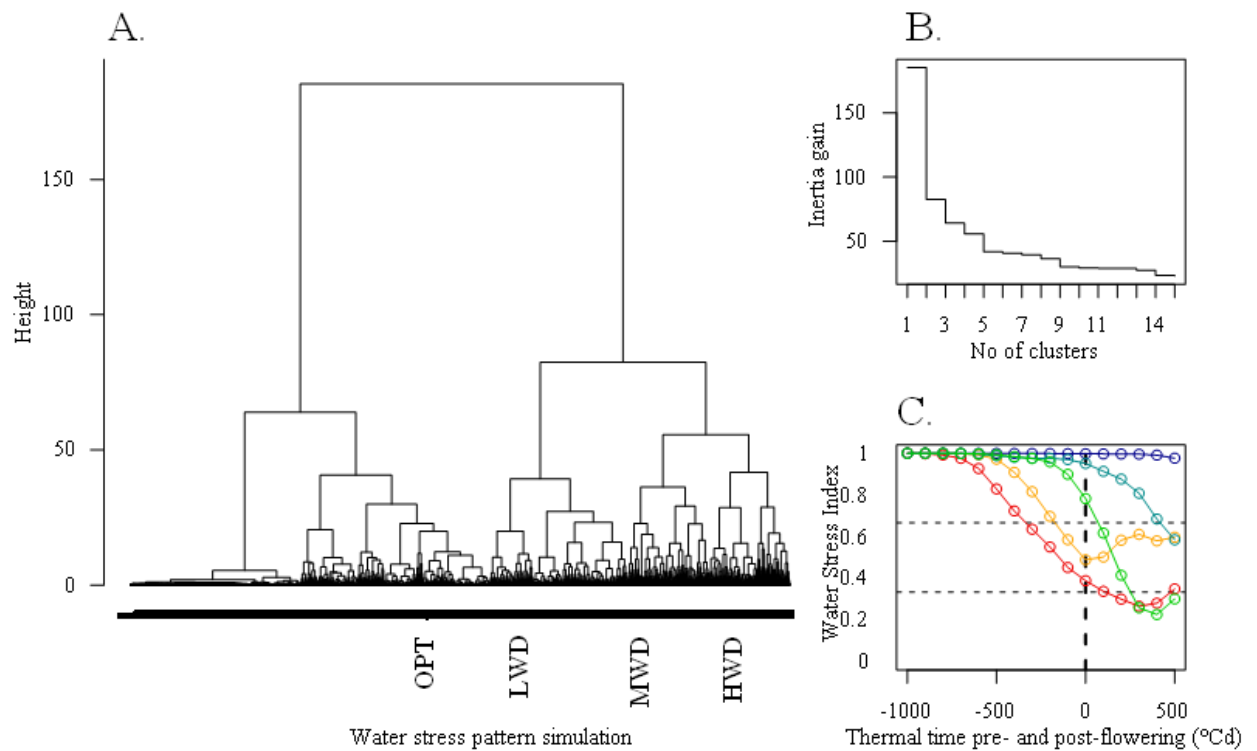


Figure III.3: Dendrogram based on the water stress index (WSI) patterns simulated for more than 40000 scenarios occurring in France plus ETs patterns

A. Ward's hierarchical clustering on the 48,762 simulation of wheat growth over France over the last 20 years; the set of water stress indexes was averaged every 100°Cd. B. Inertia gain distribution of height of tree branches. C. WSI throughout the crop cycle represented as a function of thermal time relative to flowering times. WSI were averaged for each of the five clusters.

the phenotypic data is of good quality, as C_ve ranged from 2.63% to 12.15% for Ver13RF and Wal14RF, respectively ([Table III.1](#)).

The variability of grain yield and yield components between ETs is shown in [Table III.2](#) and a visual representation in [Suppl. III.6](#). By comparing the three drought related scenarios with the optimal one, we observed a decrease of GY of 3.3% for LWD, 12.4% for MWD and 31.2% for HWD. Similar analyses conducted with yield components gave consistent results: SA was significantly affected only in the MWD and HWD clusters by 12.0% and 6.0%, respectively. GPS decreased by 9.4% in the HWD cluster compared to OPT, while it increased in the LWD cluster by 6.7%. TKW was significantly affected only in the LWD and HWD clusters by 13.0% and 17.2%, respectively. This was confirmed by adding the GxET interaction in Model 2, which was significant for grain yield and yield components, except for SA, and captured 20.7%, 5.7% and 12.7% of the GxE interaction for GY, GPS and TKW, respectively.

Significant genotypic effects were observed for grain yield and yield components ([Table III.2](#)). When examining the genetic variation (G) of grain yield, the panel showed large variability with 95% of individuals between $\pm 7.82\%$ of the mean. The variance decomposition revealed significant GxE interactions for all traits except SA. The most significant GxE interaction was for GPS where GxE interaction was $\pm 12.04\%$ of the mean. In the classical analysis (Model 1), the ratio of G on GxE interaction variances was around 46% for grain yield ([Table III.2](#)). When we applied Model 1 to GY within the ETs, this proportion increased substantially to 84%, 69% and 83% for OPT, LWD and HWD, respectively, but was lower (43%) in the MWD cluster ([Table III.3](#)).

3.4 Relationship between agronomic traits varies within environment types

Using the genotype adjusted means per ETs (Mod. 5), a biplot analysis was conducted to investigate relationships among traits and genotypes for each ETs ([Figure III.4](#)). Vector length showed the extent of variation explained by each trait in the PCA. The first two axes explained up to 62%, 64%, 57% and 54% of the total variability in OPT, LWD, MWD and HWD, respectively. In the OPT cluster, grain yield showed a positive association with HD and GPS, while in LWD and HWD grain yield was positively correlated to TKW and negatively to HD. In MWD there was no clear association between grain yield and the other yield components.

Genotypes were ranked by their yield potential in each ET and the names of the top and bottom five genotypes were shown on [Figure III.4](#). No genotype is present in the top five in at least two clusters. In contrast CENTENAIRE and ANTONIUS were in the bottom five genotypes in three and two ETs, respectively.

Pearson's correlation studies were conducted between traits within and between ETs ([Suppl. III.7](#)). Under the OPT condition, GY was positively correlated to HD ($r = 0.50$, $p < 0.001$), while the

Table III.2: Variance components of the different models for grain yield and its components

Variance components were estimated with the two models M1 and M2 in which phenotypic variance is decomposed into three or four components, E and G, the additive environmental and genetic effect, and their interaction GxE. Genetic effect by environment types interaction (GxET) is present only in Model 2 (M2). Confidence Intervals (CI) were calculated as twice the square root of the variance component divided by the general mean and expressed as a percentage. This CI shows us how the random term affects the mean. Res is the residual variance.

Trait		Mod.	Mean	E	G	GxET	GxE	Res
Grain Yield (t/ha)	Variance	M1	7.93	1.75***	0.10***		0.21***	0.99
		M2	7.93	1.67***	0.07***	0.06***	0.16***	0.99
	CI of mean in %, ($\alpha = 0.05$)	M1	7.93	33.38	7.82		11.44	25.04
		M2	7.93	32.62	6.62	6.04	10.02	25.16
Spike per area (sp/m²)	Variance	M1	529.81	9354***	951.4***		490.8***	2030
		M2	529.67	9358***	938.3***	12.88ns	478.8***	2035
	CI of mean in %, ($\alpha = 0.05$)	M1	529.81	36.7	11.70		8.41	17.08
		M2	529.67	36.7	11.62	1.36	8.29	17.11
Grain per spike (gr/sp)	Variance	M1	38.39	30.23***	7.94***		5.38***	112.15
		M2	38.39	30.27***	7.68***	0.39**	5.11***	112.40
	CI of mean in %, ($\alpha = 0.05$)	M1	38.39	28.56	14.64		12.04	55.02
		M2	38.39	28.58	14.40	3.22	11.74	55.08
Thousand kernel weight (g)	Variance	M1	40.35	22.79***	7.63***		2.23***	3.43
		M2	40.35	22.11***	7.69***	0.53***	1.93***	3.37
	CI of mean in %, ($\alpha = 0.05$)	M1	40.35	23.66	13.70		7.40	4.67
		M2	40.35	23.30	13.74	3.60	6.88	4.63
Heading date (days after 1st January)	Variance	M1	141.52	212.82***	11.79***		1.35	0.01
		M2	141.53	195.58***	12.16***	0.15***	1.28	0.01
	CI of mean in %, ($\alpha = 0.05$)	M1	141.52	20.62	4.85		1.64	0.03
		M2	141.53	19.76	4.92	0.54	1.60	0.03
Plant height (cm)	Variance	M1	81.86	93.49***	30.76***		4.89***	7.29
		M2	81.86	73.76***	30.45***	1.11	4.46***	6.85
	CI of mean in %, ($\alpha = 0.05$)	M1	81.86	23.61	13.55		5.40	6.59
		M2	81.86	20.98	13.48	2.56	5.16	3.39

LTR test : ***, *P-values* <0.001 ; **, *P-values* <0.01 ; *, *P-values* <0.05 and ns non-significant *P-value*>0.05

correlation was negative in the other clusters, with $r = -0.66, -0.26, -0.28$ ($p < 0.001$) for LWD, MWD and HWD, respectively. Regarding the relationship between grain yield and yield components, the number of grains per area showed the highest correlation with GY in OPT ($r = 0.48, p < 0.001$), while TKW was the most correlated to GY in LWD and HWD ($r = 0.52$ and 0.44 , respectively, $p < 0.001$) and SA in MWD ($r = 0.33, p < 0.001$). There was no significant correlation between GY in OPT and either LWD or HWD, but a weak positive correlation with GY in MWD ($r = 0.26$ ($p < 0.001$)).

3.5 Genetic gains for grain yield varied within environment types

The genetic progress was assessed by testing the effect of the year of registration (YR) for the different traits with Model 6 ([Suppl. III.8](#)). No significant effect of year of registration was found for SA and TKW, but a significant positive effect was found in the OPT cluster for GPS (genetic gain = $+0.13$ grains/spike/year). Regarding GY ([Figure III.5](#)), we found an increase in each cluster although it was not significant in the LWD and MWD clusters. The highest significant and positive slope were found in the OPT and HWD clusters, where the genetic gains were $+38.07 \text{ kg}\cdot\text{ha}^{-1}\cdot\text{year}^{-1}$ (± 6.6) and $+30.65 \text{ kg}\cdot\text{ha}^{-1}\cdot\text{year}^{-1}$ (± 11.8), respectively.

3.6 GWAS results

Overall, 9,908 SNPs were significantly associated (P -values < 0.001) with at least one of the six traits in at least one of the ETs. When considering LD blocks between associated SNPs and LD around associated regions, 590 QTLs were mapped with a median size of 3.8 Mb ([Suppl. III.9A](#)). The Q -value from the FDR was smaller than 0.2 for 134 of them (a presentation of the relationship between Q -values and P -values is given in [Suppl. III.9B and C](#)). Ninety percent (between the 5th and 95th percentile) of the QTLs had a size between 0.3 and 77.3 Mb, corresponding to 0.05 and 10.90% of chromosome size. The number of QTLs on genome D (126) was smaller than on genomes A (238) and B (226). Homoeologous group 2 had the highest number of QTLs with 126 QTL. [Table III.4](#) shows the distribution of the QTLs along the genome. The number of QTLs by trait ranged from 10 for SA in OPT to 34 for TKW in MWD. The variance explained by each significant SNP was 6.5% on average, ranging from 4.3% to 34.4% for HD on 5A and 2D, respectively.

3.7 Colocalization network

The number of QTLs which colocalize with at least another one QTL, varied from 48% for SA in the OPT to 92% for HD in the LWD ([Table III.5](#)). Twenty-six specific QTLs linked to GY were identified. The variance explained by all the QTLs was 6.5% on average, ranging from 4.8% to 8.8% in MWD and HWD, respectively.

Major regions of colocalization were on chromosomes 2D, 2B, 6A ([Suppl. III.10](#)). The number of common QTLs between two traits were different between environment types. For example, there were 12 QTLs in common between GY and HD in the MWD cluster, while only 2 QTLs were found in common in the HWD cluster ([Table III.5](#)). Within environment type OPT, HD was the trait with the

Table III.3: Variance components within environment types

Variance components were estimated with the model M1 within ETs, variance is decomposed into three components: E, G and GxE, i.e. the additive environmental and genetic effect and their interaction, respectively. Res is the residual variance. OPT: optimal condition; LWD: late water deficit; MWD: medium water deficit; HWD: high water deficit. GY: grain yield, SA: spike per area, GPS: grain per spike, TKW: thousand kernel weights, HD: heading date, PH: plant height.

ETs	Pheno	Units	Mean	E	G	GxE	Res
OPT	SA	Spike per unit area	535.20	5012.06***	1128.03***	656.27***	2251953.41
	GPS	Nb grain per spike	38.58	39.44***	8.48***	5.23***	118.83
	TKW	g	42.04	11.79***	7.55***	1.86***	3.69
	GY	Tonnes.ha ⁻¹	8.42	0.86***	0.14***	0.17***	0.91
	HD	Days after 1 st January	143.99	56.54***	9.18***	1.25***	0.93
	PH	cm	84.93	65.26***	31.66***	4.04***	5.06
LWD	SA	Spike per unit area	550.04	3385.66***	1062.21***	964.59***	1781810.53
	GPS	Nb grain per spike	41.19	5.99***	7.53***	8.75***	51.29
	TKW	g	36.60	13.71***	8.87***	1.48***	3.72
	GY	Tonnes.ha ⁻¹	8.15	1.15***	0.13***	0.17***	2.67
	HD	Days after 1 st January	142.16	41.51***	12.47***	2.54***	0.78
	PH	cm	73.73	29.86***	29.86***	8.05***	330.48
MWD	SA	Spike per unit area	477.21	29326.35***	534.96***	297.05**	845822.70
	GPS	Nb grain per spike	39.09	18.03***	7.41***	5.03***	46.96
	TKW	g	42.25	17.94***	10.43***	1.67***	1.22
	GY	Tonnes.ha ⁻¹	7.38	1.75***	0.09***	0.2***	2.39
	HD	Days after 1 st January	132.62	36.26***	15.92***	1.63***	0.19
	PH	cm	82.48	52.64***	26.48***	6.54***	38.24
HWD	SA	Spike per unit area	507.76	22451.28***	671.67***	572.14***	1076663.43
	GPS	Nb grain per spike	34.78	22.69***	6.71***	3.95***	85.46
	TKW	g	34.80	31.12***	6.18***	3.23***	8.35
	GY	Tonnes.ha ⁻¹	5.79	1.43***	0.1***	0.12***	2.01
	HD	Days after 1 st January	136.43	63.60***	12.55***	0.80***	1.48
	PH	cm	71.79	103.06***	26.43***	5.68***	35.61

LTR test : ***, P -values <0.001 ; **, P -values <0.01 ; *, P -values <0.05 and ns non-significant P -value >0.05

highest number of QTLs positively associated with GY, while this was the case with GPS, TKW and SA in the LWD, MWD and HWD clusters, respectively. A visual representation of the frequencies of QTL colocalization can be found in [Suppl. III.10](#). None of the QTLs identified in this study were associated with GY in any of the ETs.

A heatmap illustrating the levels of the standardized SNP effect of the minor allele can be seen in [Figure III.6](#). This figure shows a subset of 53 SNPs significantly associated with GY at least once over the different ETs. We can see three different trends: SNPs with positive effects in all the ETs(13.2%); SNPs with negative effects in all the ETs (26.4%); SNPs with contrasted effects depending on the cluster (60.3%).

4 DISCUSSION

The development of drought tolerant wheat elite genotypes can be both an asset and a goal for wheat breeders in Europe. Screening for drought resistance in the material already registered is an effective way of selecting adapted material and start crossing in order to cumulate the most favorable alleles. To this end, this study evaluated 210 elite European varieties in 35 environments for drought resistance.

4.1 Environment types improved understanding of grain yield structure and GxE interaction

Ten experiments were conducted with two water treatments (irrigated *vs* rainfed). We could have used these experiments to perform an analysis by comparing genotypes under two water regimes, i.e. irrigated *vs* non irrigated (Bennett et al., 2012; Pinto et al., 2010; Zhang et al., 2018). However, this type of analysis is only appropriate if treatments are equivalent throughout the experiments, which is not the case here as the timing of the drought stress was different between experiments. Therefore, we decided to characterize the drought scenarios in a large MET including these ten experiments to explore different drought stress scenarios. This kind of complex MET can be analyzed in different ways: (1) Yield-based characterization could be useful for breeders (Desclaux, 1996; Brancourt-Hulmel, 1999) as grain yield is the main target and integrates all the environmental factors. This kind of characterization is very efficient, but this level of integration could be an obstacle when identifying the main limiting factors. (2) Partitioning GxE interaction using quantitative environmental covariates could also be done, but selecting the main environmental covariates, either pedo-climatic variables or related to a constraint (nitrogen fertilization, frost/heat during a specific stage or biotic stresses) remains a problem. (3) Crop growth models (Chapman 2008; Chenu et al. 2011; Heslot et al. 2014; Ly et al. 2017) are also an efficient tool for identifying a limited set of environmental variables known to be biologically and agronomically relevant. All in all, choosing the most appropriate environment clustering method is central to characterizing a MET and we chose to classify our MET based only on drought stress.

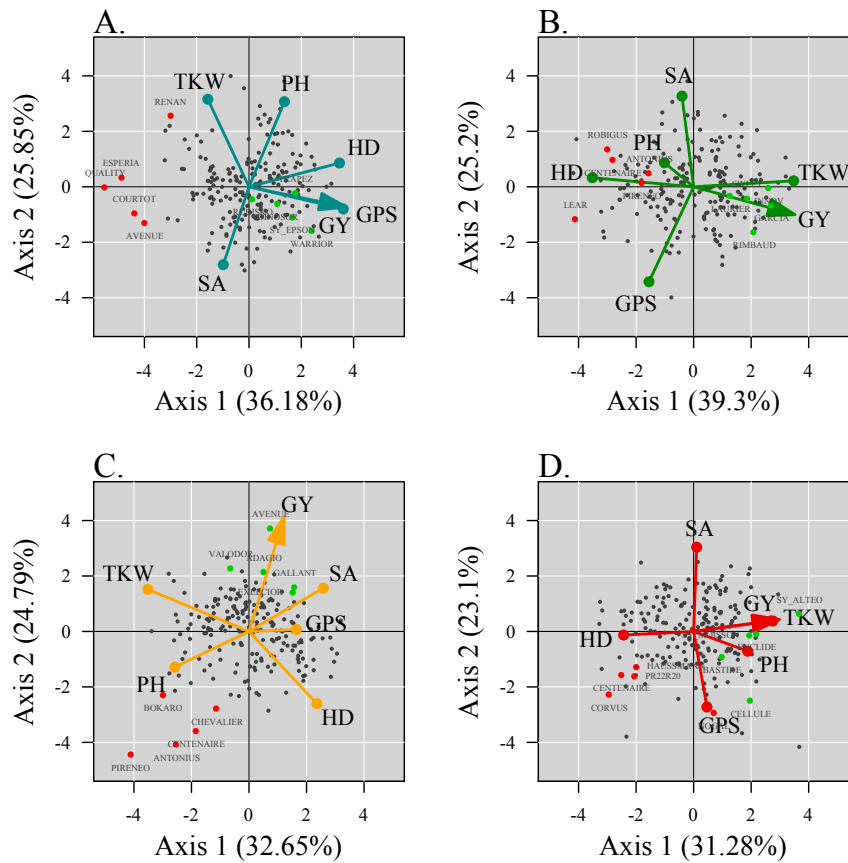


Figure III.4: Biplots based on agronomic traits and earliness for each environment type

Scatterplot showing line plots calculated on agronomic traits according to the first two principal components. Vector length showed the extent of variation explained by each trait in the PCA. Colors represent the different ETs related to water stress pattern, OPT: optimal condition (blue); LWD: late water deficit (green), MWD: medium water deficit (orange), HWD: high water deficit (red). The top and last five lines for GY are plotted in green and red dots, respectively. GY: grain yield, SA: spike per area, GPS: grain per spike, TKW: thousand kernel weights, HD: heading date, PH: plant height.

Four different patterns of drought stress were identified, based on a daily crop water stress index (ratio between water supply and demand). Sixty percent of the trials were grouped in an optimal condition (OPT) cluster with no water stress. The other forty percent were grouped in three different water stress patterns: a late water deficit (LWD) with a stress during anthesis and grain filling, a medium water deficit (MWD) with a stress at anthesis, and a high water deficit (HWD) starting during the vegetative stage and extending during the grain filling stage. The water stress in the LWD cluster mainly affected TKW, while it essentially affected the number of spikes per area in MWD, and all the grain yield components (number of spikes, grains per spike and grain weight) in the HWD cluster. The impact of stress on the grain yield components agreed with the literature (Gate, 1995; Meynard and Sebillotte, 1994).

Examining the impact of the drought stress scenario on grain yield components is the first step in validating the classification method. However, this type of method can only be useful to breeders if the environment types have a substantial influence on GxE interaction (Löffler et al., 2005; Hammer and Jordan, 2007; Chenu et al., 2011). [Table III.2](#) shows that more than 20% of the GxE interaction can be explained by incorporating the environment type effect in the mixed models. Furthermore, the ratio between G and GxE interaction almost doubled in OPT, LWD and HWD. These results are equivalent to the ratio identified in the stress-free environment type 1 (ET1) describing the North-Eastern Australia drought stress pattern defined by Chenu et al., (2011) ([Table III.3](#)). Furthermore, this is encouraging and confirms the interest of reducing GxE interaction through drought stress characterization in the European climate context.

However, the ratio between G and GxE interaction remains at the same level for grain yield in MWD. This was also found in other studies (e.g. ET2 and ET3 Chenu et al., (2011);). This could be explained by different non-exclusive factors: (1) the large difference between the stress within this cluster. Indeed, in this cluster, the number of grains per spike was impacted but rains after anthesis led to maintaining thousand-kernel weights and partly contributed to compensating the reduced of grain yield. This assumption was clearly identified in the trial Cle16RO, in which the biomasses at flowering fell by 50% compared to the rain-fed part (Cle16RF) while the grain yield was only 12% lower. The other trials within the MWD cluster showed a lighter water stress than Cle16RO. Water stress was probably not the main stress in these other trials, so other environmental factors could have impacted grain yield such as soil nitrogen availability and thus be sources of GxE interaction. (2) The WSI pattern for MWD, presented in [Figure III.2](#), showed the largest standard deviation around flowering, which is known to be the most sensitive stage in wheat. (3) The different flowering dates within the panel could also be a source of GxE interaction, especially in the cluster where the stress occurred during flowering.

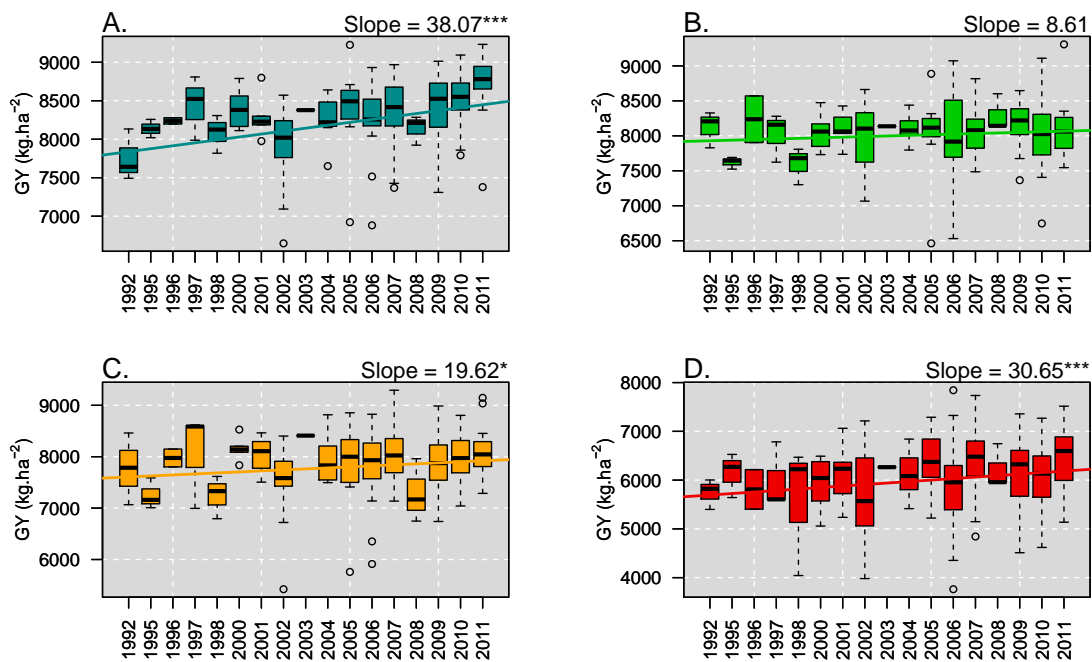


Figure III.5: Boxplot of Grain Yield performance corrected for earliness as a function of registration years of 205 wheat lines grown in the four ETs

A. optimal condition (OPT); B. late water deficit (LWD); C. moderate water deficit (MWD); D. high water deficit (HWD). Slope (kg.ha⁻².year⁻¹) of the regression lines is written on the top right of each figure and was estimated with equation 6 and tested with t-test as H₀: slope = 0, see Supp. 5. ***, P-values <0.001; **, P-values <0.01; *, P-values <0.05 and ns non-significant P-value>0.05 Colors represent the different ETs related to water stress pattern, OPT, LWD, MWD and HWD.

Also, climatic variables often vary concomitantly, such as the negative correlation between maximum temperature and the water supply/demand ratio (Sadras et al., 2012b). This means that although we conducted our clustering on water stress index, there is a chance that we clustered environments that also share the same heat stress pattern. Thus, it is complicated to distinguish between drought stress resistance and heat stress resistance. Although the clustering worked well for our targeted stress, there are several ways of improving it. We characterized the MET using an average of the WSI pattern for all the genotypes present in all the trials, but our panel contains variable heading dates (with a mean at 138 days ± 6) and the stress pattern was significantly influenced by flowering date. An early variety will not have exactly the same water stress pattern as a late one, especially in the LWD cluster where the early varieties may have escaped the drought stress.

In this study, we chose to cluster the MET on the WSI pattern found in each trial and then compared it with the historical WSI simulated over the last 20 years in France. Our results show that each environment type in the MET could be easily associated with a cluster identified with the historical data. Chenu et al., (2011) proposed a classification of the MET trials based on their similarity with the historical simulated environments in the TPE. However, although our strategy was not more efficient in explaining GxE interaction, it was easier to link the water stress pattern to a real stress impact on grain yield and yield components. For example, the OPT cluster was selected as one group, but when examining the bootstrap statistic and the historical data, it may have been separated into two groups: one without stress all along the crop cycle and one with light stress at the end of grain filling. In the latter, no loss of thousand kernel weights compared to the OPT cluster was identified in the MET. Moreover, the inertia gain ([Figure III.1B](#)) did not show any clear improvement in variance partitioning when splitting the OPT cluster into two groups.

A possibility offered by our methodology is that of weighting the genotype performance with the stress frequency of the four clusters to identify the potentially best varieties. By doing this, we found that the best French variety of our panel was “Rubisco”, which was registered in 2012. "Rubisco" was indeed the top sowing variety during the 2016/2017 season, cultivated on 12.8% of the sowing surface (FranceAgriMer, 2017). Results such as the identification of “Rubisco” strengthen our confidence in the validity of the clustering.

4.2 Genetic progress between 1992 and 2011

This work has been carried out on a large collection of elite European winter varieties, released from 1992 to 2011. The effect of selection was assessed as a function of additive genetic value and GxET interaction terms. In accordance with the recommendation from Cormier et al., (2013), we controlled variability due to precocity using heading date as a covariate.

Based on the yield stagnation observed in European countries since the end of the 1990s (P. Peltonen-Sainio et al., 2009; Finger, 2010; Brisson et al., 2010), several studies have claimed that yields are

Table III.4: QTL detected on the wheat association panel for 6 traits in the 4 environment types

Following Cormier et al., (2014), QTL boundaries were defined as the minimum and maximum physical position of SNPs belonging to the same LD block, extended by the LD decay assessed in the area. GY: grain yield, SA: spike per area, GPS: grain per spike, TKW: thousand kernel weights, HD: heading date, PH: plant height. ETs related to water stress pattern, OPT: Optimal condition (blue); LWD: Late Water Deficit (green), MWD: Medium Water Deficit (orange), HWD: High Water Deficit (red).

ETs	TRAIT	QTL		Effect (%)		MAF		Size		Group1			Group2			Group3			Group4			Group5			Group6			Group7		
		Total	% Pos.	mean	sd	mean	sd	mean	sd	A	B	D	A	B	D	A	B	D	A	B	D	A	B	D	A	B	D	A	B	D
OPT	HD	30	26.67	0.54	0.14	0.27	0.14	7.51	12.30	2	3	0	1	1	3	1	3	0	3	1	0	2	1	0	1	2	3	1	1	1
OPT	PH	25	60.00	2.05	0.45	0.22	0.16	9.04	12.78	1	0	1	1	4	2	1	0	0	1	1	2	1	1	1	1	0	0	2	5	0
OPT	SA	29	48.28	2.73	0.19	0.26	0.16	15.21	30.34	2	2	1	0	1	0	4	1	2	4	0	0	2	1	2	1	0	1	4	0	1
OPT	GPS	23	56.52	3.55	0.42	0.29	0.15	25.85	63.00	0	2	1	1	1	4	2	0	1	2	0	0	0	4	0	0	2	0	2	1	0
OPT	TKW	26	69.23	2.20	0.26	0.28	0.14	32.32	73.12	3	2	1	0	1	2	0	2	2	2	0	0	0	5	0	0	0	0	4	1	1
OPT	GY	26	26.92	0.21	0.34	0.20	0.14	20.21	56.61	2	4	0	0	4	5	0	0	0	0	0	0	0	0	1	3	0	1	4	1	1
LWD	HD	27	37.04	0.61	0.14	0.24	0.12	8.43	16.14	2	2	0	2	0	2	2	1	1	0	2	0	2	2	2	0	0	0	4	1	2
LWD	PH	26	73.08	2.99	0.92	0.25	0.14	28.31	95.51	2	0	1	5	2	0	2	1	0	0	0	4	2	4	0	0	0	1	0	1	1
LWD	SA	24	62.50	2.18	0.24	0.27	0.13	10.32	15.54	3	0	0	2	3	1	0	1	0	6	1	0	3	0	0	0	1	0	2	1	0
LWD	GPS	27	55.56	2.59	0.26	0.24	0.10	11.31	12.86	2	1	3	4	0	2	0	3	1	1	0	4	1	1	0	0	1	0	2	0	1
LWD	TKW	26	69.23	2.45	0.27	0.24	0.13	15.48	34.79	1	4	0	0	2	1	1	2	0	0	1	0	5	2	0	0	1	0	4	1	1
LWD	GY	24	29.17	0.17	0.25	0.22	0.15	6.50	8.56	1	2	0	4	1	2	1	0	0	1	0	3	1	0	1	2	0	2	2	0	1
MWD	HD	28	50.00	0.66	0.13	0.23	0.11	14.30	25.45	3	3	0	2	0	1	2	2	1	1	2	1	1	0	1	1	3	0	2	0	2
MWD	PH	18	44.44	2.34	0.23	0.25	0.11	8.66	10.37	2	1	1	2	1	0	1	1	0	0	0	2	1	2	0	0	1	0	0	3	0
MWD	SA	15	73.33	2.10	0.19	0.18	0.14	23.48	33.27	3	0	0	1	3	0	0	1	0	2	0	0	2	0	0	1	0	0	1	1	0
MWD	GPS	18	77.78	2.93	0.21	0.27	0.13	35.41	71.60	2	2	1	3	1	2	1	2	1	1	0	0	0	0	0	0	0	0	1	0	1
MWD	TKW	34	58.82	2.37	0.35	0.23	0.15	32.63	77.44	1	3	0	0	5	7	1	3	0	3	1	0	1	1	1	0	0	1	4	2	0
MWD	GY	31	29.03	0.21	0.37	0.28	0.14	7.52	9.90	2	0	0	4	0	3	1	7	0	3	0	1	3	1	0	0	1	1	3	1	0
HWD	HD	26	30.77	0.74	0.25	0.25	0.13	11.50	24.57	2	5	0	3	0	4	0	1	0	1	1	0	2	0	0	1	0	0	2	4	0
HWD	PH	28	60.71	2.63	0.49	0.24	0.15	8.81	17.04	2	0	0	3	0	3	0	0	1	0	3	3	2	3	0	1	2	0	0	4	1
HWD	SA	10	10.00	4.41	0.26	0.22	0.16	8.21	9.13	1	0	0	0	0	0	0	3	0	0	0	0	0	1	0	3	0	0	1	0	1
HWD	TKW	18	55.56	3.23	0.35	0.24	0.12	17.31	38.59	0	1	0	0	3	0	1	0	0	1	1	0	0	2	1	1	3	0	3	0	1
HWD	GY	28	25.00	3.26	0.50	0.18	0.13	11.75	24.03	0	2	0	1	3	1	0	3	0	1	1	1	0	1	0	3	2	1	2	4	2
HWD	GPS	23	52.17	0.37	0.43	0.21	0.15	21.33	65.07	1	1	0	0	6	0	1	2	0	1	1	0	2	1	1	1	2	0	0	3	0

Effect expressed in percentage of trait mean for the most associated SNP. % Pos.: Percentage of QTLs for which the most frequent allele had a positive effect on trait. MAF = minor allele frequency.

close to a genetic plateau (Calderini and Slafer, 1998; Cassman, 1999; Graybosch and Peterson, 2010). This assumption is not supported here, as modern varieties appeared better adapted than older ones. We found a positive trend of selection on grain yield in all environment types, except for LWD. The best improvement was found in the optimal water condition (OPT) with an increase of $38 \text{ kg}\cdot\text{ha}^{-1}\cdot\text{y}^{-1}$. The slope of the genetic gain seemed to be in line with studies on European winter wheat conducted previously (Brisson et al., 2010; Cormier et al., 2013). An increase of the grain number per spike (GPS) in the OPT cluster was also noticed. This observation is in agreement with the study of Brancourt-Hulmel et al., (2003), which focused on genetic progress in grain yield, by comparing 14 French winter wheat lines registered between 1946 and 1992 and concluded that grain yield improvement was mainly due to an increase in kernel number.

To our knowledge, no study has yet evaluated progress in breeding for water limited environments in a European context. However, the International Maize and Wheat Improvement Center (CIMMYT) breeds varieties adapted to marginal environments, and in this framework Lantican et al., (2001); Trethowan et al., (2002) and Manès et al., (2012) described the progress of wheat yields under low yielding and drought prone environments. These studies reported that the rate of yield increase was between 2.1 and $87.7 \text{ kg}\cdot\text{ha}^{-1}\cdot\text{y}^{-1}$. All these studies agreed that the rate of yield increase was lower in low-yielding environments than in high-yielding environments. We reached similar conclusions in the present study. Grain yield improvement was best in the OPT cluster which seems logical as French breeders and most of the wheat production areas are in the northern France, where optimal water conditions are the most frequent. Furthermore, genetic progress was limited in the intermediate water stress pattern. The gap was even more obvious in the LWD cluster, where yield improvement was not significantly different from zero. If breeders wanted to fill this gap, they would have to specifically breed for this scenario, for example by introducing QTLs specific to these particular environments. It is also interesting to note that in the HWD cluster, the slope is close to the value of the slope for the OPT cluster. This result was not expected, different assumptions might be made, for example, it may be explained by examining the positive correlation between the QTLs effect on grain yield in these two clusters.

4.3 GWAS

GWAS was carried out with six traits in the four different ETs and we found 590 QTLs. As expected, GWAS revealed a complex genetic basis and a considerable environmental influence on QTL effects. This observation was similar to other multi-site experiments (Tuberosa, 2002; Malosetti et al., 2008; Bonneau et al., 2013)). None of the QTLs had a significant and consistent effect on all the environment types.

For this study, we chose a $-\log_{10}(\text{P-value})$ threshold of 3, which may seem low regarding the number of tests performed for each trait. This choice should account for the false positive rate (Type I error).

Table III.5: Number of common QTLs between two traits

Numbers of common QTLs with opposite effects on traits are located in the lower diagonal, same sign effects are in the upper diagonal, and the ratio between the total number of QTL which colocalize with at least another QTL and the total number of QTL is on the diagonal. (e.g. 10 QTLs are specific to HD in the OPT cluster). GY: grain yield, SA: spike per area, GPS: grain per spike, TKW: thousand kernel weights, HD: heading date, PH: plant height. ETs related to water stress pattern, OPT: Optimal condition (blue); LWD: Late Water Deficit (green), MWD: Medium Water Deficit (orange), HWD: High Water Deficit (red).

		OPT					LWD					MWD					HWD									
		HD	PH	SA	GPS	TKW	GY	HD	PH	SA	GPS	TKW	GY	HD	PH	SA	GPS	TKW	GY	HD	PH	SA	GPS	TKW	GY	
OPT	HD	20/30	1		4		3	8						1	8	2		1	2	2	7	2				1
	PH	-1	18/24		2	1	1	3	7	3					3	2			2	1	2	9				
	SA	-3		12/25	1	3	2			2	1	1	1			1	1					1	1			2
	GPS	-2		-2	15/21	1	3	3	1		4	1			2			1	3	2	1	2	1	2	1	2
	TKW				-2	14/22	1		1	3	3	4			1				5			1			2	2
	GY	-1			-2		18/25	4		1	1	1	2		3				2	2	5	1	1	1		7
LWD	HD	-2	-1	-1	-2		-1	25/27	1		1	1	1	15	1	1		2	2	8	3				1	1
	PH		-2				-1		15/25		2	1		1	3	1				1	4					1
	SA	-1	-2		-1	-1	-1		-1	17/24	2	1				4	1	1	2			1	1			2
	GPS		-1				-1	-1	-4		19/27	1	4	2			3	1	1			1				
	TKW	-2		-1	-2						-1	-2	16/24	1				8			1	1		2	2	2
	GY	-2	-6		-1		-1	-4	-4		-1	-1	-1	18/24			1	1	7		1		2	1	1	2
MWD	HD	-1			-1			-1		-2	-1		-3	21/27		1		1	1	7	1					1
	PH		-1	-2			-1		-1	-1	-1		-2		10/18					2	4					
	SA	-1					-1			-2		-1	-2	-1		11/13						2	1			3
	GPS			-2		-2	-1				-2		-2				12/15									
	TKW	-4	-1		-2		-2	-1	-1	-1	-1	-2	-2	-1	-1	-1		21/29	3	1	2		2	2		1
	GY	-8	-4		-2		-3	-11	-3		-2	-1	-2	-11	-2	-1	-2	-1	25/31	2	2	1				4
HWD	HD	-2	-1		-2		-1	-2	-3	-2		-3	-3	-1		-1		-4	-8	17/24	2					
	PH	-1	-2		-1	-1		-2	-3	-1	-2		-4	-1				-1	-3	-1	19/27					
	SA										-1					-1	-1					8/10				4
	GPS			-1			-1	-1	-1		-1										-1	-1	11/22			1
	TKW	-1			-1		-1	-2	-1		-1			-2		-1	-1	-1		-1				11/17		
	GY	-1					-1	-1	-1	-1	-1				-1	-1		-2		-2	-3		-1	-3		20/27

Applying the Bonferroni correction to our study raised the threshold to 5.84 with a Type I error at 0.05. But this correction assumed that the tests were independent, which is generally not the case in GWAS due to the LD between markers. Gao et al., (2009) proposed a permutation method to calculate the effective number of independent tests, which could be applied to GWAS. With this method, the threshold decreased to 4.99 with the Type I error at 0.05. If we applied this threshold to our study the Type II error would increase dramatically and we would not be able to detect QTLs with small effects. Cormier et al., (2014) applied a detection power with a similar panel (80% of the varieties are common to our panel) and found that increasing the P-value threshold higher than 3 considerably reduced the power of detection, even more if the SNP explained less than 10% of the variance. These minor QTLs are not really crucial for breeders because they imply a considerable breeding effort for a small gain, but in future works they could be useful for meta-QTL analyses on European material in a context of abiotic stresses. Moreover, choosing this threshold helped us to build a colocalization network that could prove useful for understanding the genetic correlation between traits within and outside ETs (Butler et al., 2005; Mathews et al., 2006) and above all the colocalization between flowering and grain yield QTLs.

Several major genes controlling flowering date and plant height were detected. The reduced height gene *Rht-D1*, on chromosome 4D, colocalized with QTL of PH in each cluster. *Rht* genes have been largely associated with the “green revolution”. Shorter wheats will increase grain yield at the expense of straw biomass and are more resistant to damage by wind and rain (Peng et al., 1999). The effect of the minor allele of this QTL increased plant height by 3.4±1.1cm and the fraction of total genetic variance explained by this QTL averaged 17±9.4% in the four ETs. Within this QTL, we found a negative association with grain yield in the LWD and MWD cluster but not in OPT and HWD, meaning that *Rht-D1* could have a significant negative effect on grain yield depending on the ETs.

Regarding major genes for precocity, the photoperiod sensitivity gene *Ppd-D1* on chromosome 2D colocalized with a QTL of heading date in each cluster. With a minor allele frequency of 0.5, *Ppd-D1* was totally balanced in our panel. The fraction of total genetic variance explained by this QTL averaged 24 ±8.3% for heading date in the four clusters. *Ppd-D1* also colocalized with grain yield in three clusters (OPT, LWD and MWD), but the effect was different depending on the cluster considered. Early genotypes had a lower grain yield in the OPT cluster and the opposite effect was observed in LWD and MWD clusters. In cluster OPT, late genotypes had more grains per spike, while in MWD the thousand-kernel weight was higher in the early genotypes than in the late ones. As expected, fine tuning flowering time could be useful to escape drought stress (Fischer and Maurer, 1978; Bennett et al., 2012). No colocalization was found with vernalization genes, because our panel contained exclusively winter wheat varieties and trials were sown in autumn.

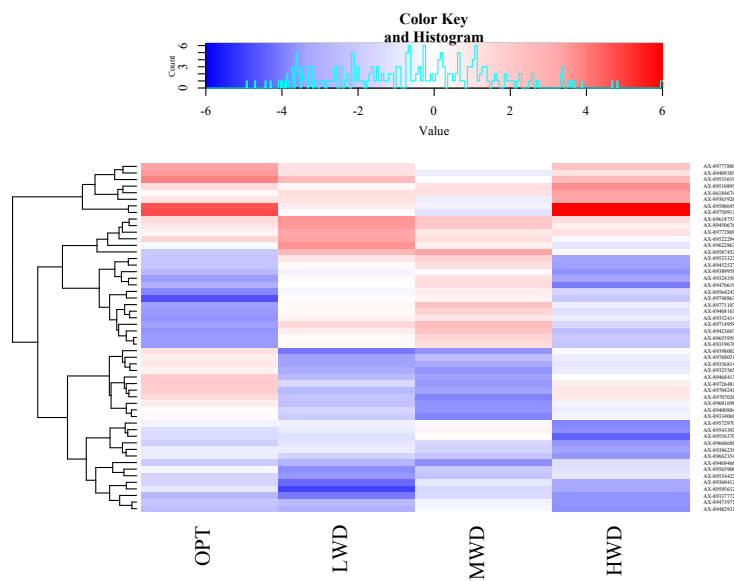


Figure III.6: Heatmap of the standardized effects of the 53 QTLs of grain yield in the four ETs

The color indicates the sign of the effect of the minor allele for positive effects in red, no effect in white, and negative effects in blue. The SNPs were selected based on the absence of co-location with heading date and plant height. A hierarchical clustering based on the Euclidian distance of these effects was used to sort the SNPs. ETs related to water stress pattern, OPT: Optimal condition (blue); LWD: Late Water Deficit (green), MWD: Medium Water Deficit (orange), HWD: High Water Deficit (red).

As we worked on a panel composed of elite varieties, the method used to select QTL candidates to improve grain yield progress should be discussed. Anthesis corresponds to a physiological transition which has a major effect on GxE interaction (Langer et al., 2014). Varieties were tested in a wide range of environments, in particular those with terminal stresses affecting the late varieties more than the early ones. This created a confounding effect of major phenology genes (M. Reynolds et al., 2009) which are more likely to be associated with agronomic traits. We also chose not to take into account the grain yield QTLs that colocalized with plant height and have a contrary effect, since plant height is easily controlled by breeders through major genes and because increasing straw biomass often penalizes grain yield. Thus 53 QTLs not linked to heading date and plant height, of which 26 were specific to one ETs, were identified (see the heatmap of their standardized effects in all the ETs [Figure III.6](#)). We could see three main QTL patterns: QTLs with a positive effect in all the clusters, QTLs with a negative effect in all the clusters, and QTLs with contrasting effects. As the effect taken into account is that of the minor allele, it seems pertinent for European breeders to increase the frequency of the first group, *i.e.* QTLs with a positive effect in all clusters, and to reduce the frequency of the QTLs with a negative effect. The other QTLs are more or less specific to one cluster, so they could be useful depending on the target of the breeders.

Nevertheless, we could highlight QTLs interesting for breeding. The first is located in the telomeric region of the short arm of chromosome 2B. Within this area two SNPs were associated with grain yield in OPT and HWD, and the part of the variance explained by these QTLs are around 10 and 15%, respectively. As we said previously, this area was not associated with heading date and plant height, but with the number of grains per spike and TKW. In addition, the minor allele (frequency around 10%) contributes the positive effect, meaning that increasing the frequency of the minor allele in the French elite germplasm could help to improve grain yield in optimal and high water stress conditions.

On the short arm of chromosome 4A, another QTL could be highlighted for grain yield in the LWD and MWD clusters, which explained 7 and 6% of the genotypic variance. Although the minor allele contributed the negative effect, this QTL was quite balanced in our panel (MAF = 0.47) and equally spread between recent and old varieties. However, an opposite effect was observed in the OPT cluster, and although this effect was not significant, we could expect that this QTL improves grain yield in water stressed environments and decreases it in optimal environments. This kind of QTL could be useful for breeding only if the target is a group of environments with a high frequency of drought events. Furthermore, Bouffier (2014) reported that in the literature, chromosome 4A was one of the chromosomes most frequently linked to yield under irrigated, drought and heat conditions.

The genetic basis of drought resistance has been screened extensively in spring wheat germplasm from CIMMYT and from Australia (Pinto et al., 2010; Pinto and Reynolds, 2015; Tricker et al., 2018), but such studies are not common in elite European germplasms.

The use of QTLs identified in this germplasm could enhance selection for drought resistance in wheat. Furthermore, examining the pleiotropic effect between grain yield and its components could help breeders to identify ideotypes specific to the agricultural practices and climatic conditions of their market.

CONCLUSION

In a context of climate change, we investigated drought stress resistance using a European panel of elite winter wheat varieties. Using a crop model-based water stress index we identified four drought scenarios and reported differential genetic progress as a function of the water scenario, demonstrating the higher yield potential of recent varieties. We also reported the genetic determinism of yield and its components within the different water stress scenarios. Among the 590 QTLs identified, we highlighted several QTLs that were linked to one or several of the drought stress scenarios. By selecting some of these QTLs, it may be possible to increase genetic progress in water stress conditions. However, we focused only on drought stress although other stresses can also impact grain yield. New environment types, addressing a larger variety of environmental factors should be developed in order to characterize more diverse conditions.

Chapitre IV : Tolérance aux stress thermiques

Chapitre IV : Genome-wide identification of QTLs affecting terminal heat stress responses in wheat (*Triticum aestivum* L.)

Gaëtan Touzy^{1,2} – Jacques Le Gouis³ – Vincent Lievin² – Xavier Decoopman⁴ – Elise Redondo² – Sébastien Praud² – Stéphane Lafarge²

¹ Arvalis-Institut du végétal, Biopole Clermont Limagne, 63360 Saint-Beauzire, France.

² Biogemma, Centre de Recherche de Chappes, Route d'Ennezat CS90216, 63720 Chappes, France.

³ INRA, UCA UMR 1095, GDEC, Clermont-Ferrand, France.

*: corresponding author: stephane.lafarge@biogemma.com

Abstract

Heat stress is critical abiotic stress, especially for winter wheat for bread-making (*Triticum aestivum* L.) at the flowering grain filling stage, by limiting its growth and productivity around the world. In a climate change context, the acquisition of new high-yield and stress-tolerant varieties is necessary and requires improved understanding of the physiological and genetic bases of heat tolerance. This study was conducted to identify genomic areas associated with a physiological trait and morphologic grain characteristics under heat stress treatment.

A panel of 199 elite European varieties was evaluated under controlled conditions in 2016 and 2017. A split-plot design was used to test the effects of high temperature for ten days after the flowering date. A genome-wide association study (GWAS) was carried out to test the main effect of SNP and the effect of SNP by treatment interaction, using genotyping data from a 280K SNP chip. Two different classes of traits were measured: physiological traits at flowering time, the number of productive spikes per plant, stay green traits and morphologic grain traits on the main shoot and tillers. Strong correlations were observed in the same class of traits. Genotypes by treatment interaction were only observed for morphologic traits in grain in both main shoots and tillers. The GWAS identified several quantitative traits loci (QTLs) associated with the main effect and interaction with the treatment.

1 INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most important crops cultivated in the world. From the late '60s, a steady growth trend for wheat yields has been observed in many areas (Calderini and Slafer, 1998). However, a decline in this trend has been observed in several countries over the last two decades (e.g. Peltonen-Sainio et al., 2009; Brisson et al., 2010). In France, this has mainly been explained by climate variability, such as drought stress during booting and high temperatures during grain filling (Brisson et al., 2010). Lobell et al., (2011) succeeded in showing that climatic variation such as precipitations and heat stress may have led to a reduction of 5.5% in grain yield since 1980, which corresponds to a loss of 35M tonnes on a world scale (equivalent to the French annual wheat harvest). In the context of climate change, temperatures are expected to increase across many regions in the world. By the end of the 21st century, the average global temperature is predicted to increase by 0.3°C-1.7°C in the lowest greenhouse gas emission scenario, and 2.6°C-4.8°C in the highest one (Stocker et al., 2013). Combining climate predictions with a wheat simulation model, Semenov and Shewry, (2011) showed that not drought, but the frequency and magnitude of heat stress around flowering time and during grain filling could significantly impact wheat yield. This may be especially the case for wheat varieties commonly grown in Europe as they were not specifically selected for their tolerance to heat stress. For this reason, for improving our understanding of the physiological and genetic bases of the response of European wheat varieties to heat stress, it would seem important to propose efficient tools to facilitate breeding in this context.

Heat stress can be characterized by three different parameters for defining a scenario: (1) timing, *i.e.* at what stage the plant was affected by stress; (2) the duration of stress; (3) the intensity, *i.e.* to what levels of temperature the plant has been exposed. Moreover, two kinds of stress are associated with yield loss: “heat shocks” which are often short stress periods (a couple of hours) but with very high temperatures (>40°C) and “chronic heat stresses” where average daily temperatures are higher than the optimum during a longer period. A large range of crop responses to high temperature has been reported in literature. These responses depend on the stress scenario, the different germplasms (*e.g.* unsuitable European material vs suitable Australian material) and the kind of experiment: controlled conditions or field trials. Porter and Gawith, (1999) reviewed more than 65 papers to describe the response of wheat to extreme temperatures and provided optimal temperatures for the different stages. They determined that 20.7°C (± 1.4) was the optimal temperature at the grain filling stage and 35.4°C (± 2.0) was the maximum temperature beyond which growth stops. To evaluate the impact of heat stress on plants, numerous studies count the number of days during which the maximum temperature exceeds 25°C at different growth stages (*e.g.* Sofield et al., 1977; Hunt et al., 1991; Gate, 1995; Lecomte, 2005). Gate et al., (2010) applied this threshold to a French multi-environment trial network and showed that for each day with a maximum temperature higher than 25°C, from heading to heading +750-degree days (°Cd), the yield loss may be up to 0.14t/ha.

Heat stress affects plant development and yield through numerous mechanisms. Firstly, higher temperatures during the crop cycle induce an acceleration in biological processes (Johnson and Thornley, 1984) and, as a result, an earlier occurrence of phenological stages can be observed (Porter and Gawith, 1999) then it limit the biomass accumulation. Despite these earlier translated growth stages, wheat in Europe was still impacted by several heat stress events during grain filling (Gate, 2007; Gate et al., 2008). The impacted physiological mechanisms could be split into three types, which were reviewed by Cossani and Reynolds, (2012): (1) light interception with a reduction of the leaf area index and green area duration. *E.g.* Xu et al. (1995) showed that a reduction of leaf chlorophyll content during a post-anthesis heat stress led to earlier senescence; (2) radiation use efficiency (RUE) by altering proteins structure and then reducing the intensity of photosynthesis and by decreasing CO₂ solubility and then increasing photorespiration; (3) the partitioning of total assimilates. If the stress is around anthesis the number of grains per spike could be impacted by inducing pollen sterility or grain abortion (Tashiro and Wardlaw, 1990), and if the stress is after flowering the water-soluble carbohydrates remobilization and starch synthesis and consequently grain weight could be impacted (P. J. Stone and Nicolas, 1994; R. Wheeler et al., 1996; Calderini et al., 1999). High temperature during cell division in grain formation could also reduce the final wheat grain (Nicolas et al., 1984; Commuri and Jones, 2001; Barnabás et al., 2007).

Senescence is a normal process for destroying chlorophyll and remobilizing nutriment from the source to the target organ (Taiz and Zeiger, 2002). In monocarpic species, later initiation or/and a decrease in the senescence rate compared to a standard reference is defined as the “stay green” ability. Stay-green plants may be able to maintain photosynthesis during grain filling (Thomas and Howarth, 2000). Different techniques were used to monitor senescence and extract the main parameters (*e.g.* start, duration, rate and end) defining the stay-green trait, such as visual scoring (Lim et al., 2007a), measurement of the chlorophyll content with a SPAD meter (Borrell et al., 2000; Christopher et al., 2008) and an estimation of the normalized difference vegetative index (NDVI) (Lopes and Reynolds, 2012; Bouffier et al., 2015; Christopher et al., 2016)). Under heat and drought stresses, stay-green was interpreted as a way to quantify tolerance (Olivares-Villegas et al., (2007); Pinto et al., (2010); Bouffier, (2014)).

Improving new wheat varieties could be facilitated by deciphering the genetic bases of heat tolerance. Several quantitative genetic studies were conducted to identify quantitative trait loci (QTL), associated with different agronomical and physiological traits in heat stress conditions. Reviews of these studies could be found in Bouffier (2014), Farooq et al. (2014), Tricker et al. (2018). Many of them were performed with bi-parental spring wheat populations under arid environments. It seems that no study has been conducted to dissect genetic determinants of heat tolerance with a European winter wheat panel. Whilst the introgression of QTLs from European germplasm is more appropriate to a European

breeding program, since the genetic background is closer to their material, the risk of introducing negative QTLs at the same time is lower.

However, evaluating the impact of heat stress in field trials without any drought stress could be challenging. Numerous field studies for QTL detection were conducted by the CIMMYT in Mexico (Pinto et al., (2010); Bouffier et al. (2015); Pinto et al., (2016)). They used two different sowing dates, an optimal date and a late date. Both experiments were well irrigated, so delayed sowing exposed plants to terminal heat stress due to high summer temperatures, but they did not suffer from any drought stress. Compared to the CIMMYT facilities, performing field trials for screening varieties for heat tolerance is more complicated in Europe. Indeed, even if late sowing may increase the likelihood of heat stress, Europe's temperate climate is too unstable to risk these kinds of experiment. Moreover, testing winter panels that are composed of both photoperiod-sensitive and insensitive genotypes may bias the results. Both cold and photoperiod requirements may cause varieties to flower at different times and then undergo different stresses, with a possible confusion between escape and tolerance. Most of the studies quantifying the impact of heat stress alone have used controlled environments (P. Stone and Nicolas, 1994; Gibson and Paulsen, 1999; Spiertz et al., 2006; Shirdelmoghanloo et al., 2016).

The objective of this study was to quantify the effects of high temperatures during grain filling on elite European winter wheat. Our study was conducted in controlled conditions and each plant was exposed to a heat treatment after it had flowered, without any acclimation period. Flowering time, leaf chlorophyll content, the number of productive spikes, grain number, grain weight and grain size were measured. A GWAS was then performed to identify genomic areas associated with these traits and their response to heat treatment. We identified two main QTL associated with thousand kernel weight tolerance to heat. We expect this work to contribute to future progress in breeding for adapting to climate change.

2 MATERIALS AND METHODS

2.1 Plant material and growth conditions

One-hundred and ninety-nine elite European winter wheat varieties, registered in Europe from 1974 to 2010, were assembled in a diversity panel ([Suppl. IV.1](#)). These varieties were evaluated in the greenhouse during the spring in 2016 and 2017 at Biogemma's site (Chappes, France).

Plants were sown on 02-22-2016 and 03-02--2017 in 1cm deep trays containing commercial potting TS 3, from Klassman, France. The seedlings were raised in the greenhouse maintained at 21°C (daytime maximum) / 15°C (night-time minimum). After two weeks, seedlings were vernalized for 49 to 56 days at 8°C with an 8h photoperiod. Then, plants were transplanted into 1L pots (12 cm long and 11 cm wide). The rooting medium in the pots was commercial "Fleurissement 384" from Klassman,

France, plus a controlled-release fertilizer “Nutricote type 70” from Fertil, England, with 13%N, 13%P₂O₅ and 13%K₂O at 0.48g per pot and with minor nutriments added before transplanting. Pots were watered daily and kept in trays containing water about 1cm deep throughout the experiment to avoid water stress. Bio-control and chemical pesticides were used to control diseases and pests. The greenhouse was maintained at 21/15°C and relative humidity was set at 60%. Air temperature and relative humidity were continuously monitored throughout the experiment at 15-min intervals using a temperature and humidity probe.

The photoperiod was 16 h and sodium-vapor lamps (adjusted to be 100cm away from the plant canopy) provided a photosynthetic complement if natural luminosity was under 90W/m². The transition period between the daytime maximum and night-time minimum temperature was 2h. Daytime maximum temperature was held for 12h from 08:00 am to 08:00 pm. Similarly, night-time minimum temperature was held for 8h from 10:00 pm to 06:00 am.

2.2 Split-plot design

A split-plot design was used. The first effect (whole-plot) was two temperature treatments: non-stress (NS) and stress (S) temperature regimes. The second effect (sub-plot) was to use different varieties from the panel. Each treatment × variety combination was replicated three times. Pots were spread out on 15 tables (100 x 200cm) organized in a 3 by 5 matrix ([Suppl. IV.2](#)). One hundred and twenty pots were arranged on each table, in an 8 by 15 matrix, which resulted in a density of 60 plants per m². Spring wheat varieties were positioned all around the tables to minimize border effects. Knowing the greenhouse temperature and luminosity gradient, the three replicates were arranged in an East/West direction, *i.e.* replicate 1 on tables 1, 4, 7, 10 and 13; replicate 2 on tables 2, 5, 8, 11 and 14; and replicate 3 on tables 3, 6, 9, 12 and 15. Each treatment was arranged on two and a half tables, *e.g.* within the replicate 1 the tables 1, 4 and half of table 7 were assigned to one treatment, while tables 10, 13 and the other half of table 7 were assigned to another treatment ([Suppl. IV.2](#)). The tables were split into two, defining 5 sub-blocks per whole-plot, where 3 checks were replicated. Temperature treatment location was switched between 2016 and 2017 experiments and different randomization was drawn up for genotypes and check locations.

Forty thermo-buttons from the Progres Plus company, France were used for monitoring temperature every half hour throughout the growth cycle, there were two buttons per table. The *krige* function from the package “gstat” (Pebesma, 2004) was used to perform an ordinary kriging operation and assign a mean temperature value for each pot.

2.3 Heat treatment

Three days after anthesis (DAA), plants targeted to be stressed, were put into another greenhouse for ten days, with the same photoperiod, but with a temperature regime of 29/23°C. During this time, the

original plan was not respected, but all the plants with the same flowering date were together on the same table. After heat treatment, the plants were returned to their former location.

The quality of temperature control was assessed by thermo-buttons. The average daytime temperatures (between 07:00 and 21:00 hours) and night-time temperatures (between 21:00 and 07:00 hours) in the 2016 experiment were for S ($30.4^{\circ}\text{C}\pm 2.9$ and $24.4^{\circ}\text{C}\pm 2.6$), and for NS ($20.0^{\circ}\text{C}\pm 2.6$ and $16.4^{\circ}\text{C}\pm 2.9$); and in the 2017 experiment were for S ($29.0^{\circ}\text{C}\pm 3.0$ and $23.8^{\circ}\text{C}\pm 3.5$), and for NS ($21.7^{\circ}\text{C}\pm 1.7$ and $17.6^{\circ}\text{C}\pm 2.3$). On average, thermal time was $27.8^{\circ}\text{C}\cdot\text{day}^{-1}$ and $27.1^{\circ}\text{C}\cdot\text{day}^{-1}$ in S treatment and $19^{\circ}\text{C}\cdot\text{day}^{-1}$ and $20.6^{\circ}\text{C}\cdot\text{day}^{-1}$ in NS treatment, for 2016 and 2017 respectively. Relative humidity during daytime and night-time was $49.16\%\pm 9.2$ and $61.0\%\pm 8.4$ for S treatment and $74.0\%\pm 10.3$ and $80.2\%\pm 8.3$ for NS treatment in 2016. It was $60.0\%\pm 9.6$ and $70.3\%\pm 9.0$ for S treatment and $77.14\%\pm 8.0$ and $82.8\%\pm 6.3$ for NS treatment in 2017. By combining temperature and humidity, we have calculated the vapour pressure deficit VPD (VPD). VPD was slightly higher in 2016 than in 2017. VPD during daytime and night-time was at $2.13\text{kPa}\pm 0.38$ and $1.86\text{kPa}\pm 0.23$ for S treatment and at $1.71\text{kPa}\pm 0.21$ and $1.48\text{kPa}\pm 0.15$ for NS treatment in 2016. It was $2.40\text{kPa}\pm 0.37$ and $2.07\text{kPa}\pm 0.32$ for S treatment and $1.99\text{kPa}\pm 0.16$ and $1.67\text{kPa}\pm 0.16$ for NS treatment in 2017. Therefore, there was a difference in temperature and VPD between stressed and non-stressed conditions, regardless of the year of the experiment.

2.4 Measurement of yield components

For each pot, flowering time (D.Z65) corresponding to the Zadok stage Z.65 (Zadoks et al., 1974) was read daily on the main shoot. At the end of the experiments, the number of fertile spikes per plant (SPP) was counted, the main shoot was separated from the others and grains were harvested using a small plot harvester. The grains were dried for one week at 35°C to homogenize the moisture content of the grain before beating/counting. Using a MARVIN digital seed analyser (www.gta-sensorik.com), the number, width, length and area of dry grains were determined for both the main shoot and the tillers. The grains were weighed (GW) and then the thousand kernel weights (TKW) calculated.

2.5 Measurement of chlorophyll content and estimation of senescence traits

A DUALEX SCIENTIFIC™ (www.force-a.com) sensor was used to measure chlorophyll (CHL) content ($\mu\text{g}\cdot\text{cm}^{-3}$) from light transmission as follows:
$$\text{CHL} = \frac{\text{Near.infrared transmission} - \text{Red transmission}}{\text{Red transmission}}$$
 CHL was measured weekly at a minimum of five different dates for each pot, starting from flowering time until maturity. Measurements were taken at five different points along the flag leaf blade.

Inspired by Christopher et al. (2014) wheat senescence was modelled for each plant using the following logistic function:

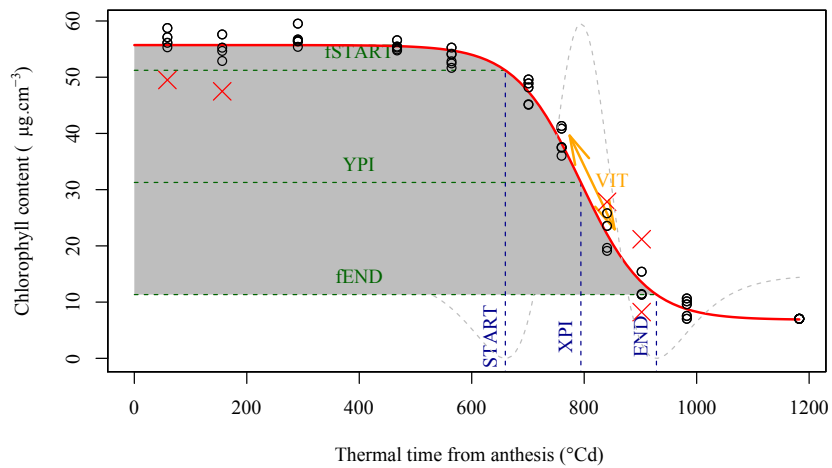


Figure IV.1: Example of the fitted logistic curve and the estimation of stay-green traits

The red curve is an example of the fitted logistic curve for plant 455 in the 2016 experiment; the black open dots represent the raw data; the red crosses represent data identified as outliers; green and blue segments with dashes are respectively the y-axis and x-axis coordinates of the start, the point of inflection and the end of senescence; the grey curve with dashes is the third derivative of the function which, as a minimum, defines the start and end of senescence and which, as a maximum, defines the x-axis of the point of inflection; the orange arrow is an illustration of the maximal senescence rate, the grey area represents the area under the senescence curve (AUC)

$$[1] \quad CHL = D + \frac{K}{1+a \times \exp(r \times t)}$$

Where CHL is a vector of the chlorophyll content ($\mu\text{g}\cdot\text{cm}^{-3}$) measurements from the flowering date to maturity (at least 5 dates and 5 measurements per flag leaf blade), D is the final CHL of the dead plant, K corresponds to the difference in CHL between the maximum and final values, a and r are an indicator of the senescence rate and t is the thermal time from plant anthesis. To enable proper data fits, an extra CHL value corresponding to the fully senesced plant, was added $200^\circ\text{C}\cdot\text{days}^{-1}$ after the last measurement. Fitted logistic functions were obtained using the *nls* function (Bates and Watts, 1988; Bates and Chambers, 1992) from package *stat* (R Development Core Team, 2011). A first run was performed on raw data for identifying outliers. To do this, residuals distribution was compared to normal distribution and outliers were identified using method I of the *getOutliers* function (Van Der Loo, 2010) from the *extremevalues* package (R Development Core Team, 2011). A second run was then performed without the outliers for estimating the different parameters of the logistic model. Pseudo R^2 were calculated for evaluating the strength of fitted models as per the following Efron (1978) formula:

$$[2] \quad R^2 = 1 - \frac{\sum_{i=1}^N (CHL_i^2 - \overline{CHL})^2}{\sum_{i=1}^N [(CHL_i - \overline{CHL})^2]}$$

Where N is the number of observations in the model, CHL is a vector of the chlorophyll content ($\mu\text{g}\cdot\text{cm}^{-3}$) measurements from flowering to maturity, \overline{CHL} is the mean of the CHL values, and \widehat{CHL} is the value predicted by the model.

The logistic function in Eq[1] provided a close fit to the experimental data (e.g. [Figure IV.1](#)). The model converged on more than 95% of the data in both experiments. A second data cleaning operation was performed by hand to ensure that the adjustment of the kinetics corresponds correctly to the senescence process. Removal of outliers was based on senescence parameters distributions, visual curve fitness and R^2 . All in all, it resulted in 199 plants (7.8%) with no senescence data. The mean of R^2 for the remaining plants was 0.94 (± 0.05) and 0.98 (± 0.02) for the 2016 and 2017 experiments, respectively. All the senescence parameters presented normal distribution characteristics except for a .

Additional senescence traits were obtained using the properties of the logistic function. The x-axis and y-axis of the point of inflection were obtained as $XPI = \frac{|\ln(a)|}{r}$ and $YPI = \frac{K}{2}$. The maximal rate of senescence was obtained with the first derivation as $VIT = f'(XPI)$. The beginning (START) and the end (END) of senescence were determined as the thermal time when the third derivative of the function attained its minimums. μ The chlorophyll content at these two points was noted as fSTART and fEND. We calculated the area under the senescence curve from anthesis until the completion of senescence (AUCK). It was calculated as the integral from anthesis to end of senescence for the x-axis, and from K to 0 for the y-axis. AUCK is related to the start, the speed of senescence and the CHL

maximum. We also calculated the area under the curve without the CHL maximal, named AUC. AUC was represented on the [Figure VI.1](#) as the grey area.

2.6 Statistical analysis

The same experimental designs were used in both experiments (2016 and 2017), we decided then to use the same baseline model to estimate genetics by treatment interaction (GxT). This model was specific to the split-plot design of our study and was adapted to a multi-year analysis and SNP by GWAS treatment interaction. The models were fitted using a mixed model written in R using the *ASReml-R* package (Butler et al., 2009).

The baseline model for the GxT estimate in one experiment was written as:

$$[3] \quad Y_{ijkl} = \mu + R_j + T_k + T_k(R_j) + B_l(R_j T_k) + G_i + G_i \times T_k + \varepsilon_{ijkl}$$

Where Y_{ijkl} is the phenotypic value of genotype i ($i=1:199$) in the sub-block l of the treatment k of the replicate j , μ is the general mean trait, R_j is the fixed effect of the replicate j ($j=1:3$), T_k is the fixed effect of treatment k with $k = 1 =$ non-stress treatment (NS) and $k = 2 =$ stress treatment (S), $T_k(R_j)$ is the fixed effect of treatment k within replicate j , $B_l(R_j T_k) \sim \mathcal{N}(0, \sigma_l^2)$ is the random effect of the incomplete sub-block l ($l=1:5$) within treatment k of replicate j , $G_i \sim \mathcal{N}(0, \sigma_g^2 K)$ is the random genotype effect where K is the genomic relationship matrix calculated using VanRaden (2008) formula (see below), $G_i \times T_k$ is the random genotype-by-treatment interaction effect with $G_i \times T_k \sim \mathcal{N}(0, \sigma_{g.t}^2 \cdot K \otimes I_{nT})$ where $K \otimes I_{nT}$ is the Kronecker product between the K matrix and the I_{nT} , identity matrix and ε_{ijkl} is the residual error $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

The variance/covariance of the random polygenic effect (G_i) was modelled with a kinship (K) estimated with molecular markers. Pairwise kinship coefficients were computed according to the first method described by VanRaden (2008), with a subset of 27,680 SNP obtained by removing highly correlated SNP. In brief, we computed r^2 for each pair of SNP located on the same chromosome. A dendrogram was built with this matrix of r^2 and one SNP was randomly selected from each cluster with a distance < 0.1 (with distance = $1 - r^2$).

Narrow sense heritability (h^2) was estimated as:

$$[4] \quad h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{g.t}^2 / t + \sigma_\varepsilon^2 / tr}$$

where σ_g^2 , $\sigma_{g.t}^2$ and σ_ε^2 are genotype, genotype by treatment interaction and residual variance, and t and r are the number of treatments and replicates.

For a combined analysis across the two years, the baseline model was enhanced with a year effect as:

$$[5] \quad Y_{ijklm} = \mu + A_m + T_k + T_k \times A_m + A_m(R_j) + T_k(A_m R_j) + B_l(A_m R_j T_k) + G_i + G_i \times T_k + G_i \times A_m + \varepsilon_{ijklm}$$

where the new terms A_m and $G_i \times A_m \sim \mathcal{N}(0, \sigma_{g.a}^2 \cdot K \otimes I_{nA})$ are the effect of the m th years and their interaction with genotype, respectively. In this model, we must stress that a part of the $G_i \times T_k \times A_m$ interaction was included in the model residual, resulting in an underestimation of the specific influence of $G_i \times T_k$ and $G_i \times A_m$.

Narrow sense heritability (h^2) was estimated as:

$$[6] \quad h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{g.t}^2/t + \sigma_{g.a}^2/a + \sigma_\varepsilon^2/atr}$$

where σ_g^2 , $\sigma_{g.t}^2$, $\sigma_{g.a}^2$ and σ_ε^2 are genotype, genotype by treatment interaction, genotype by years interaction and residual variance, respectively, and t, a and r are the number of years, treatments and replicates.

To test the stress impact and relationship between traits the genomic best linear unbiased predictions (gBLUP) for each genotype and treatment were extracted from equations 3 or 5 as:

$$[7] \quad \overline{GGT}_{ik} = \hat{\mu} + \hat{G}_i + \widehat{G_i \times T_k}$$

Based on the \overline{GGT}_{ik} the relationships between traits were estimated with the Pearson correlation and principal component analysis (PCA).

The stress index was calculated as:

$$[8] \quad SI = 100 \times \overline{GGT}_{1.1} / \overline{GGT}_{2.2}$$

Where $\overline{GGT}_{2.2}$ is the average of gBLUP in no-stress treatment (NS) and $\overline{GGT}_{1.1}$ is the average of gBLUP in stress treatment (S).

Finally, the baseline model was adapted to multi-year GWAS split-plot analysis as:

$$[9] \quad Y_{ijklm} = \mu + A_m + T_k + T_k(A_m) + \alpha_i + \alpha_i \times T_k + A_m(R_j) + T_k(A_m R_j) + B_l(A_m R_j T_k) + G'_i + G'_i \times T_k + G'_i \times A_m + \varepsilon_{ijklm}$$

where the new terms $\alpha_i + \alpha_i \times T_k$ represent the allelic fixed effect of the genotype i at SNP α and interaction between marker α and treatment k , respectively. As part of the genetic effect is captured by the SNP, the genetic background is noted G'_i , it is defined in the same way as the genetic effect for Model 5.

All the varieties were genotyped with the Affymetrix Axiom 280K SNP array (Rimbert *et al.*, 2018). Only the Polymorphic High-Resolution SNP were used in this analysis. All SNP were physically

mapped on the genome reference sequence RefSeq V1.0 (IWGSC et al., 2018). Heterozygous loci were considered as missing data. Monomorphic SNP, SNP with more than 10% missing data, and SNP with a minor allele frequency (MAF) below 5% were discarded. Missing data were imputed with Beagle 4.1 (Browning and Browning, 2016). Finally, 164,198 SNP were tested with Equation 9.

To avoid proximal contamination, SNP were tested using a K matrix computed with SNP that were not located on the same chromosome as the tested SNP (Rincent et al., 2014).

Variance components of random factors were tested one by one using the likelihood ratio test (LRT) (Kendall and Stuart, 1979), based on log-likelihood (Lmax) differences between the complete and the reduced model without the test factor. A Wald test was performed on the complete model for testing the SNP main effect and SNP by treatment interaction. Following Gao et al., (2009), SNP were considered significantly associated with a trait if the $-\log_{10}(\text{Pvalue})$ exceeds a threshold of 5.26. The estimation of the part of the phenotypic variance explained by the SNP was calculated using the R^2 resulting from the regression of the $G_i \times T_k$ coefficients in function of the SNP's effect.

QTL boundaries from GWAS results were defined following the method described in Cormier et al., (2014). Briefly, SNP belonging to the same linkage disequilibrium (LD) cluster were defined as a group of quantitative trait nucleotides (QTN). Clustering was performed by averaging r^2 distances and the tree was cut at 1-“critical LD” (critical LD = 0.24 (Bresgello, 2005)). QTL boundaries were defined as the maximal and minimal map positions of the SNP within the QTN extended by the LD decay specific to the genomic area of the QTL.

3 RESULTS

3.1 Phenotypic analysis and variance decomposition

The variance components associated with the different effects extracted from Model 3 for 2016 and 2017 experiments are presented in [Suppl. VI.3](#). Significant genotypic effects were observed for all traits except for senescence parameters D and a in both years. Traits with significant genotypic effects showed highly variable heritability, ranging from 0.20 for fEND in 2017 to 0.95 for D.Z65 in 2016 and 2017.

As the heat treatment was applied to each plant three days after anthesis, as expected, no significant impact was observed on flowering time in the 2016 and 2017 experiments. However, a light effect was observed on the number of fertile spikes per plant: the SI was 3.14% (ns) and 4.85% ($P\text{-value}<0.05$) in 2016 and 2017, respectively. The number of grains on the main shoot was not affected, the stress index was 7.15 ($0.05<P\text{-value}<0.01$), but there was an effect on tillers: light in 2016 with an SI of 8.26% (ns) and strong in 2017 with an SI of 13.92% ($P\text{-value}<0.01$). These results were expected due to the delay in flowering between the main shoot and the tillers.

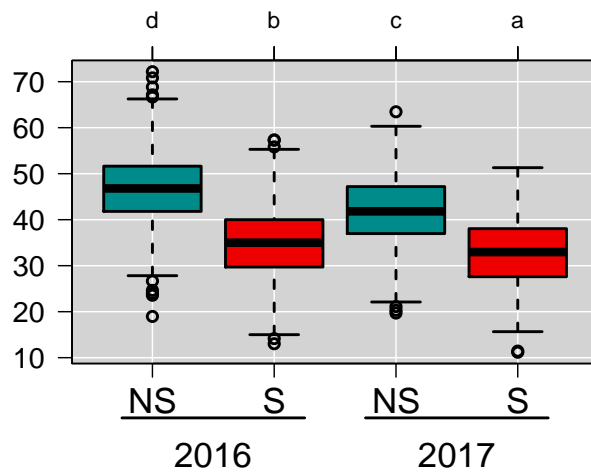


Figure IV.2: Boxplot of TKW for 199 wheat cultivars grown over two years (2016 and 2017) with two heat stress treatments

gBLUP values were calculated from Equation 2 and used to build the boxplot. Quartiles and medians are used to construct the box. Whiskers extend to 1.5 times the interquartile range from the box. Boxplot means with the same letter above are not significantly different ($\alpha=0.001$), based on Tukey HSD P-value.

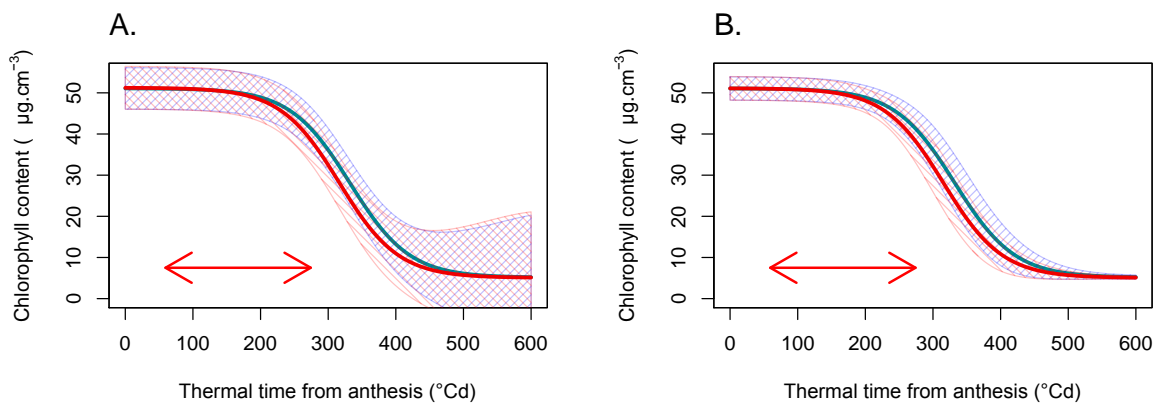


Figure IV.3: Change in senescence kinetics over time presented for all genotypes in the two heat-treatments

Red and blue curves represent the average of the fitted logistic senescence curve for all genotypes in stress and no stress treatments, respectively. Red and blue dashed areas represent the mean \pm standard deviation (Sd). (A) curve mean and Sd were calculated on the basis of raw data; (B) curve mean and Sd were calculated on the basis of gBLUP parameters. The red arrow represents the period during which the stressed plants were exposed to heat treatment.

Considering the senescence traits, only the x-axis of the point of inflection (XPI) and the end of senescence (END) in 2016 were significantly affected by the treatment with an SI of 6.89 and 6.44%, respectively. Heat stress treatment impacted the grain morphology, grain length and width, in both experiments, and for both the main shoot and the tillers. For the TKW of the main shoot, the means were 46.77g in 2016 and 41.96g in 2017 for NS treatment while they were 34.88g and 32.86g for S treatment ([Figure VI.2](#)). TKW of the main shoot was then affected by 25.73% and 21.88% in 2016 and 2017, respectively.

Variance decomposition revealed significant GxT interactions for TKW, grain area, width and length for both the main shoot and the tillers in 2016 or 2017 ([Suppl. VI.3](#)). GxT interaction was the highest for TKW and grain width on the main shoot, representing 22.45 and 25.61% of the genetic variance in 2016 and 2017, respectively. For other traits, the GxT interaction varied between the different years.

3.2 Multi-year phenotypic analysis and variance decomposition

Variance components associated with different effects extracted from Model 5, which combined the 2016 and 2017 experiments, are presented in [Table IV.1](#). Significant genotypic effects were observed for all traits except for senescence parameters D, a and r, and fEND. Traits with significant genotypic effects showed highly variable heritability ranging from 0.33 for the rated speed of senescence (VIT) to 0.88 for the flowering date (D.Z65).

No significant effect of treatment was observed for D.Z65, SPP and senescence parameters, and a light effect was observed for XPI, START and AUC. The SI for this trait were 3.75, 5.39 and 3.75%, respectively. The mean senescence curve slightly varied between treatments ([Figure IV.3](#)). The shape of these curves seems to be equivalent, but in heat-treatment senescence started earlier. A light significant effect was observed for the number of grains on the main shoot or on tillers, the SI being 7.15 and 11.24%, respectively. For other traits such as grain weight, TKW, grain area, width and length, a significant effect was observed with an SI varying from 4.16% for grain length on tillers to 28.74% for the grain weight on the main shoot.

We also can see in [Table IV.1](#) that senescence traits were the most affected in the two years, while there was only a weak effect on the flowering date.

When looking at the GxY and GxT interactions, we can see that a significant GxY interaction was found for all traits, except for senescence parameter *a*. The genetic variance of the senescence rate (VIT) was the trait with the highest GxY interaction. The grain parameters showed high, significant GxT effects.

For TKW, the variance explained by GxT and GxY was relatively equivalent for the main shoot, 4.24g² and 4.39g² and slightly lower for the tillers with 2.93g² and 3.21g², respectively. This strong

Table IV.1: Variance components from the multi-year analysis

Mean is the mean for fitted values, Sd is the standard deviation for fitted values, H² is heritability, SI is the stress index, varT is the variance explained by the treatment, varG is the variance explained by the genotype, varGxT is the variance explained by the GxT interaction, varGxY is the variance explained by the GxY interaction and varRes the residual variance.

		Units	Mean	Sd	H ²	SI	ssY	ssT	ssYT	varG	varGT	varGY	varRes
Flowering Date	D.Z65	Days	175.68	6.84	0.88	-0.06	32.17.	0.31	0.57	24.29***	0.14	5.03***	9.04
Senescence Parameters	D	µg.cm ⁻³	5.10	1.09	0.05	1.66	118.08***	0.4	4.09	0.04	0.25**	0.25*	6.27
	K	µg.cm ⁻³	46.14	3.16	0.77	-0.27	15.5	2.74	40.45	7.62***	2.07E-06	1.62***	17.44
	a		6.15E-04	4.90E-04	4.31E-03	-29.22	1.08e-04*	4.35E-06	1.99E-06	7.68E-09	1.66E-07	1.07E-07	1.95E-05
	r		2.28E-02	4.94E-03	0.18	-1.67	1.36e-03***	4.05E-06	8.36E-05	2.16E-06	2.71e-06**	5.19e-06***	6.70E-05
Senescence traits	XPI	°C.days	656.90	82.54	0.81	3.75	164647.34***	25379.47.	20941.51.	3778.31***	143.5	508.22***	6675.48
	YPI	µg.cm ⁻³	28.17	2.08	0.74	0.11	67.25***	0.03	0.46	2.09***	0.13*	0.62***	4.32
	VIT	µg.cm ⁻³ per °C.days	-0.26	0.06	0.33	-2.28	0.09***	8.18E-04	0.01	6.47e-04**	3.69e-04***	9.31e-04***	0.01
	START	°C.days	532.33	67.39	0.74	5.39	18764.63.	28689.27*	6275.82	2896.58***	221.61.	665.03***	6819.53
	fSTART	µg.cm ⁻³	47.00	3.28	0.81	-0.04	64.1**	0.03	5.3	6.58***	0.06	1.57***	9.13
	END	°C.days	781.47	107.61	0.74	2.62	350664.14***	23832.57	63877.27*	4814.29***	225.93	1050.55***	12937.70
	fEND	µg.cm ⁻³	9.33	1.16	0.14	0.81	121.98***	0.3	1.9	0.11	0.22**	0.27**	5.26
	AUC		26827.84	4175.61	0.79	3.75	244029819.34***	48250208.85.	13692149.37	9649425.58***	247760.81	2040704.51***	16651766.35
AUCK		34147.54	5745.61	0.81	3.72	582017317.1***	54314087.41	29880570.37	15481553.65***	629865.14*	2995929***	20730986.64	
Number of Spikes	SPP		4.79	1.05	0.70	4.01	2.67	2.07	1.07	0.64***	0.06*	0.21***	1.67
Main Shoot	NbGrains		39.47	8.11	0.80	7.15	184.54.	190.65.	65.76	39.5***	4.29***	5.48***	57.16
	GW	g	1.54	0.45	0.72	28.74	0.49*	3.3***	0.19	0.06***	0.01***	0.01***	0.11
	TKW	g	39.15	7.29	0.64	23.90	166.16*	1582.09***	141.67*	12.3***	4.24***	4.39***	32.20
	Grain.Area	mm ²	14.55	1.65	0.72	14.40	6.53*	70.3***	2.84	0.85***	0.2***	0.2***	1.53
	Grain.Width	mm	3.42	0.24	0.66	10.01	0.07	1.83***	0.16*	0.01***	4.62e-03***	4.99e-03***	0.03
	Grain.Length	mm	5.84	0.29	0.85	4.53	0.26*	1.24***	0.01	0.04***	4.54e-03***	4.49e-03***	0.04
Tillers	NbGrains		138.06	35.38	0.71	11.24	162.56	8587.89*	3586.11	765.63***	111.84***	207.08***	1813.77
	GW	g	5.37	1.64	0.68	26.24	0.62	57.14***	0.36	1.14***	0.25***	0.34***	2.97
	TKW	g	38.92	6.04	0.69	17.54	27.79	676.46***	96.72*	10.9***	2.93***	3.21***	20.80
	Grain.Area	mm ²	13.96	1.43	0.78	10.68	4.25*	30.97***	6.59*	0.85***	0.15***	0.13***	1.10
	Grain.Width	mm	3.37	0.20	0.75	6.65	0.03	0.68***	0.23**	0.01***	3.36e-03***	2.24e-03***	0.02
	Grain.Length	mm	5.71	0.29	0.86	4.16	0.19*	0.94***	0.07	0.05***	3.93e-03***	5.48e-03***	0.03

Wald and LTR test: ***, P-values <0.001; **, P-values <0.01; *, P-values <0.05 and '.', P-values <0.1 and non-significant P-value>0.05

GxY could be explained by uncontrolled environmental factors such as global radiation, which was far lower in 2016 than in 2017, even with the sodium-vapor lamps complement.

3.3 An overview of senescence traits

The response of senescence traits is detailed in [Table IV.1](#). However, it could be possible to get an idea of the evolution of heritability and the response of senescence all along the grain filling period. [Figure IV.3](#) compares the senescence curve variability based on raw data (pots) and based on gBLUP senescence parameters. For raw data, the extent of the confidence interval (CI) around the mean (as \pm one standard deviation (sd)) seems to be equivalent along the senescence period except at the end of senescence, where the CI area increases significantly ([Figure IV.3A](#)). The same pattern can be observed for the gBLUP based senescence variability ([Figure IV.3B](#)), but with the opposite behaviour at the end of senescence. In fact, varieties are declared as a random effect in the mixed model, so if there are no significant differences among genotypes the shrinkage factor will tend towards infinity and breeding value predictions will tend to the mean (low heritability). [Figure IV.3](#) shows that if there are differences between genotypes at the beginning of senescence, all the genotypes converged to values which are not discriminating. Furthermore, the figure presents the distribution of senescence kinetics for all genotypes in both treatments and we can see that heat treatment only had a slight effect on the senescence curves (the stress index is at 5.39 for the START and 3.75 for the XPI).

3.4 Genetic correlations

Relationships among traits and genotypes for each heat treatment, were investigated using Pearson's correlation and biplot analysis using the gBLUP per treatment from Eq. 7. [Figure IV.4](#) shows a subset of traits related to the grain characteristics of the main shoot, senescence traits, D.Z65 and SPP (correlations on the whole set are presented in [Suppl. IV.5](#). Vector length shows the extent of variation explained by each trait in the PCA. The first two axes explained up to 62 and 64% of total variability in NS and S treatment, respectively. The relationships between traits seem equivalent whatever the treatment. Two main trait clusters could be easily identified: (1) traits related to the grain as TKW, grain area, length and width, and (2) senescence traits. No significant correlation was detected between D.Z65 and TKW in NS treatment ($r = 0.08$, $p > 0.1$), but a weak and significant one was detected between these two traits in S treatment ($r = -0.24$, $p < 0.001$). We also detected weak, but significant ($p < 0.05$), correlations between TKW and senescence traits YPI ($r = 0.19$), fSTART (0.20), END ($r = -0.16$) and VIT ($r = 0.19$) in NS treatment. In S treatment, we still detected correlations between TKW and YPI ($r = 0.26$), fSTART ($r = 0.27$), and VIT ($r = -0.18$) but no correlation was detected between TKW and END. A new positive correlation was detected with AUC ($r = 0.19$).

3.5 GWAS

The single locus genome-wide association study, involving a total of 164,198 SNP, was performed using raw data for both years. Then, the main effect and the SNP by heat treatment interaction were

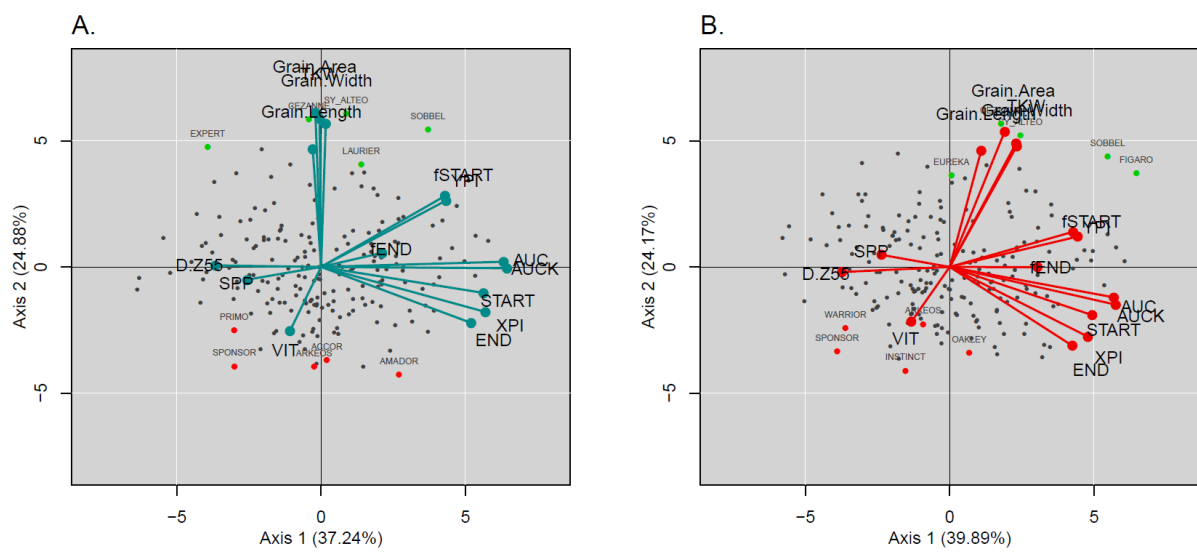


Figure IV.4: Biplots based on a subset of traits measured on 199 wheat varieties grown in two heat treatment situations

Scatterplot showing varieties' names calculated on agronomic traits as per the first two principal components. (A) represents NS treatment and (B) represents S treatment. The top and last five lines for the main shoot TKW are respectively plotted in green and red dots.

tested using Model 9. All phenotypic traits were tested, except the senescence parameters a and D , and $fEND$ which did not show significant genotypic effect. Four genomic areas were significantly associated with the SNP main effect for the flowering date considering the estimated 5.26 threshold ([Figure IV.5A, Table IV.2](#)). Whilst no SNP by treatment interaction was significantly associated with the flowering date ([Figure IV.5B](#)). TKW from the main shoot presented only one genomic area associated with the SNP main effect on chromosome 7A and to the SNP by treatment interaction effect on chromosome 4B ([Figure IV.5C, IV.5D, Tables IV.2, IV.3](#)).

[Figure IV.6](#) represents allelic effects for the most closely associated SNP for the flowering date (2D) and the most closely associated SNP for the TKW (4B). Firstly, we can see that there is no genotype dispersion around the bisecting line for the flowering date ([Figure IV.6A](#)), because there is no genotype by treatment interaction, whereas, a large dispersion could be observed for TKW ([Figure IV.6B](#)). Secondly, the figure shows the difference in the dispersion of both SNP alleles. For an SNP significantly associated with the main effect ([Figure IV.6A](#)), the allele separates genotypes only on the bisecting line. Whereas, for an SNP significantly associated with treatment interaction, the allele separates genotypes following the second bisecting line ($y=-x$).

Out of the four QTLs associated with the flowering date ([Table IV.2](#)), the highest one is located on chromosome 2D and it colocalizes with the photoperiod sensitivity gene *Ppd-D1* (Beales et al. 2007). This QTL also colocalized with two senescence traits, the start of senescence and the x-axis of the point of inflection. The effect of the minor allele had opposite effects, meaning that earlier genotypes had a delayed start to senescence. Furthermore, two more QTL were associated with senescence traits: one QTL located on chromosome 2A, which was associated with the maximum senescence rate (VIT), and one QTL located on chromosome 4D which was associated with the initiation and the end of senescence. The 4D QTL colocalizes with the *Rht-D1* dwarfing gene (Peng et al. 1999). This QTL is also associated with the number of grains per spike in the main shoot and tillers, but with an opposite effect compared to senescence traits. This means that early senescing plants had more grains and heavier spikes than later ones. Lastly, one QTL localized on chromosome 7A is associated with TKW for both the main shoot and the tillers. None of the QTL associated with the main SNP effect were associated with the SNP by treatment interaction.

Seven QTL were significantly associated with the SNP by treatment interaction effect ([Table IV.3](#)). Four are related to senescence traits, one located on chromosome 4B is associated with TKW and related traits (grain area, length, and width) of the main shoot, and two located on chromosomes 3B and 6B were associated with grain weight and TKW traits for tillers. All the effects of the major allele of the most significant SNP by QTL are presented in [Table IV.3](#). For example, varieties with allele G at QTL 4B, which is the major allele with a frequency at 0.64, have a lower performance in stressed

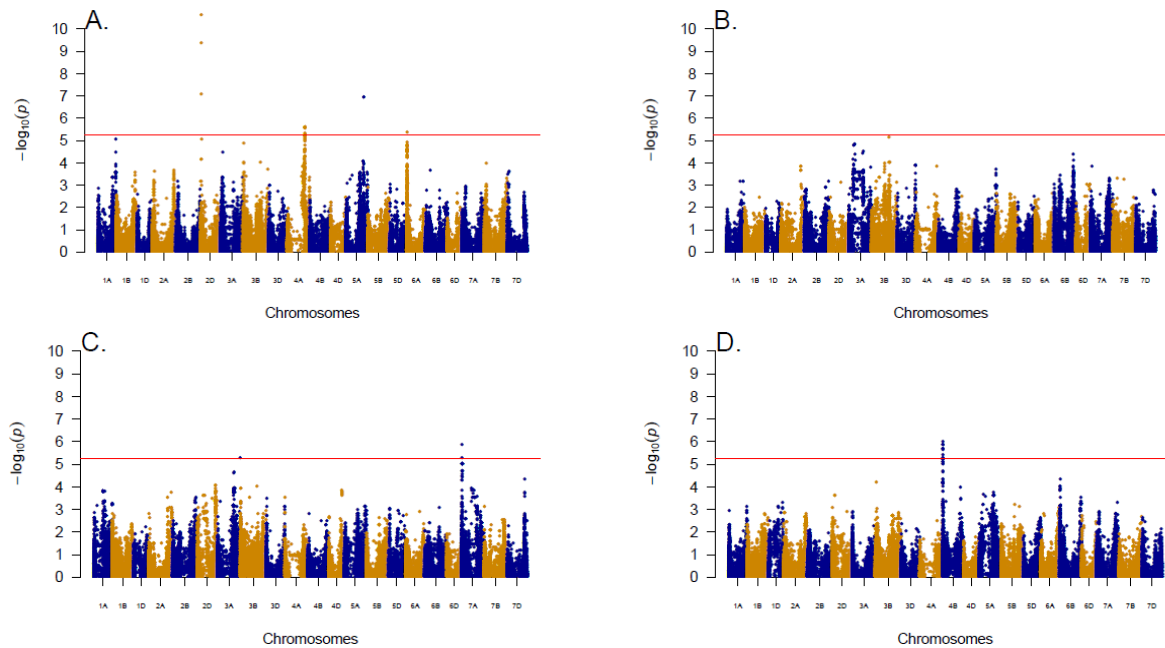


Figure IV.5: Manhattan plot of GWAS conducted on the SNP main effect and the SNP by treatment interaction for D.Z65 and TKW

A. SNP main effect for D.Z65; B. SNP by treatment interaction for D.Z65; C. SNP main effect for TKW; D. SNP by treatment interaction for TKW. The red line represents the 5.6 significant threshold calculated as per Gao et al. (2015).

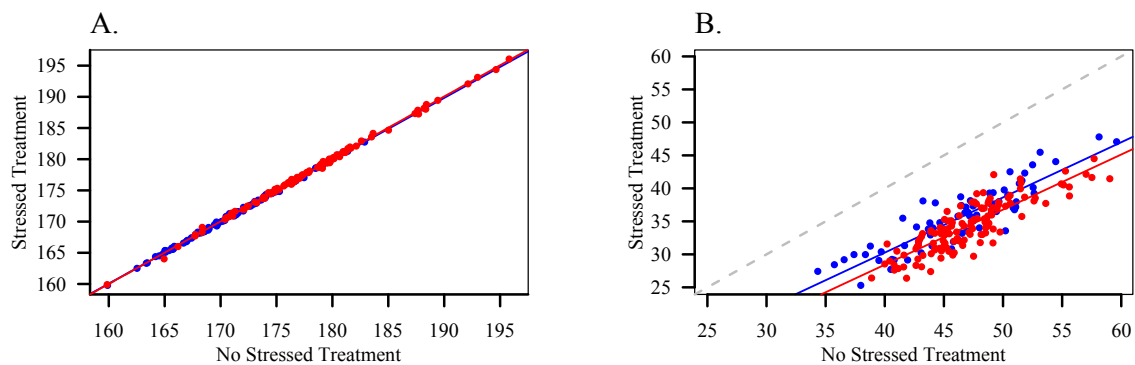


Figure IV.6: gBLUP values for SNP main effect and SNP by treatment interaction effect

A. Correlation between flowering dates for stressed and non-stressed plants: in blue, the minor allele effect of QTL 2D and in red the effect of the other allele. B. Correlation between TKW for stressed and non-stressed plants: in blue the minor allele effect of QTL 4B and in red the effect of the other allele

treatment than varieties with allele A. This effect can be visualized in [Figure IV.6B](#), it amounts to 3.51g (± 2.15).

As mentioned before, none of the QTL associated with an SNP by treatment interaction show a significant SNP main effect.

4 DISCUSSION

This study explored the impact of post-anthesis heat stress on a European elite wheat panel in controlled conditions. Physiological traits such as flowering time, senescence and grain morphologic traits were recorded. GWAS was performed to dissect the genetic determinants of heat tolerance.

4.1 Experiment validation and stress impact

Before interpreting GWAS results, it seems important to correctly describe the stress undergone by plants. Each plant in a heat stress condition undergoes a post-anthesis elevation of temperature, whose average daily temperature is 8.6°C in the daytime and 5.8°C at night-time. Heat stress applied for 10 days, reduced grain weight by 24% for the main shoot compared with an optimum temperature of 21/17°C. This result is equivalent to other reports on grain weight loss under post-anthesis heat stress (Stone and Nicolas, 1994; Gibson and Paulsen, 1999; Spiertz et al., 2006). The TKW decrease was more correlated to a reduction in width (SI = 4.5%) rather than in length (SI=10.0%). The higher sensibility of width over length has already been reported (Jamil et al., 2017). As expected, we observed no impact on the number of grains on the main shoot.

We decided to apply stress conditions 3 days after flowering, so as not to impact the number of grains on the main shoot. This choice was made to limit the impact of heat stress on only one of the yield components, grain filling. In fact, it has been shown that high pre-flowering temperatures may cause flower sterility (*e.g.* Farooq et al., (2011); Barber et al., (2017)). Since yield components are formed one after the other and there are multiple compensation mechanisms (Gate, 1995) liable to bias our interpretation. Furthermore, when plants are exposed to a long period of stress and they survive it, they can set up various mechanisms to acclimatize to this stress (Gaspar et al., 2002). As these two points could largely disturb physiological and genetic heat-resistance mechanisms, our experiment was conducted to limit them to identify solely heat stress QTLs specific to grain filling.

Furthermore, we chose to analyse separately grains from the main shoot and from tillers. Heat stress may actually have a different impact on the main shoot and tillers, due to the delay in flowering dates, which may be up to nine days (Jones et al., 2017).

Heat stress can also impact the number of grains on tillers by reducing pollen fertility (Farooq et al., 2011), while fertilization has already occurred on the main shoot. Stress decreased the number of grains on tillers by more than 5% in both experiments (2016 and 2017). Nevertheless, heat treatment reduces the grain weight for both the main shoot and tillers, with stress indices at 23.9% and 17.5%

Table IV.2: list of associated SNP following a GWAS conducted on 199 wheat varieties grown for two years with two post-flowering temperature treatments: SNP with no interaction effect. Position is the position on Chinese Spring refSeq1.0 (IWGSC et al 2018)

	Trait	Units	Chr.	Position (bp)	SNP	Allele	Allele Freq.	SNP			QTL boundary		
								LOD	Effect	Effect SD	Nb. SNP	Min. (bp)	Max. (bp)
Flowering Date	D.Z65	Days	4A	576159104	AX-89667446	A	0.36	5.62	3.98	2.35	13	575485603	581337826
	D.Z65	Days	5A	584428207	AX-89390208	C	0.15	6.93	6.16	2.42	2	583039479	585816937
	D.Z65	Days	6A	13010065	AX-89584501	G	0.45	5.39	3.63	2.34	1	10570285	15449845
	D.Z65	Days	2D	34995194	AX-89382235	T	0.48	10.62	-5.71	2.22	3	32510457	36442547
Senescence	XPI	°C.days	4D	17083871	AX-89398511	A	0.33	5.77	-65.29	29.72	4	16992006	18830860
	XPI	°C.days	2D	34995194	AX-89382235	T	0.48	6.11	50.98	28.82	1	33532314	36458074
	VIT	µg.cm ⁻³ per °C.days	2A	45486914	AX-89761338	G	0.16	5.54	0.04	0.02	1	44809106	46164722
	START	°C.days	2D	34995194	AX-89382235	T	0.48	6.24	56.48	26.04	2	33217357	36453954
	END	°C.days	4D	17083871	AX-89398511	A	0.33	7.06	-83.51	34.52	16	16826225	23189173
Number of Spikes	SPP		2B	25953023	AX-89528410	A	0.18	6.56	1.00	0.41	3	24831122	26711709
	SPP		2B	619396366	AX-89729195	T	0.09	6.47	1.82	0.44	170	605382213	652111584
Main Shoot	NbGrain		4D	19694106	AX-89363486	G	0.36	6.88	7.44	3.28	21	16840726	23189154
	GW	g	4D	17083871	AX-89398511	A	0.33	9.37	0.38	0.14	32	16664691	25447819
	TKW	g	7A	16603832	AX-89332971	G	0.32	5.87	3.60	1.84	3	15907587	17300104
	Grain.Area	mm	7A	16603832	AX-89332971	G	0.32	5.66	0.87	0.46	2	15907587	17300104
	Grain.Width	mm	7A	16603832	AX-89332971	G	0.32	6.13	0.12	0.06	3	15907587	17300104
Tillers	GW	g	4D	17430704	AX-89345211	A	0.36	8.99	1.69	0.61	28	13163753	24863780
	NbGrain		4D	19459578	AX-89445201	G	0.39	6.44	35.82	15.11	9	16839627	22372934
	TKW	g	7A	25929273	AX-89765935	C	0.37	5.86	3.13	1.71	21	25228927	26941996

D.Z65 = flowering date, XPI and YPI are the x-axis and y-axis of the inflexion point of senescence curve, VIT = rate of senescence, START = beginning of senescence, END = end of senescence, SPP = number of fertile spikes per plant, NbGrain = number of grains, GW = Grain weight, TKW = thousand kernel weight, Grain.Area = area of grain, Grain.Width = width of grain.

respectively. The lower decrease for tillers could be due either to compensation for the number of grains and/or a shorter stress period due to later flowering.

For monocarpic species in terminal heat stress, it is useful to determine whether the stress damaged the photosynthetic source more than the reproductive sink. Heat stress accelerates the rate of grain filling whereas its duration is shortened (Sofield et al., 1977; Dias and Lidon, 2009). It has been estimated that for every °C above the optimal growing temperature, the duration of grain filling is reduced by 2-3 days (Streck, 2005). In this study, we monitored leaf chlorophyll content along grain filling. A light impact of stress was observed on senescence traits. This impact was visible when plotting the decrease of chlorophyll content as a function of Julian days for both treatments ([Suppl. IV.6](#)). Looking at senescence dynamics as a function of degree days highlights this difference (Figure 3, [Tables IV.1](#)). The only senescence parameters which had a significant treatment effect were the number of °C for the start and the middle of senescence, with a stress index of 5.26% and 3.75% respectively. These observations confirm the fact that high temperatures shorten the plant cycle and limit the amount of light intercepted (Sofield et al 1977), more than changing the rate of senescence. Wheat grain development is usually separated into three phases: the lag phase, the filling phase and the maturation phase (Gate, 1995; Acevedo et al., 2006). During the lag phase, which lasts for about 15 days, the number of endosperm cells is fixed. It was proposed that the number of cells in the endosperm is regulated by the supply of assimilates available to grain during the first 2 weeks after anthesis (Brocklehurst 1977). This number will then regulate the rate of accumulation of dry matter during the filling phase. It is then likely that the stress we applied has modified the number of cells and then impacted the grain filling.

All in all, heat treatment has strongly impacted grain filling on the main shoot and tillers. The number of grains was only affected on tillers and heat treatment has only slightly affected senescence. A large variance and heritability range was observed among grain morphology traits and stay-green traits. This means that varieties within this panel, showed differences in terms of grain-filling and stay-green, so QTLs linked with these traits might be discovered. Furthermore, a large variance of GxT interaction was identified for grain weight, meaning that plant response to heat stress could be different depending on the genotypes and that QTLs linked with this response may also be discovered. The performance of genotypes for TKW and their responses to heat stress are presented in the [Suppl. IV.4](#). This result could be useful for breeders to check the performance of the varieties and their response to the stress applied in our experiment.

4.2 GWAS

4.2.1 QTL associated with traits

All in all, for the main SNP effect, we detected 9 QTL associated with at least one trait. Four of them were associated with the flowering date. The QTL on chromosomes 4A and 2D were previously found

Table IV.3: list of associated SNP following a GWAS conducted on 199 wheat varieties grown for two years with two post-flowering temperature treatments: SNP with a significant interaction effect. “Position” is the position on Chinese Spring refSeq1.0 (IWGSC et al 2018)

	Trait	Units	Chr.	Position (bp)	SNP	Allele	MAF	SNP by treatment			QTL boundary		
								LOD	Effect	Effect SD	Nb. SNP	Min. (bp)	Max. (bp)
Senescence	XPI	°C.days	7B	693020724	AX-89578121	C	0.09	5.71	60.40	34.72	2	692252452	693926665
	XPI	°C.days	1D	17642951	AX-89766697	A	0.38	5.38	-32.45	31.54	1	13902574	21383328
	YPI	µg.cm ⁻³	1B	631711683	AX-89638269	G	0.19	6.69	1.24	0.82	2	630767028	632656378
	AUCK		4A	328128331	AX-89642074	A	0.09	5.55	-3119.83	2355.64	36	38974540	472729606
Main Shoot	TKW	g	4B	1841022	AX-89364414	G	0.36	5.99	-3.51	2.15	11	1339055	2852831
	Grain.Area	mm ²	4B	1853262	AX-89500218	A	0.37	6.42	-0.78	0.53	34	1384350	3700062
	Grain.Width	mm	4B	1853262	AX-89500218	A	0.37	5.83	-0.11	0.07	9	1310533	2334232
	Grain.Length	mm	4B	1854364	AX-89738651	T	0.36	5.49	-1.14E-01	1.10E-01	25	1438558	3697489
Tillers	GW	g	3B	150697370	AX-89468413	C	0.09	5.40	1.48	0.74	1	141244647	160150093
	TKW	g	6B	695794696	AX-89726711	A	0.42	5.76	2.53	1.96	7	694970068	696420120
	Grain.Area	mm	6B	695794696	AX-89726711	A	0.42	5.85	0.58	0.51	8	694970068	696420120
	Grain.Length	mm	6B	695794696	AX-89726711	A	0.42	6.13	1.01E-01	1.10E-01	9	694970068	696420120

XPI and YPI are the x-axis and y-axis of the inflection point of the senescence curve, AUCK = total surface under the senescence curve, GW = Grain weight, TKW = thousand kernel weight, Grain.Area = area of grain, Grain.Width = width of grain, Grain.Length = length of grain.

to be associated with the flowering date in the study of Touzy et al. (Subm.), which have explored QTL for drought resistance on the same panel. Both QTL associated with the number of spikes per plant, located on chromosome 2B, were also identified in the study of Touzy et al. (Subm.), but not in association with the same trait.

Three QTL were associated with senescence traits, and two of them were co-located with the major genes *Rht-D1* and *Ppd-D1*. QTL 4D is co-located with the *Rht-D1* height gene. *Rht* genes (or dwarfing genes) have many effects on wheat development (eg Borrell et al., 1991; Peng et al., 1999). The G allele of molecular marker *AX-89398511*, which is the most associated SNP within this QTL, was positively associated with plant's height in Touzy et al. (Subm.), meaning that plants with the G allele carried the wild form of *rht-D1*. In our study, this allele is associated with an early middle and end of senescence, and a higher number of grains (for the main shoot and tillers). As plants were placed at random in the greenhouse, we could discard the idea that small plants were in competition for light with higher ones, which could have impacted the number of grains and senescence. Christopher et al., (2018) have also found that senescence QTL were associated with *Rht* genes. As was the case in our study, they could not conclude that *Rht-D1* is directly associated with senescence or whether it is due to a genetic linkage with other genes.

Regarding major genes for precocity, the photoperiod sensitivity gene *Ppd-D1* on chromosome 2D co-localizes with a QTL of the flowering date and QTLs for the start and the middle senescence. These co-localizations between phenology and senescence traits were largely observed in literature (Verma et al., 2004; Bogard et al., 2011; Christopher et al., 2016; Pinto et al., 2016). Verma et al., (2004) identified a QTL which located on the short arm of chromosome 2D and was associated with greater yield and delayed senescence. In their study, they used a Beaver × Soissons bi-parental population and found that the alleles responsible for delayed senescence came from Soissons. As we have Soissons in our panel, we checked the polymorphism of the highest associated molecular marker within QTL 2D and found a similar result. This means that earlier varieties, with the wild version of *ppd-D1*, have a longer stay-green for flag leaf. Based on different studies, Bogard et al., (2011) presented four assumptions to decipher the relationship between wheat precocity and senescence. Among these assumptions, three involved the direct or indirect effect of comparing varieties with different levels of precocity in the same field. In field experiments, late plants often undergo terminal stress, and as senescence is highly dependent on environmental conditions such as water, temperature and nitrogen availability, it could be expected that late plants have shorter stay-green capacities due to environmental effect. However, as our experiments were in controlled conditions (no water stress and non-limiting nitrogen) we could discard these assumptions and prioritize the hypothesis presented by Wingler et al., (2010). This assumption presents the relationship between precocity and leaf senescence, in *A. thaliana*, as a change in the metabolism occurring soon after anthesis. In fact,

dismantling leaf structural components releases sugars in the leaves after anthesis, which could trigger senescence.

Only one QTL on chromosome 7A was associated with TKW. Unfortunately, no co-location between this QTL and the QTL identified in Touzy et al. (Subm.) study was identified. This observation does not put all the work that has been carried out in any of these studies into question as they were conducted in very different conditions. The study of a complex trait such as heat tolerance is easier in controlled conditions, but the representativeness of this kind of studies in a breeding or agriculture context is certainly lower than in field experiments. However, as explained by Tardieu, (2012), both kinds of studies are important as they enable different behaviours to be accessed.

4.2.2 QTLs associated with a response to heat stress

Seven QTL associated with the response to heat treatment were identified. None of them are associated with flowering time, which confirmed our protocol and allowed us to identify the QTL linked to heat stress tolerance, without risking any confounding effect due to escape.

Four QTL are associated with senescence traits. On chromosome 4A, the QTL associated with the area under the curve (AUCK) is centromeric. As linkage disequilibrium is high in this region, the probability of identifying co-localization with other QTL or genes from literature are therefore high. For the remaining QTL, it is interesting to check co-localization with the study made by Touzy et al. (Subm.), because they have used the same panel of varieties tested in field trials. They grouped the different trials, from a large multi-environmental network, depending on their water stress scenarios. They have identified QTL in four different water stress scenarios, an optimal one with no stress, terminal stress, water stress during anthesis and a scenario with stress from booting to harvest. Even if the QTL which have been detected in their study are suspected of being associated with drought stress, they did not check for heat stress and they also pointed out that the drought stress is often associated with heat stress.

Three QTLs associated with the response of grain weight to heat stress were identified in our study. Two of them present a co-localization with the study made by Touzy et al. (Subm.). The most associated SNP of the QTL located on chromosome 3B has also been associated with grain yield in environments with water stress during the anthesis period (Touzy et al. Subm.). The QTL located on chromosome 4B, which is associated with the response of main shoot TKW, co-localized with a QTL associated with the TKW in all the water stress scenarios and in optimal conditions. More than 14% of the plants' stress responses could be explained by this QTL. We have checked numerous studies that present QTLs located on the chromosome 4B for the TKW and grain yield in heat condition, many of these studies are reviewed by Zhang et al., (2010), Bouffier, (2014) and Tricker et al., (2018). Pinto et al., (2010) also reported two QTL in heat conditions for TKW and plant height on chromosome 4B. They also have highlighted the fact that numerous studies identified this QTL but suggest that this

region may co-localized with the Rht-B1 gene. However, they have identified this QTL with a biparental Seri and Babax population, which did not segregate for any Rht genes. Then, we checked the position of our QTL and Rht-B1 ([Suppl. IV.7](#)) and found that no genetic link between them. Therefore, it seems possible that one or more genes may be linked to tolerance to thermal stress in this zone, independently of Rht-B1.

Finding co-localizations between experiments conducted in controlled conditions and in the field allowed us to validate our approach. Even if these QTL did not explain a large part of the variability of the response to stresses in the field or in the greenhouse, we can rely on these results for improving our understanding of genetic responses to abiotic stresses.

CONCLUSION

In a context of climate change, we have investigated terminal heat stress tolerance using a European panel of elite winter wheat varieties. Optimal conditions and post-anthesis heat treatment was applied to plants grown in the greenhouse, which enabled us to identify varieties capable of tolerating this stress and perform GWAS for identifying the genetic determinant of heat tolerance. Grain weight was strongly impacted by stress, while senescence traits were not. We particularly identified a significant SNP by treatment interaction for TKW that revealed a QTL on chromosome 4B. This study only focused on one particular terminal heat stress although other stress scenarios may impact crops. However, our findings could help to identify genes linked with heat stress tolerance or help breeders to select specific QTL.

Chapitre V : Prospection multi-stress

Chapitre V : Prospection des effets alléliques en fonction de facteurs de stress dans un réseau d'essais multi-environnemental

Gaëtan Touzy^{1,2} – Renaud Rincint³ – Stephane Lafarge² – Willem Kruijer⁴ – Fred Van Eeuwijk⁴ – Jacques Le Gouis⁴ – Sébastien Praud^{2*}

¹ Arvalis-Institut du végétal, Biopôle Clermont Limagne, 63360 Saint-Beauzire, France.

² Biogemma, Centre de recherche de Chappes, Route d'Ennezat CS90216, 63720 Chappes, France.

³ INRA, UCA UMR 1095, Génétique, Diversité et Ecophysiologie des Céréales, 24 Avenue des Landais, 63177 Aubière Cedex, France.

⁴ Wageningen University, 6700AA Wageningen, Netherlands (W.K., F.V.E.)

* : corresponding author : sebastien.praud@biogemma.com

1 INTRODUCTION

Après avoir décrit le comportement et les bases génétiques du panel élite face à un ou plusieurs scénarios de stress hydriques et thermiques, nous avons exploré une approche visant à identifier les déterminants génétiques de la réponse des plantes à de multiples conditions environnementales. En effet, les Chapitres III et IV se sont focalisés spécifiquement sur des stress hydriques et thermiques et comme il a été présenté dans le Chapitre I, il existe de nombreuses stratégies d'adaptation des plantes à ces stress (Reynolds et al., 2006; Tardieu, 2012). Chacune des stratégies d'adaptation peut avoir un effet positif dans des scénarios de stress et un effet négatif dans d'autres cas ([Table I.3](#)). Par exemple, la fermeture des stomates est l'une des stratégies d'évitement de stress hydriques chez les plantes. Cette stratégie est efficace dans des scénarios de stress terminaux sévères, car elle permet de conserver une humidité du sol qui pourra être utilisée par la suite, en contrepartie elle a pour effet d'augmenter la température du couvert, car une quantité plus faible d'eau est évaporée. On voit bien dans cet exemple toute la nécessité d'élaborer une approche de détection de déterminants génétiques liés à un ou l'autre des stress, ou même à une combinaison de ces stress. Cependant, il serait dommage de prendre en compte les stress thermiques et hydriques en occultant les autres types de stress qui pourraient interagir avec les plantes.

De nombreuses interactions entre stress peuvent apparaître dans les expérimentations au champ (Suzuki et al., 2014). Les interactions entre nutrition azotée et nutrition en eau en font parties. Sans faire une description détaillée de ces interactions, il semble important d'en relever les points principaux, car elles sont certainement les plus importantes dans notre réseau d'expérimentation. En effet, du point de vue de la croissance de la plante, les interactions entre nutrition en eau et azotée peuvent avoir de multiples conséquences. Par exemple, le manque d'eau (ou d'azote) réduit la croissance des plantes et donc la demande en azote (ou en eau). Au contraire, un déficit de nutrition azotée peut réduire le développement de la plante, donc réduire la surface de recouvrement du sol et ainsi augmenter l'évaporation de l'eau du sol. Enfin, l'absorption et l'assimilation en azote par les racines sont directement en lien avec les transports d'eau dans la racine et dans la plante. Ces interactions n'ont été abordées que rapidement dans cette thèse *via* la présentation des équations d'élaboration de la biomasse et de la surface foliaire qui sont co-limitées par un facteur de stress hydrique ou de stress azoté. Plusieurs études y font référence et notamment la revue de Sadras et al., (2016) qui présente les impacts des interactions entre l'eau et l'azote, aussi bien en terme agronomique, physiologique, économique et de sélection et de modélisation.

Après avoir identifié les différentes causes de l'impact de stress biotique au sein d'un réseau d'expérimentations multi-environnemental et surtout leurs impacts sur le classement des variétés (donc sur l'interaction GxE), la question qui s'ouvre dans ce chapitre est d'analyser plus finement les régions génomiques dont l'effet, sur le phénotype, dépend de l'environnement. En somme, existe-t-il de l'interaction QTLxE dans notre panel et serons-nous capables d'associer cette interaction QTLxE à

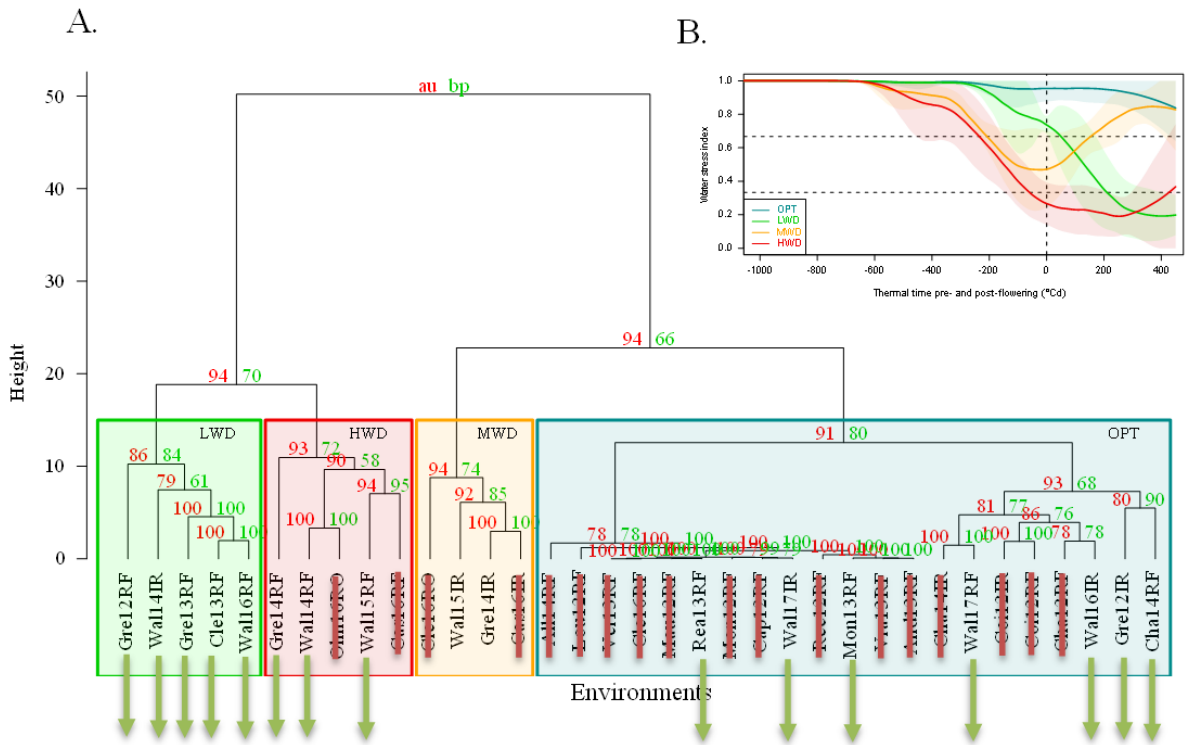


Figure V.1: Présentation de la sélection des environnements d'après la classification d'essais réalisée dans Touzy et al. (en prep.)

La figure A. représente le dendrogramme issu de la classification des environnements faite par Touzy et al. (en prep). Cette classification a été réalisée suivant le profil d'un index de stress hydrique calculé tout au long du cycle de la culture. La figure B. présente les profils d'indices de stress hydrique qui ont été moyennés pour chacun des quatre groupes. Les flèches vertes représentent les environnements qui ont été sélectionnés pour cette étude.

une variation quantifiable des environnements du MET, par exemple, l'évolution d'un ou de plusieurs facteurs de stress. Plusieurs études ont présenté des QTL d'interaction répondant à une ou des variations de variables environnementales (Malosetti et al., 2013; Bouffier, 2014; Millet et al., 2016). En 2010, van Eeuwijk et al., présentent une revue de l'ensemble des méthodes (notamment les modèles mixtes) pour identifier et extraire le maximum d'informations sur l'interaction QTLxE. Ils présentent aussi l'utilisation de modèles de culture pour décrire les facteurs environnementaux qui impacteraient les plantes et interagiraient avec des régions génomiques. Il est intéressant de noter que la quasi-totalité des études qui ont été menées sur les interactions QTLxE ont été réalisées avec des populations biparentales. L'une des raisons majeures de l'absence de ce type d'analyses avec les panels d'association est la prise en compte des corrélations entre environnements et des relations entre individus (apparentement), ce qui mène à des temps de calculs trop longs pour analyser plusieurs milliers de SNP. Cependant, Xu et al., (2015) ou encore Millet et al., (2016) ont présenté une approche en plusieurs étapes pour résoudre ce problème. Dans cette partie nous nous sommes largement inspirés de ces études pour tester les interactions QTLxE.

Le but de ce chapitre est donc d'identifier des régions génomiques en lien avec des variations de facteurs environnementaux dans un réseau d'essais. Cependant, pour rester dans le sujet de cette thèse, une sélection des environnements a été réalisée pour maximiser nos chances d'identifier des QTL de réponses à des températures élevées et/ou aux stress hydriques, sans pour autant oublier les autres types de stress qui peuvent intervenir dans ce réseau multi-environnemental. La diminution du nombre d'environnements nous permettra aussi de réduire les temps de calculs. Les travaux qui sont présentés dans ce chapitre ont été réalisés en collaboration avec l'équipe de Fred van Eeuwijk, notamment Willem Kruijer, durant un séjour de trois mois dans leur laboratoire. La démarche scientifique qui est présentée ici est largement inspirée des différents travaux issus de leur laboratoire (Boer et al., 2007; van Eeuwijk et al., 2010; Malosetti et al., 2013; Millet et al., 2016). Ainsi, trois étapes majeures peuvent être identifiées dans ce chapitre : (1) description du réseau multi-environnemental (MET), qui passe notamment par une sélection de covariables environnementales qui impactent le rendement et ses composantes ; (2) détection d'interactions QTLxE via des modèles multi-environnementaux et multi-locus ; (3) mise en relation des effets des interactions QTLxE avec les valeurs de covariables environnementales pertinentes.

2 MATERIEL ET METHODES

2.1 Matériel végétal et expérimentation

L'ensemble des données génotypiques et phénotypiques utilisées dans cette étude ont été présentées dans le Chapitre III. La [Figure V.I](#) présente les environnements qui ont été sélectionnés pour représenter un maximum de scénarios hydriques, et aussi des scénarios de stress thermiques. Le choix de la sélection des environnements a aussi été réalisé au regard du nombre de génotypes présents dans

Table V.1: Indicateur statistique de la qualité des expérimentations et variabilité du rendement en grain dans le MET

Les indicateurs statistiques ont été extraits du premier stade des analyses GxE (Chapitre III). ETs: Type d'environnement avec : OPT condition optimale en terme de disponibilité en eau, LWD déficit hydrique tardif (post-floraison), MWD déficit hydrique modéré pendant floraison and HWD Fort déficit hydrique pré-floraison. Moyenne : moyenne générale dans les environnements (en t/ha) CVe: coefficient variation environnementale. Res.: variance résiduelle (en (t/ha)²). H_g² héritabilité généralisée au sens large. AIC: Akaike Information Criterion.

Environnements	ETs	Nb genotype	Moyenne	CVe	Variances			AIC
					Genetic	Res.	H _g ²	
Gre12IR	OPT	174	7.52	3.09	0.27	0.05	0.89	1722.38
Rea13RF	OPT	174	8.57	5.3	0.48	0.21	0.89	4145.26
Mon13RF	OPT	174	8.36	3.72	0.32	0.1	0.85	1984.42
Cha14RF	OPT	174	7.56	3.55	0.45	0.07	0.91	2120.04
Wal16IR	OPT	174	9.67	3.88	0.7	0.14	0.90	2047.56
Wal17IR	OPT	174	8.88	3.84	0.25	0.12	0.81	1856.29
Wal17RF	OPT	174	7.84	4.19	0.22	0.11	0.8	1817.3
Gre12RF	LWD	174	6.75	5.38	0.15	0.13	0.66	1945.45
Gre13RF	LWD	174	9.00	5.53	0.21	0.25	0.6	2357.1
Cle13RF	LWD	174	9.17	3.55	0.38	0.11	0.86	2150.41
Wal14IR	LWD	174	8.24	8.52	0.35	0.49	0.59	2311.52
Wal16RF	LWD	174	7.08	4.78	0.68	0.11	0.91	2099.07
Gre14IR	MWD	174	7.87	5.03	0.37	0.16	0.81	2149.09
Wal15IR	MWD	174	8.96	5.48	0.76	0.24	0.86	2240.98
Gre14RF	HWD	174	4.26	8.14	0.05	0.12	0.44	1814.23
Wal14RF	HWD	174	5.23	12.15	0.19	0.4	0.48	2202.48
Wal15RF	HWD	174	6.25	7.14	0.98	0.2	0.9	2276.87

chaque essai (pour avoir un MET équilibré), de la mesure de chacun des traits phénotypiques et de la qualité des données (par exemple, une forte héritabilité du rendement [Table V.1](#)). C'est au final 174 lignées de blé tendre élites européennes qui ont été évaluées dans 17 environnements différents, un environnement étant défini par la combinaison de l'année, du lieu et du type d'itinéraire technique (irrigation ou non).

Dans chaque environnement, quatre traits phénotypiques ont été mesurés : la date d'épiaison (D.Z55) en nombre de jours après le 1^{er} janvier (stade décrit par Zadoks et al. (1974)), le rendement en grain (GY) en quintaux par hectare rapporté à 0% d'humidité, le poids de mille grain (TKW) en grammes et enfin, le nombre d'épis par mètre carré (SA). Le nombre de grains par épi (GPS) a été calculé suivant la formule $GY/(TKW \times SA)$.

De la même manière que dans le Chapitre III, l'analyse du réseau multi-environnemental a été réalisée en deux étapes : (1) une première analyse statistique a été faite essai par essai. Elle a eu pour but d'appliquer le meilleur modèle statistique pour identifier et nettoyer les données, puis d'extraire les moyennes ajustées et enfin, calculer les héritabilités. (2) dans un second temps, les moyennes ajustées ont été utilisées dans les analyses de l'interaction GxE présentées par la suite.

Dans l'ensemble des données du MET, seulement 2% des données sont manquantes. Il en résulte donc un jeu de données assez bien équilibré, c'est-à-dire que chaque génotype a été évalué dans tous, ou pratiquement tous, les environnements. Les données manquantes sont le fait de différents facteurs : soit la variété n'a tout simplement pas été implantée dans le champ, soit la variété a subi des dégâts et les données sont inexploitable. Pour faciliter les calculs et limiter un maximum les biais, une imputation des données phénotypiques a été réalisée grâce au package R *missMDA* (Josse et al., 2012; Josse and Husson, 2016).

2.2 Données moléculaires

Toutes les variétés ont été génotypées avec une puce Affymetrix Axiom SNP de 280K présentée par Rimbart et al., (2018). Seuls les SNP polymorphes à haute résolution ont été utilisés dans cette étude. Tous ces SNP ont été physiquement cartographiés sur la séquence de référence du génome RefSeq V1.0 (Appels et al., 2018). Les loci hétérozygotes ont été considérés comme des données manquantes. De plus, les SNP pour lesquels il manquait plus de 10 % des données et les SNP dont la fréquence des allèles mineurs (FAM) était inférieure à 5 % ont été retirés de l'analyse. Enfin, les données manquantes ont été imputées avec Beagle 4.1 (Browning and Browning, 2016). En résumé, ce sont 161,772 SNP qui seront testés dans les analyses de GWAS présentées par la suite.

L'effet polygénique a été modélisé à l'aide d'une matrice d'apparentement (K). Les coefficients de parenté ont été calculés par paires selon la première méthode décrite par VanRaden (2008), avec un sous-ensemble de 27 680 SNP obtenus par élimination des SNP fortement corrélés. En bref, nous

avons calculé le r^2 pour chaque paire de SNP situés sur le même chromosome. Un dendrogramme a été construit à partir de cette matrice de r^2 et un SNP a été choisi au hasard dans chaque groupe avec une distance $<0,1$ (avec distance = $1 - R^2$).

A partir des résultats de GWAS, les limites des bornes des QTL ont été définies selon la méthode décrite dans Cormier et al (2014). En résumé, les SNP appartenant au même groupe de déséquilibre de liaison (LD) ont été définis comme un groupe de nucléotides à caractère quantitatif (QTN). Le regroupement a été effectué en faisant la moyenne de la distance des r^2 et l'arbre a été coupé au seuil critique de LD (LD critique = 0,24 (Bresgello, 2005)). Les limites des QTL ont été définies comme étant la position maximale et minimale sur la carte physique des SNP à l'intérieur du QTN. Ces positions ont été étendues par la décroissance locale de LD, qui est spécifique à la région génomique du QTL.

2.3 Description des analyses multi-environnementales

La première étape de l'analyse du MET consiste à modéliser les données phénotypiques *via* un modèle mixte multivarié. Le but de ce modèle est d'identifier et d'estimer les composantes de la matrice de variance/covariance (VarCov) en incluant possiblement les hétérogénéités des variances génétiques dans chaque environnement et les hétérogénéités des covariances génétiques par paire d'environnements. Suivant les résultats présentés dans le Chapitre III, une grande variabilité de la date de floraison existe dans notre panel et cette variabilité peut être expliquée à plus de 30% par la variabilité au QTL bien connu localisé sur le chromosome 2D, parce qu'il colocalise avec le gène de sensibilité à la photopériode *Ppd-D1*. Il a donc été choisi dans cette étude de prendre en compte une partie de cette variabilité en incluant un marqueur diagnostique de *Ppd-D1* en covariable dans notre modèle. Ce choix a été fait pour limiter les phénomènes d'échappement au stress qui sont souvent observés avec un panel et dans ce type de réseau d'essais. Au final le modèle est défini par :

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_j \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 1_n + \alpha_1 Ppd \\ \vdots \\ \mu_j 1_n + \alpha_j Ppd \end{pmatrix}, \Sigma_G \otimes K + \Sigma_E \otimes I_n \right) \quad \text{Mod. [1]}$$

où n est le nombre de génotypes ($I=174$), j le nombre d'environnements ($J=17$), Y_1, \dots, Y_j sont les vecteurs-colonnes avec n lignes contenant les valeurs phénotypiques dans les J environnements, μ_1, \dots, μ_j sont les moyennes de chaque environnement et $\alpha_1, \dots, \alpha_j$ sont les effets spécifiques du SNP *AX-89636338* pour chaque environnement, ce SNP est diagnostique des effets des allèles du gène *Ppd-D1* sur la phénologie de la plante (cf. résultat Chapitre III). Les matrices de VarCov sont déterminées par la matrice K de dimension $n \times n$ qui a été présentée précédemment, la matrice Σ_G de dimension $J \times J$ qui décrit les variances et covariances génétiques entre environnements et la matrice diagonale Σ_E de dimension $J \times J$ qui décrit les variances résiduelles spécifiques à chaque environnement. Le

Table V.2 : les différentes formes de la matrice Σ_G

<i>Model</i>	$var(g_{ij})$	$cov(g_{ij}; g_{ij}^*)$	n_{par}	n_{par} avec $J=17$
<i>ID</i>	σ_{GE}^2	0	1	1
<i>DG</i>	σ_{GEj}^2	0	J	17
<i>FA1</i>	$\lambda_{1j}^2 + \sigma_{GEj}^2$	$\lambda_{1j}^2 \lambda_{1j}^{2*}$	$2J$	34
<i>FA2</i>	$\lambda_{1j}^2 + \lambda_{2j}^2 + \sigma_{GEj}^2$	$\lambda_{1j}^2 \lambda_{1j}^{2*} + \lambda_{2j}^2 \lambda_{2j}^{2*}$	$3J-1$	50
<i>FA3</i>	$\lambda_{1j}^2 + \dots + \lambda_{3j}^2 + \sigma_{GEj}^2$	$\lambda_{1j}^2 \lambda_{1j}^{2*} + \dots + \lambda_{3j}^2 \lambda_{3j}^{2*}$	$4J-3$	65
<i>FA4</i>	$\lambda_{1j}^2 + \dots + \lambda_{4j}^2 + \sigma_{GEj}^2$	$\lambda_{1j}^2 \lambda_{1j}^{2*} + \dots + \lambda_{4j}^2 \lambda_{4j}^{2*}$	$5J-6$	79
<i>UN</i>	σ_{Gj}^2	$\sigma_{j,j}^{2*}$	$J(J+1)/2$	153

symbole \otimes est le produit de Kronecker. Ce modèle a été ajusté grâce au logiciel ASREML-R développé par (Butler et al., 2009).

Dans cette étude, plusieurs formes de matrice Σ_G ont été testées pour décrire le plus justement possible les variances génétiques au sein de chaque environnement et les relations génétiques entre environnements. Suivant le type de relations, certaines formes de matrices sont plus ou moins réalistes et adaptées à notre jeu de données. Cette approche est similaire à l'approche qui a été présentée par Boer et al., (2007). Ce sont donc six modèles qui ont été testés : le modèle avec des variations génétiques identiques (ID) qui est le plus simple, où il n'y a pas de relations entre les génotypes et où la variance génotypique est la même dans tous les environnements ; le modèle de variation génétique hétérogène (DG) où les variances génotypiques sont spécifiques à chaque environnement mais où il n'y a toujours pas de covariance ; enfin quatre autres modèles qui sont dit facteur analytique (FA) d'ordre 1 à 4 et qui seront décrits par la suite. Le modèle le plus complet est un modèle où chacun des termes de la matrice Σ_G serait estimé individuellement, c'est un modèle dit non-structuré (unstructured). Cependant, ce type de modèle demande un nombre important de paramètres à estimer ($npar = J \times \frac{J-1}{2}$), où J est le nombre d'environnements. Dans notre cas, ce type de modèle n'a pas convergé après 1000 itérations. Les modèles FA sont ceux qui s'approchent le plus de la réalité tout en limitant le nombre de paramètres à estimer (Boer et al., 2007).

Dans les modèles FA, qui ont été présentés par (Piepho, 1998, 1997; Smith et al., 2001), la matrice Σ_G est définie par la formule suivante $\Sigma_G = \Gamma\Gamma^T + \Psi$, où Γ est une matrice des facteurs de taille $J \times q$, avec J le nombre d'environnements et q le nombre de facteurs qui sont impliqués, Ψ est la matrice diagonale des variances spécifiques à chaque environnement. Le nombre maximal de facteurs (Q) est égal au nombre d'environnements moins un. Le nombre de paramètres à estimer pour chaque modèle est présenté dans la [Table V.2](#). Le choix des modèles a été réalisé en essayant de minimiser le critère d'information bayésien (BIC) (Hastie et al., 2001; Broman and Speed, 2002) et en essayant de minimiser le facteur d'inflation de la régression des QQ-plot issus du modèle de GWAS multivariés qui est présenté par la suite. A partir de la matrice de VarCov Σ_G , les corrélations génétiques ont été calculées suivant le calcul matriciel $cor(g) = diag(\Sigma_G)^{-1/2} \times \Sigma_G \times diag(\Sigma_G)^{-1/2}$.

2.4 Calcul des covariables environnementales

La caractérisation des facteurs limitant le développement du blé tout au long de son cycle est centrale dans cette étude. Une multitude de variables susceptibles d'apparaître dans un contexte de culture européenne a été progressivement définie par les agronomes (Meynard and Sebillotte, 1994; Brancourt-Hulmel et al., 1999; Lecomte, 2005). Le calcul de ces variables (notées EC pour « Environmental Covariate ») a été réalisé grâce au modèle de culture CHN, qui est un modèle de culture développé par Arvalis. CHN repose sur des principes physiologiques décrits par Gate (1995), Meynard (1997) and Chatelin et al. (2005) et sont présentés par Le Bris et al., (2016). Une description

de l'application de CHN au réseau d'expérimentation a été présentée dans les Chapitres I et III. Grâce à CHN, deux types de variables vont être calculées : (1) les variables classiques, ou indicateurs limitants, qui sont calculées à partir de la prédiction des stades de phénologie des variétés et des variables climatiques classiques comme les températures minimales et maximales en °C, les précipitations en mm, l'évaporation potentielle (Penman) en mm, et le rayonnement global en calorie par m² ; (2) les variables dites intégrées, qui sont des variables de stress liées à la nutrition azotée et hydrique de la plante et impactant la biomasse ou la surface foliaire (LAI). Ces variables sont directement issues du modèle de culture CHN. Chaque EC a été calculé pour chacune des variétés présentes dans l'essai. Si une variété n'était pas présente dans l'essai, alors une imputation a été réalisée en prenant les valeurs des EC de la variété Soissons. Au final, un total de 212 EC ont été calculées sur les 174 génotypes. Les EC ont été moyennées par environnement pour former une matrice de taille J x 212. Par la suite, cette matrice (Ω) sera centrée/réduite pour permettre la comparaison entre les EC. Une description de chacune des covariables est présentée dans l'Annexe V.1. Parmi ces EC, seules 18 ne présentent pas de variabilité dans notre jeu de données et ont été éliminées.

2.5 Sélection des covariables

Le nombre et la diversité des environnements présents dans notre réseau vont définir la pertinence d'une covariable par rapport à l'autre. En effet, même si les covariables que nous avons calculées ont été définies par plusieurs études réalisées sur du blé en condition de culture européenne, il est peu probable qu'elles soient toutes pertinentes dans notre MET. De plus, il est possible qu'il y ait une forte redondance au sein des EC, et que certaines covariables soient fortement corrélées à d'autres du même type ou d'un type différent.

Cette sélection des EC a été réalisée en deux temps : (1) la première sélection a été inspirée de la procédure présentée dans l'étude de Bouffier et al., (2015). La matrice Ω a été regroupée en k-clusters grâce à la fonction *clara* (R Development Core Team 201) qui est une méthode de partitionnement autour des médoïdes. Dans chaque groupe, l'EC la plus représentative a été extraite. Cette EC est équivalente au barycentre de chaque groupe, elle va en effet partager un minimum de dissimilarité avec les autres covariables du même groupe. Le nombre K de groupes a été fixé à 17 (équivalent au nombre d'environnements du MET). (2) La seconde sélection est une procédure pas à pas qui vise à identifier le meilleur sous-set de ces 17 covariables qui permettrait de reconstruire au mieux la matrice de VarCov génétique entre environnements (Σ_G) issue du modèle [1], cette procédure a été présentée dans l'étude de Rincent et al. (en prep). Pour cela, une matrice de similarité entre environnements (Σ_{EC}) a été construite suivant la formule : $\Sigma_{EC} = \mathbb{1}_J - \frac{D_{\Omega'}}{\max(D_{\Omega'})}$, avec $D_{\Omega'}$ la matrice de distance euclidienne calculée à partir de la matrice Ω' . La matrice Ω' a été construite pas à pas en intégrant à chaque étape l'EC qui permet d'augmenter au mieux la corrélation entre Σ_{EC} et Σ_G . Le sous-ensemble

optimal de covariables environnementales a été défini à partir du moment où l'ajout d'une covariable dans la matrice Ω' diminue la corrélation entre Σ_{EC} et Σ_G .

2.6 Détection des SNP associés aux différents traits phénotypiques dans le réseau multi-environnemental

La détection des SNP associés à un ou plusieurs environnements du MET a été réalisée en trois temps. La première prospection a été réalisée par un modèle de GWAS classique mono-locus pour chaque trait dans chaque environnement. Ensuite, un modèle mono-locus et multi-environnements (MET-GWAS) a été réalisé pour chaque trait sur l'ensemble des environnements simultanément. Enfin, une sélection des SNP a été réalisée grâce un modèle multi-environnemental multi-locus qui intègre pas à pas le SNP le plus significatif (stepwise forward). Dans ce dernier modèle, c'est seulement un sous-ensemble de SNP qui a été testé. Ce sous-set est issu de l'identification de QTL associés aux traits (P -value < 0.001) dans les deux modèles précédents. Les QTL ont été définis suivant la méthodologie présentée précédemment (Cormier et al., 2014) et uniquement le SNP le plus significatif a été utilisé dans la troisième étape.

Cette démarche a été présentée dans l'étude de Millet et al., (2016), le but est d'identifier et d'estimer l'évolution des effets des SNP dans le MET. De la même manière que pour l'analyse phénotypique multi-environnementale, le marqueur diagnostique de l'effet du gène *Ppd-D1* a été pris en compte dans chacun des modèles.

2.6.1 GWAS classique par trait et par environnement

Dans un premier temps, les études d'association ont été réalisées pour chaque trait dans chaque environnement. Les 161,772 SNP ont été testés avec le modèle mixte mono-locus présenté par (Yu et al., 2006):

$$Y_i = \mu + \alpha Ppd + x\beta + g_i + \varepsilon_i \quad \text{Mod. [2]}$$

où Y_i est le vecteur des valeurs phénotypiques dans un environnement donné, μ est la moyenne générale, Ppd est le vecteur des scores du SNP diagnostique de *Ppd-D1* et α est l'effet additif de *Ppd-D1*, x est le vecteur des scores du SNP à tester et β est l'effet additif de ce SNP, g_i représente les effets aléatoires polygéniques, dont la matrice de variance/covariance est estimée par la Kinship (K_{chr}) suivant $g_i \sim \mathcal{N}(0, \sigma_g^2 K_{chr})$, et ε_i les effets aléatoires résiduels définis par $\varepsilon_i \sim \mathcal{N}(0, \sigma_e^2)$. Les effets des SNP ont été estimés par la méthode des moindres carrés généralisée et la significativité ($H_0 : \beta = 0$) a été testée avec un test de Fisher. Le logiciel FaST-LMM v2.07 (Lippert et al., 2011b) a été utilisé pour réaliser cette GWAS.

2.6.2 GWAS multi-environnements par locus et par trait

Dans un second temps, les études d'association multi-environnementales ont été réalisées suivant la méthode de Zhou and Stephens, (2014). Cette méthode utilise le même modèle que dans la partie précédente, c'est-à-dire un modèle mixte où les phénotypes suivent une distribution normale multivariée, plus l'effet d'un SNP dans chaque environnement qui va être testé. Etant donné que ce type de modèle peut être extrêmement long à converger (plusieurs minutes ou heures dans certains cas) et qu'il faut l'exécuter pour chaque SNP, Zhou and Stephens, (2014) proposent de modéliser les matrices de VarCov sans l'effet du SNP et d'utiliser ces matrices pour estimer les effets des SNP. Cependant, Zhou and Stephens, (2014) précisent bien que cette méthode d'analyse n'est pas adaptée à la détection de QTL avec de gros effets, qui entraîne une variation des matrices de VarCov, d'où l'utilisation du marqueur diagnostique de *Ppd-D1*. Finalement, le modèle est déclaré de la façon suivante :

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_n \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 1_n + \alpha_1 Ppd + \beta_1 x \\ \vdots \\ \mu_j 1_n + \alpha_j Ppd + \beta_j x \end{pmatrix}, \Sigma_G \otimes K + \Sigma_E \otimes I_n \right) \quad \text{Mod. [3]}$$

où les éléments de l'équation sont les mêmes que les éléments du Modèle [1] avec la matrice Σ_G qui est la matrice de variance covariance sélectionnée à partir du Modèle [1] et x qui est un vecteur ($n \times 1$) contenant les scores du SNP testé pour les n génotypes, et β_1, \dots, β_j qui sont les effets spécifiques du SNP pour chaque environnement. La significativité ($H_0 : \beta_1 = \dots = \beta_j = 0$) a été testée avec un test de Fisher. Ce modèle de GWAS multivarié a été développé sous le logiciel R par Willem Kruijer de l'université de Wageningen, il a été utilisé pour la première fois dans l'étude de Millet et al., (2016). Pour chaque trait, ce modèle de GWAS a été réalisé avec chacune des matrices de VarCov Σ_G et Σ_E qui ont été estimées à partir du Modèle [1]. Les $-\log_{10}(\text{P-value})$ sont représentés sous forme de QQ plot dans les annexes V.2. Le facteur d'inflation a été estimé à partir de ces QQ-plots, il est égal à la valeur de la pente de la droite de régression, avec un intercept fixé à zéro.

2.6.3 Multi-environnements multi-locus et par trait

Une procédure pas à pas, qui intègre le SNP avec la P-value la plus significative à chaque étape, a été réalisée à partir d'un sous-ensemble de SNP. Le modèle multi-environnements multi-locus se présente sous cette forme :

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_n \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_1 1_n + \alpha_1 Ppd + \beta_{11}x + \dots + \beta_{1L}x \\ \vdots \\ \mu_j 1_n + \alpha_j Ppd + \beta_{j1}x + \dots + \beta_{jL}x \end{pmatrix}, \Sigma_G \otimes K + \Sigma_E \otimes I_n \right) \quad \text{Mod. [4]}$$

où les éléments de l'équation sont les mêmes que les éléments du Modèle [3], avec un nombre de SNP allant de 1 à L. Le sous-ensemble de SNP testé est issu de la détection réalisée dans les deux étapes de

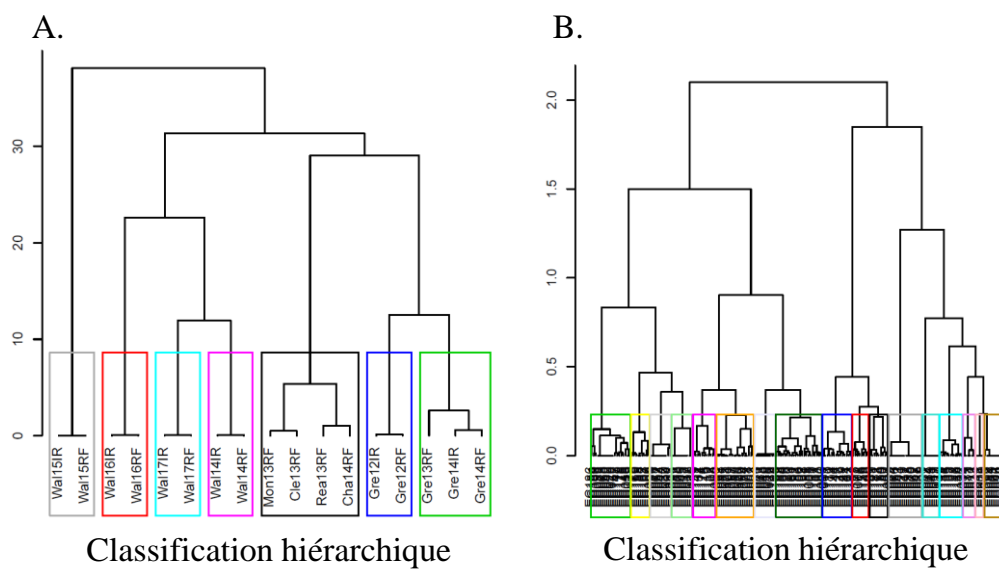


Figure V.2 : Distribution des covariables environnementales et clusterisation du MET

A. Dendrogramme sur la matrice D_{Ω} , les différents carrés représentent le regroupement des environnements proposé par la fonction *HCPC* du package R *FactoMineR*. B. Dendrogramme sur la matrice $D_{\Omega t}$, les différents carrés représentent le regroupement des covariables environnementales en 17 groupes.

GWAS précédentes. A la suite de ces deux étapes, les SNP ont en effet été regroupés en QTL, suivant la méthodologie présentée dans le Chapitre II et seulement le SNP avec la P-value la plus faible dans chaque QTL a été testé dans le modèle multi-environnements et multi-locus. L'incorporation de nouveaux SNP est arrêtée lorsqu'aucun SNP ne présente une P-Value < 0.001 (P-value issue d'un test de Wald). Comme pour le Modèle [4], pour réaliser cette régression pas à pas les matrices de VarCov sont issues du Modèle [1]. Cependant, une fois que le lot de SNP optimal a été identifié par la régression pas à pas, le Modèle [4] est ajusté une dernière fois en estimant les matrices de VarCov Σ_G et Σ_E .

De ce dernier modèle, seront extraits les effets des SNP spécifiques à chaque environnement ($\beta_{11}, \dots, \beta_{jL}$). Enfin, une corrélation de Spearman entre les effets des SNP et les covariables environnementales est testée pour essayer de relier la variation des effets des SNP aux covariables environnementales.

3 RESULTATS

3.1 Description du MET

Le jeu de données présente une large variabilité en termes de potentiel de rendement, allant d'une moyenne de 4.26 t/ha (Gre14RF) à plus de 9.67 t/ha (Wal16IR). De plus, les héritabilités au sens large présentent des valeurs importantes allant de 0.44 pour Gre14RF à plus de 0.90 pour Cha14RF, Wal16IR et Wal16RF (Table V.1). La Figure V.2 montre le regroupement des environnements en se basant sur les covariables environnementales. Les environnements d'une même localisation et d'une même année sont très proches. En effet, ces environnements ne sont différenciés que par une ou plusieurs irrigations (Annexe III.2.B). De plus, on peut voir que les expérimentations qui ont été réalisées à Gréoux-les-Bains (France) et Walla Walla (USA), les deux sites les plus représentés dans ce réseau d'expérimentation et les plus stressants d'un point de vue hydrique et thermique, sont bien regroupés. On peut noter que la distance entre les essais de Walla Walla en 2015 est aussi importante qu'avec les autres essais de Walla Walla et les essais de Gréoux ou d'autres sites français.

La Table V.3 présente l'ensemble des paramètres et index qui ont été calculés suite à l'application du Modèle [1] sur l'ensemble des traits avec différentes matrices de VarCov Σ_G . Pour chaque trait, six modèles ont été testés et un seul, indiqué par les nombres en gras dans la Table V.3, sera utilisé pour la suite des analyses. En regardant l'évolution du BIC en fonction de la complexité des matrices de VarCov, il apparaît clairement que les modèles ne prenant pas en compte de covariance génétique entre les environnements sont moins pertinents. Par exemple, le BIC du modèle DG pour le rendement s'élève à 13,567 alors que les BIC pour les modèles FA sont tous inférieurs à 13,000. Cependant, il semble important de regarder l'aspect visuel des QQ-plots issus du modèle de MET-GWAS en utilisant les différentes matrices de VarCov. Au final, le modèle FA4 a été sélectionné deux fois, le modèle FA3 deux fois et le modèle FA1 une seule fois.

Table V.3 : Comparaison des matrices des BIC pour choix des variance/covariances

<i>Trait</i>	<i>MODEL</i>	<i>N</i>	<i>J</i>	<i>DF</i>	<i>AIC</i>	<i>BIC</i>	<i>logREML</i>	<i>Apha</i>
<i>DOE</i>	ID	174	17	18	9481	9588	-4722	
	DG	174	17	34	9417	9620	-4675	
	FA1	174	17	51	5408	5713	-2653	1.31
	FA2	174	17	67	5197	5598	-2532	1.16
	FA3	174	17	82	5103	5594	-2470	1.13
	FA4	174	17	96	5014	5588	-2411	1.05
<i>SA</i>	ID	174	17	18	26539	26647	-13252	
	DG	174	17	34	26545	26748	-13238	
	FA1	174	17	51	25604	25909	-12751	1.61
	FA2	174	17	67	25513	25913	-12689	1.37
	FA3	174	17	82	25458	25949	-12647	1.39
	FA4	174	17	96	25427	26001	-12617	1.10
<i>GPS</i>	ID	174	17	18	12023	12131	-5993	
	DG	174	17	34	12037	12240	-5985	
	FA1	174	17	51	11110	11415	-5504	1.18
	FA2	174	17	67	11072	11473	-5469	1.59
	FA3	174	17	82	11036	11526	-5436	1.08
	FA4	174	17	96	10977	11551	-5392	2.25
<i>TKW</i>	ID	174	17	18	10003	10111	-4984	
	DG	174	17	34	10012	10215	-4972	
	FA1	174	17	51	7427	7732	-3663	1.23
	FA2	174	17	67	7249	7650	-3558	1.10
	FA3	174	17	82	7211	7702	-3524	0.97
	FA4	174	17	96	7193	7767	-3501	0.92
<i>GY</i>	ID	174	17	18	13373	13481	-6669	
	DG	174	17	34	13364	13567	-6648	
	FA1	174	17	51	12622	12927	-6260	1.86
	FA2	174	17	67	12308	12709	-6087	1.41
	FA3	174	17	82	12208	12699	-6022	1.02
	FA4	174	17	96	12190	12764	-5999	0.95

La [Figure V.3](#) montre les corrélations phénotypiques entre environnements pour le rendement, ainsi que les corrélations génétiques qui ont été extraites du meilleur Modèle [1]. La matrice de corrélation phénotypique a été triée en fonction des distances moyennes entre environnements ; le même classement a été utilisé pour représenter les corrélations génétiques. Les corrélations phénotypiques s'élèvent à plus de 0.79 entre Wall15RF et Wall16RF et des corrélations négatives (jusqu'à -0.43) peuvent être observées, notamment avec l'environnement Mon13RF qui est l'environnement le plus au nord de la France. De plus, il est apparu clairement que les expérimentations localisées à Gréoux-les-Bains et à Walla Walla sont groupées entre elles et définissent deux ensembles qui semblent différents des localisations plus au nord de la France. Il est aussi intéressant de noter que tous les environnements qui ne se différencient que par leur conduite de culture (c'est-à-dire irrigation) ne sont pas forcément regroupés ensemble, ce qui laisse penser que le traitement peut impacter le classement des variétés. La représentation des corrélations génétiques accentue la structuration en deux groupes : Gréoux-les-Bains et Walla Walla d'un côté, et les autres environnements de l'autre.

3.2 Sélection des covariables environnementales

L'ensemble des covariables qui ont été calculées sur le MET a été regroupé en 17 clusters (nombre de cluster équivalent au nombre d'environnements). Une visualisation de cette classification est présentée dans la [Figure 2B](#) et les résultats pour chaque EC sont présentés dans l'[Annexe V.1](#). Étant donné que le nombre d'environnements et leurs diversités sont assez limités dans ce MET, il semble important de regrouper les variables en groupes de variables, pour leur donner un sens qui se rapproche des conditions climatiques subies par les plantes. Par exemple, dans le groupe 1, il apparaît un nombre important de covariables en lien avec des températures faibles à des stades précoces de développement. Au contraire, il est plus compliqué d'attribuer une catégorie d'évènements climatiques dans le groupe 3, car ce groupe rassemble des covariables qui n'ont *a priori* aucune relation entre elles, par exemple le nombre de jours avec des températures gélives entre le semis et le stade épis 1 cm est corrélé positivement ($r=0.60$, $Pval<0.001$) à la somme des températures supérieures à 25°C pendant la méiose.

En plus d'un premier tri sur la similarité des EC, un autre tri a été réalisé pour essayer d'extraire et d'identifier les sous-ensembles d'EC qui permettraient de construire pour chaque trait une matrice de similarité entre environnement (Σ_{EC}), proche de la matrice des VarCov génétiques entre environnements (Σ_G). La [Figure V.3A](#) nous montre l'évolution de la corrélation entre les matrices Σ_G et Σ_{EC} pour le rendement. Ainsi, un sous-ensemble de 7 EC a été identifié et sélectionné pour le rendement. La corrélation entre les deux matrices (Σ_G et Σ_{EC}) atteint un score de 0.45 ($Pval<0.001$) et il est intéressant de noter qu'un nombre plus important d'EC pour construire la matrice Σ_{EC} diminue de façon claire la corrélation entre Σ_{EC} et Σ_G . Au final et dans l'ordre, les EC qui ont été sélectionnées pour le rendement sont EC125, EC053, EC119, EC137, EC058, EC161 et EC064. En les ajoutant au fur et à mesure, ces EC permettent de construire une matrice Σ_{EC} corrélée à 0.399, 0.408, 0.422, 0.425,

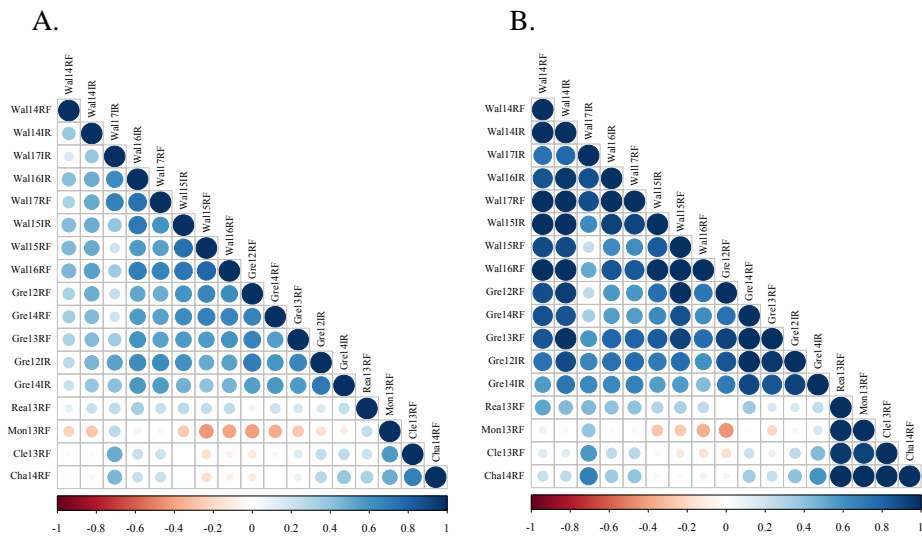


Figure V.3 : Corrélation phénotypique et génétique entre environnement
 A. Corrélation phénotypique. B. Corrélation génétique.

0.435, 0.441, 0.448, respectivement ([Figure V.3B](#)). Les évolutions des corrélations pour les autres traits sont présentées dans l'[Annexe V.3](#). La corrélation maximale entre Σ_G et Σ_{EC} a été identifiée pour la date de floraison, avec un $r=0.56$ ($P_{val}<0.001$). Au contraire, la corrélation minimale a été identifiée pour le nombre de grains par épis, avec un $r=0.32$ ($P_{val}<0.001$). Plusieurs covariables se retrouvent dans les sous-ensembles d'EC de plusieurs traits. Par exemple, la covariable EC181 est systématiquement première covariable pour les composantes du rendement. Malheureusement, cette variable est représentative du groupe de covariables 4, qui mélange plusieurs types de stress (stress azoté post-remplissage, déficit hydrique précoce, température maximale à la méiose). Il est donc difficile de conclure clairement sur l'impact de cette covariable sur les relations génétiques entre environnement. Au contraire, la covariable EC137 a été sélectionnée pour le TKW et le rendement. Cette covariable qui représente le groupe de covariables 5, qui groupe de nombreux indicateurs d'un stress température post-épiaison. Un stress thermique post-épiaison impacterait donc les relations génétiques entre environnements pour le TKW et pour le rendement.

3.3 GWAS et identification de QTLxE

Les deux détections de SNP sur l'ensemble du génome nous ont permis d'identifier 301 SNP associés avec au moins un trait en GWAS classique et/ou en GWAS multi-environnement. Par la suite, la sélection pas à pas *via* le modèle multi-locus multi-environnement (Modèle 4) nous a permis de sélectionner 10 SNP pour le rendement, 11 pour le TKW, 18 pour le nombre de grains par épi, 15 pour le nombre d'épis par m² et 8 pour la date d'épiaison. Pris tous ensemble, plus *Ppd-D1*, ils permettent d'expliquer 63%, 42%, 68%, 34% et 34% de la variance phénotypique de D.Z55, SA, GPS, TKW et GY, respectivement. Pris individuellement, la variance phénotypique expliquée par les SNP va de moins de 1% à plus de 14% pour le rendement ([Table V.4](#)). Seulement deux SNP ont été identifiés significativement associés avec la date de floraison dans tous les environnements. Sans surprise, il s'agit du SNP diagnostique de *Ppd-D1* et d'un SNP localisé sur le chromosome 2B qui est lié avec plusieurs marqueurs diagnostique de *Ppd-B1* ([Annexe V.5A](#)).

Si l'on se focalise sur le rendement, la [Table V.4](#) nous montre qu'aucun SNP n'est significativement associé au rendement dans tous les environnements. Le SNP qui se rapprocherait le plus d'un SNP à effet systémique est le SNP localisé sur le chromosome 4B. Il présente un effet positif pour l'allèle minoritaire dans 15 environnements sur 17 et explique 5% de la variance phénotypique. Il est aussi intéressant de noter que quatre SNP localisés sur les chromosomes 2D, 3A et 5A présentent un effet positif dans certains environnements et négatif dans d'autres. Ces SNP reflètent une source d'interaction GxE importante dans des environnements contrastés, car ils peuvent impacter positivement ou négativement le rendement. L'[Annexe V.4](#) présente les résultats issus du Modèle [4] pour chacun des traits. Le calcul des bornes de QTL par l'étendue du LD dans la zone permet d'identifier les potentielles colocalisations entre les composantes du rendement et le rendement. Seulement deux QTL présentent une colocalisation avec le TKW et le GPS : le QTL localisé sur le

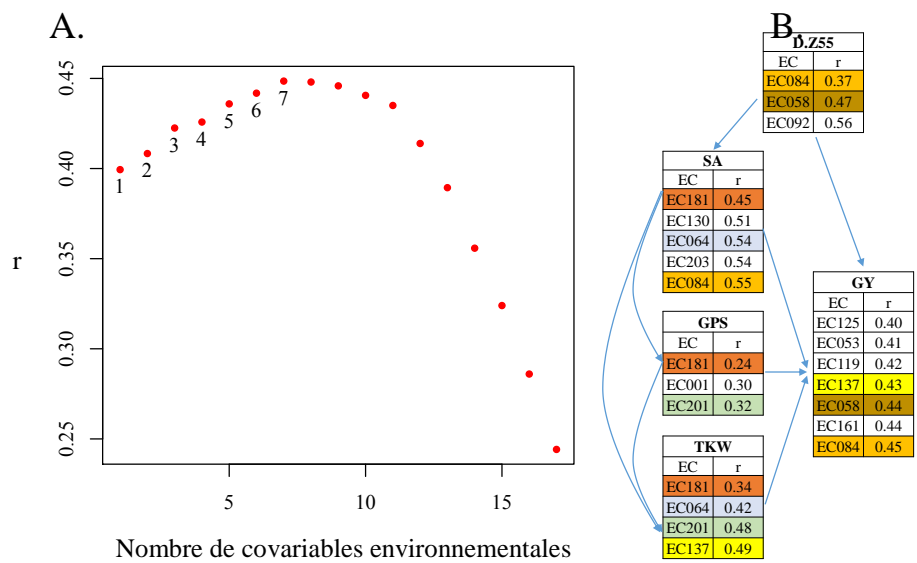


Figure V.4 : Corrélation entre la matrice de VarCov Σ_G et la matrice de distance Σ_{EC}

A. La figure présente l'évolution de la corrélation entre Σ_{EC} et Σ_G en fonction du nombre de covariables environnementales (EC) utilisées pour construire la matrice Σ_{EC} pour le rendement. B. La figure présente la sélection des EC pour chacune des composantes, les flèches présentent les relations de cause à effet entre le rendement et les composantes. Les EC qui ont été identifiées associées avec plus d'un trait sont en couleur.

chromosome 2A et le QTL localisé sur le chromosome 6B (SNP AX-89364136). Hormis Ppd-D1, il semble intéressant de noter qu'aucun des QTL associés avec une composante du rendement ou le rendement ne co-localise avec la date d'épiaison. Ce qui nous permet d'avancer qu'aucun QTL n'est associé à une stratégie d'évitement du stress par la plante.

La [Table V.4](#) présente les corrélations de Spearman entre les effets des SNP dans chaque environnement (β_j) et les valeurs du sous-ensemble de covariables environnementales (EC) préalablement sélectionnées. 19 SNP sur les 54 SNP associés au rendement ou ses composantes présentent une corrélation significative avec au moins une EC. Pour chaque trait, une représentation visuelle de la relation entre effet du SNP et la covariable environnementale est présentée dans la [Figure V.4](#). Pour réaliser cette figure, la plus forte corrélation par trait a été prise et il est possible de voir que pour les quatre traits, la relation entre β_j et EC semble linéaire. Cette relation nous permet de calculer le gain éventuel d'effet du SNP sur le trait en fonction de la valeur de l'EC. Par exemple, le SNP AX-89695675 localisé sur le chromosome 4B semble être un locus de réponse à un stress thermique. En effet, l'allèle favorable permet un gain de rendement de 4 kg/ha pour chaque °C au-dessus de 25°C pendant la phase de remplissage (entre Z55 et Z92) par rapport à l'allèle défavorable.

4 DISCUSSION

La détection de déterminants génétiques liés à la réponse aux variations environnementales (QTLxE) est une des pistes qui a été suivie par de nombreuses études pour aider les sélectionneurs à développer des outils pour adapter leur matériel à des conditions stressantes spécifiques. Cependant, très peu de ces études ont été réalisées avec des panels de diversité en prenant en compte la non-indépendance des environnements, surtout du fait de la durée des temps de calculs et des ressources qui doivent être mobilisées. Dans ce chapitre, nous présentons une suite d'analyses statistiques visant à estimer l'évolution des effets des QTL en fonction de la valeur des covariables environnementales à partir d'un MET. L'objectif étant de fournir aux sélectionneurs des outils (QTL) pour développer des variétés plus tolérantes à des stress ciblés.

4.1 Description du MET

L'utilisation de modèles mixtes, tels qu'ils ont été mis en place dans ce chapitre, nous permet de modéliser des relations génétiques entre environnements (matrice de variance/covariance) et de modéliser une structure d'erreur pour chaque environnement, tout en prenant en compte les relations entre individus (K). L'estimation de chacun des paramètres de la matrice de VarCov génétique peut être une procédure longue qui converge plus ou moins difficilement. Par exemple, l'algorithme d'estimation des paramètres du modèle non structuré ("unstructured"), qui n'impose aucune contrainte à la matrice de variance/covariance, n'a pas convergé après plusieurs heures de calcul. Plusieurs formes de VarCov génétiques plus contraintes ont donc été testées pour chacun des traits phénotypiques. Il s'avère que les formes les plus complexes de matrice de VarCov, c'est-à-dire des

Table V.4 : Description des SNP sélectionnés dans le modèle multi-environnement multi-locus (Mod [4]) pour le rendement et relation avec les covariables environnementales

Les positions des QTL ont été obtenues en ajoutant ou soustrayant l'étendue du LD dans la zone ; le LOG = $-\log_{10}(\text{Pvalue})$ et les effets sont issus du Modèle [4], le R² pour chaque SNP et le R² de la régression entre la matrice d'effet du SNP et la matrice des valeurs $Y_{ij}-\mu_j$; Le nombre d'environnements où le SNP a été identifié comme significativement associé au caractère (Pval<0.05) est présenté dans la table, ainsi que le minimum, maximum, moyenne de l'effet et le nombre de fois où le SNP a un effet positif ou négatif. Pour chaque covariable sélectionnée, le coefficient de régression de Spearman est représenté avec son niveau de significativité, ce coefficient est issu de la régression des effets des SNP dans chaque environnement et de la valeur des covariables dans chaque environnement.

SNP	QTL								Effets					Covariables Environnementales						
	Chr	Pos	Min	Max	MAF	LOG	R ²	Nenv	Min	Max	Moy.	Nb.<0	Nb.>0	EC053	EC058	EC084	EC119	EC125	EC161	EC137
PPD	2D	33,955,686	31,523,326	36,388,046	0.47	>15	0.13	15	-3.93	12.08	4.78	1	14	0.26	0.04	-0.23	-0.13	-0.49*	0.67**	0.44.
AX-89605455	2D	62,036,179	55,074,186	68,998,172	0.47	4.68	0.00	5	-1.92	1.75	-0.46	4	1	0.16	0.43.	0.04	0.56*	0.15	-0.1	-0.4
AX-89596791	2D	427,917,052	427,541,064	428,293,040	0.15	3.01	0.03	5	-4.61	1.96	-1.15	5	0	-0.13	0.25	-0.17	0.21	0.33	-0.17	-0.46.
AX-89309859	3A	1,034,036	428,240	1,639,832	0.10	3.69	0.02	5	-2.97	3.97	0.83	1	4	-0.14	-0.16	0.48.	0.17	0.53*	-0.14	0.05
AX-89442221	3A	20,792,908	20,176,745	21,409,071	0.12	3.38	0.10	5	-2.76	2.82	0.36	1	4	0.58*	0.55*	-0.33	0.08	-0.3	0.5*	-0.10
AX-89695675	4B	532,256,479	508,483,271	556,029,687	0.33	6.49	0.05	15	-0.11	6.31	3.31	0	15	-0.23	-0.4	0.00	-0.22	-0.36	0.51*	0.62**
AX-89654042	4D	4,075,309	4,075,309	4,075,309	0.34	4.71	0.03	3	-3.23	1.30	-0.55	3	0	-0.03	0.31	-0.12	0.34	-0.32	-0.09	-0.03
AX-89775787	5A	482,318,596	481,238,453	483,398,739	0.17	5.12	0.14	5	-3.15	4.35	1.07	1	4	0.31	0.21	-0.16	0.12	-0.31	0.36	0.21
AX-89563611	5D	389,597,594	389,475,330	389,719,858	0.17	6.58	0.05	6	-5.26	1.50	-1.59	6	0	0.44.	0.05	-0.01	0.01	-0.32	-0.04	0.25
AX-89730209	5D	504,510,030	502,999,179	506,020,881	0.40	4.19	0.01	7	-3.70	1.13	-1.16	7	0	0.28	0.09	0.14	0.28	-0.42.	0.08	0.19
AX-89505927	7A	85,431,910	83,513,483	87,350,337	0.22	5.11	0.01	3	-1.75	2.31	0.00	0	3	0.22	0.19	-0.61**	-0.41	-0.39	0.15	0.07

***, P-values <0.001 ; **, P-values <0.01 ; *, P-values <0.05 and ns non-significant P-value>0.05

matrices qui ont une variance génétique propre à chaque environnement et des covariances entre environnements, sont ceux qui permettent de construire le modèle le plus juste. C'est donc des modèles dits « Factor Analytic » qui ont été utilisés pour modéliser les VarCov génétiques. En effet, ces modèles permettent de construire une matrice de VarCov assez représentative de la réalité (variance et covariance hétérogène) tout en estimant un nombre de paramètres restreints qui permettent d'effectuer les calculs en un temps raisonnable (Boer et al., 2007; Meyer, 2009; van Eeuwijk et al., 2016).

Il semble important de tester plusieurs types de VarCov, car la diversité des relations génétiques à l'intérieur de la matrice de VarCov nous renvoie à l'importance de l'interaction GxE au sein du MET (van Eeuwijk et al., 2016). En effet, plus les corrélations génétiques sont importantes et moins il y a des chances d'observer des changements de classement des variétés entre les environnements (cross-over interactions). Si le GxE pour certains traits n'est pas très important une matrice de VarCov type DG peut être suffisante, ce qui permettrait de gagner du temps de calcul et de la puissance de test. Pour l'ensemble des traits, il semble que l'interaction GxE soit importante pour que les modèles FA soient sélectionnés ([Table V.3](#)).

La [Figure V.3B](#) montre que les corrélations génétiques pour le rendement sont assez étendues ($-0.4 < cor(g) < 1$). De plus, il apparaît clairement deux groupes dans les corrélations génétiques : (1) le plus petit des groupes est apparenté à des environnements de type « Centre ou Nord » de la France, avec Mon13RF qui est l'environnement avec la latitude la plus élevée et qui montre des corrélations génétiques négatives avec certains environnements de l'autre groupe. (2) le second groupe d'environnements est plus représentatif des conditions méditerranéennes françaises (présence des expérimentations réalisées à Gréoux-les-bains). Il faut aussi remarquer un regroupement des environnements issus de Gréoux-les-bains et Walla Walla, même si par exemple Wal15RF a une corrélation génétique plus importante avec Gre14RF qu'avec Wal15IR, ce qui montre une interaction Génotype-Irrigation importante. Au final, il faut retenir qu'il y a deux groupes dans ce MET, un groupe qui est assimilé à des conditions sèches méditerranéennes et un groupe plutôt optimal, ce qui n'a rien d'étonnant quand on connaît comment la sélection des environnements a été réalisée. Il est aussi intéressant de noter que dans ce MET, la présence prépondérante d'environnements de type méditerranéen est à l'opposé des MET qui ont été utilisés dans les études de Ly et al., (2018) et de Rincent et al. (en prep), où les essais de Gréoux-les-bains étaient minoritaires. Ce qui nous permet d'avancer que ce type de réseau expérimental semble être adapté à une exploration des comportements génétiques face à des conditions de sécheresse et de températures élevées (stress hydrique et thermique).

4.2 Sélection des covariables environnementales

Pour construire une matrice de similarité entre environnements proche de la matrice de VarCov génétique, il semble important de sélectionner un sous-set de covariables environnementales pour identifier les plus pertinentes ([Figure V.3](#)). Cette observation est en accord avec les résultats présentés dans Rincet et al. (en prep). Cette étude, intitulée “ *Using crop growth model stress covariates and AMMI decomposition to better predict genotype by environment interactions*”, est basée sur des approches de prédiction génomique des interactions GxE. De la même façon que les études de Jarquín et al., (2014) et Heslot et al., (2014), cette étude utilise les covariables environnementales pour construire une matrice de VarCov entre environnements, ce qui en théorie permettrait de prédire les performances de nouvelles variétés dans de nouveaux environnements. Le but de notre étude est sensiblement différent, dans le sens où nous ne voulons pas prédire les performances des variétés, mais identifier et estimer l’effet des déterminants génétiques. Dans ce chapitre, nous n’avons donc pas construit une matrice de similarité entre environnements pour prédire les interactions GxE, mais pour essayer de les reconstruire avec un minimum de variables environnementales, pour au final tester l’effet des SNP en fonction de ces variables qui ont été sélectionnées.

Cependant, le problème rencontré dans ce type de MET est la redondance des EC due au faible nombre d’environnements. Nous avons donc choisi de faire une première sélection des covariables environnementales selon la méthode présentée par Bouffier et al. (2015). Cette sélection nous permet d’identifier un nombre restreint d’EC représentatives de l’ensemble des EC. De plus, comme nous avons pu le voir, cette méthode de clusterisation des EC nous permet d’identifier des groupes d’EC et d’essayer de conclure quant à leurs effets sur le rendement. Par exemple, agronomiquement parlant, il est difficile d’imaginer que le nombre de jours de stress azoté entre Z55 et Z75 influence fortement le nombre d’épis par m² qui est censé être préalablement établi. Cependant, s’il l’on regarde ce que représente cette covariable (EC130), il apparaît qu’elle est l’EC représentative du cluster de covariable numéro 12. Ce cluster regroupe de nombreuses covariables de stress hydrique et azoté en phase post épiaison qui ont été calculées grâce au modèle de culture CHN. Il est évident que ce type de stress fort n’est pas arrivé d’un coup à épiaison, mais est la résultante de l’installation d’un stress hydrique ou d’un stress de nutrition azotée plus précoce.

4.3 Identification des QTLxE et relation avec les covariables environnementales

Les QQ-plot du modèle multi-environnemental (Mod [3], [Annexe V.2](#)) nous montrent que la prise en compte et le choix des matrices de VarCov génétiques entre environnements est primordial pour limiter l’inflation des Pval observées, c’est-à-dire pour limiter la détection de faux positifs (erreur de Type D). La plupart des études comparent et sélectionnent l’ordre des modèles FA en minimisant le BIC (Boer et al., 2007; van Eeuwijk et al., 2016). Comme nous avons vu précédemment, la comparaison du BIC entre les matrices de VarCov type ID ou DG et les matrices FA permet de choisir facilement la plus adaptée, cependant, le choix de l’ordre du modèle FA reste plus compliqué. C’est

pourquoi, il nous paraît important d'observer le comportement des détections de QTL issus de ces modèles. Pour le rendement, par exemple, l'inflation des Pvalues du modèle avec une matrice de VarCov FA1 ou FA2 (qui sont les moins complexes des matrices FA) nous montre une déviation très importante par rapport à la droite $X=Y$. Il semble clair qu'il faille écarter les modèles avec FA1 et FA2, cependant le choix entre FA3 et FA4 peut être arbitraire. Il faut donc tenter de conserver un équilibre entre la diminution de l'inflation des Pvalues et la sur-structuration du modèle pour conserver des chances de détecter quelques QTL (communication de Willem Kruijer lors de mon stage à WUR).

La mise en place d'une étape de validation des QTL *via* une procédure pas à pas (stepwise forward) est complètement liée à la problématique présentée précédemment. En effet, le but de notre détection de QTL mono ou multi-environnementaux était de réaliser un premier tri dans les SNP pour réaliser ensuite une sélection avec une procédure pas à pas. Ce type de procédure "stepwise" a été largement utilisée pour sélectionner des sous-sets de QTL et limiter les erreurs de Type 1 (Cordell and Clayton, 2002; Segura et al., 2012; Millet et al., 2016; Li et al., 2017). Cependant, pour faire de la sélection de variables, les régressions "stepwise" ne semblent pas faire l'unanimité chez les statisticiens (Flom et al., 2007). Ceci pour de multiples raisons, dont le problème de la colinéarité entre les variables à sélectionner et de la multiplication des tests. Dans la littérature, de nombreuses approches ont été présentées pour régler ces problèmes, par exemple, la ridge regression ou LASSO (Hoggart et al., 2008; Malo et al., 2008; Cho et al., 2010). Dans tous les cas, ces approches (stepwise comprise) sont plus performantes qu'une simple détection mono-locus (Ayers and Cordell, 2010). Cependant, elles peuvent être compliquées à mettre en place en GWAS, du fait de la prise en compte de la structuration des données (matrice K). Les différentes procédures que nous présentons dans ce chapitre sont donc une première étape vers une meilleure détection de QTL, mais elles peuvent être encore améliorées.

Seulement deux QTL ont un effet significatif dans l'ensemble des environnements. Ces QTL sont associés avec la date d'épiaison et localisés sur les chromosomes 2B et 2D. Ils co-localisent avec les gènes bien connus de précocité *Ppd-D1* et *Ppd-B1*. Il est intéressant de noter que si l'effet de *Ppd-D1* a largement été identifié dans les chapitres III et IV et des études de GWAS qui ont été réalisées sur un panel équivalent au panel élite (Le Couvreur et al., 2011; Cormier et al., 2014), il n'en est pas de même pour l'effet de *Ppd-B1*. En effet, ce QTL n'a été identifié associé à la précocité que dans le scénario de condition hydrique optimale dans le chapitre III. Ce type de résultat, en accord avec les nombreuses études d'association multivariées sur des phénotypes corrélés réalisées (Korte et al., 2012; Stephens, 2013; Zhou and Stephens, 2014), permet de formuler l'hypothèse que la prise en compte de l'ensemble des données dans un modèle multi-varié (multi-environnemental) augmente la puissance de détection de QTL, qui pourraient avoir un effet plus faible mais plus robuste. Ce type de résultat nous conforte donc dans notre stratégie de détection de QTL *via* un modèle multi environnemental.

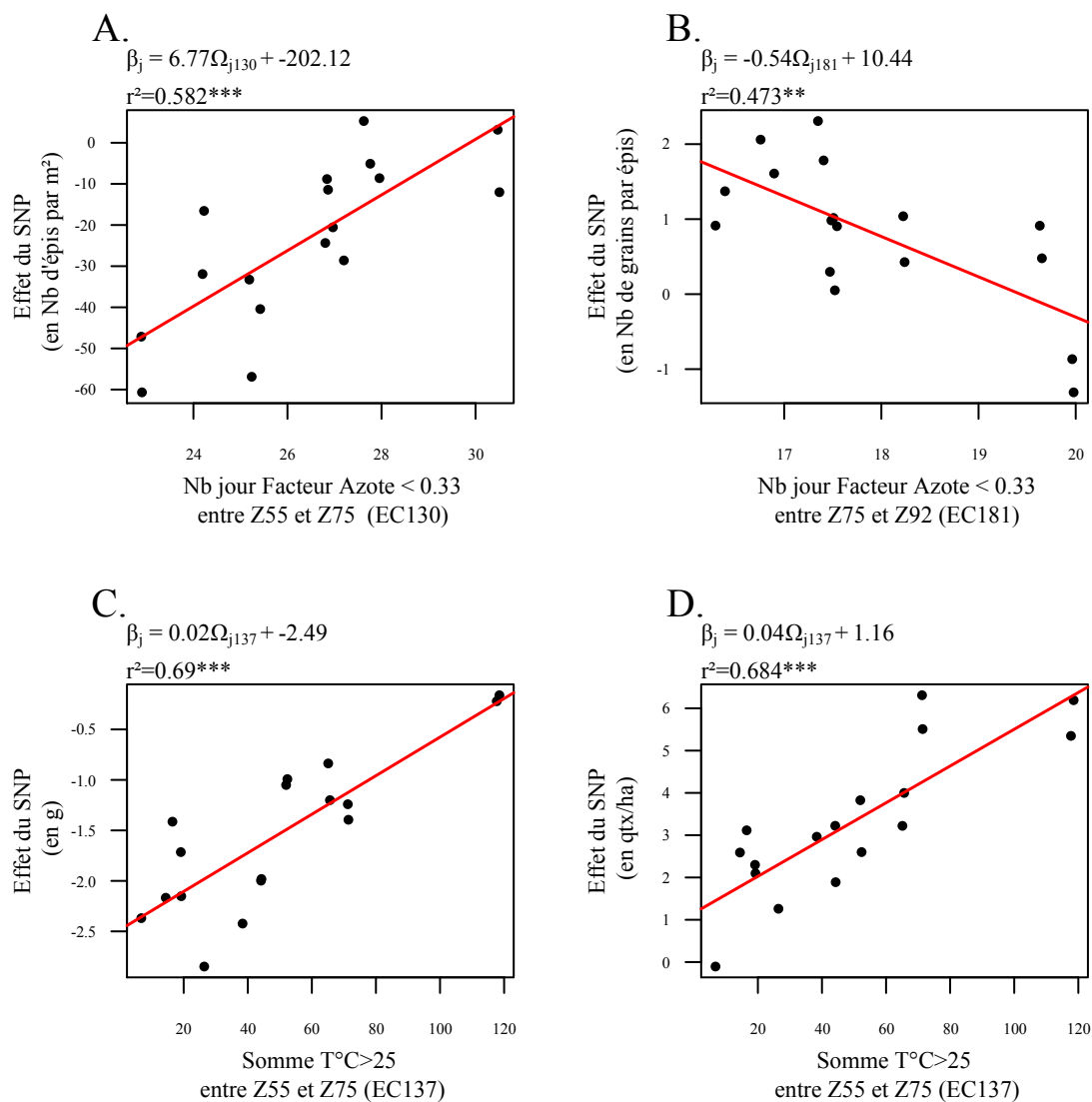


Figure V.5 : Visualisation de l'effet des SNP en fonction de l'évolution de covariables environnementales dans le MET

Pour illustrer l'évolution des effets des SNP vs EC, le couple qui présente les plus fortes corrélations pour chaque trait est présenté dans ces figures. Les figures A, B, C et D présentent donc la relation effet du SNP en fonction de la valeur d'une EC pour le nombre d'épis par m², nombre de grains par épi, le TKW et le rendement en grain, respectivement.
 ***, *P-values* < 0.001 ; **, *P-values* < 0.01

Les autres QTL présentent donc une instabilité d'un environnement à l'autre. Cette observation est en accord avec de nombreuses études qui ont été réalisées sur le blé dans des MET et les chapitres précédents (Malosetti et al., 2008; Bennett et al., 2012; Cormier et al., 2014). Cependant, l'objectif de ce chapitre n'était pas de montrer la présence de QTL systémiques pour le rendement, mais d'essayer d'identifier des QTL qui répondraient aux variations des conditions pédoclimatiques. Après la détection de QTLxE, nous avons donc choisi de confronter les effets de ces QTL à la fluctuation de variables environnementales. Il aurait été possible de tester ces variations directement dans le Modèle [4], sous forme de régression factorielle (van Eeuwijk et al., 2010). Cependant, nous sommes encore une fois confrontés au problème de temps de calcul, et surtout biologiquement parlant, il semble peu probable que la relation entre l'effet d'un QTL et une covariable environnementale soit totalement linéaire. C'est pourquoi nous avons choisi d'extraire les effets des QTL et de tester leurs corrélations avec le sous-set de variables environnementales précédemment sélectionné, grâce au coefficient de Spearman. Les résultats de cette prospection ([Table V.4](#) et [Annexe V.4](#)) nous montrent que tous les SNP ne répondent pas aux covariables que nous avons sélectionnées. De plus, la [Figure V.5](#) montre la relation entre les SNP et les covariables qui sont les plus corrélés, il semble que pour ces cas-là, les relations soient linéaires.

A ce niveau d'avancement de notre étude, il faut bien se rappeler que les corrélations entre les effets des QTL et les covariables environnementales ne sont pas des relations de cause à effet et qu'il serait extrêmement imprudent de conclure que le QTL sur le 4B, par exemple, est bien un QTL de réponse au stress thermique post-épiaison. Il aurait été très intéressant que ce QTL soit aussi identifié comme étant associé aux composantes du rendement comme le nombre de grains par épi ou le poids de mille grains, ou encore qu'il soit identifié dans l'étude présentée dans le chapitre IV. Cependant, ce n'est pas le cas, mais de nombreuses études ont identifié des QTL de réponse au stress thermique sur le chromosome 4B ([Annexe I.1](#)).

CONCLUSION

La détection de QTL de réponse aux variations de l'environnement est un enjeu majeur de l'adaptation de nouvelles variétés à de nouvelles conditions de cultures, qui semblent devenir de moins en moins favorables du fait d'évènements extrêmes (sécheresse, vague de chaleur...). L'approche qui est présentée dans ce chapitre tente donc de détecter ces QTL qui répondent aux variations environnementales que l'on peut retrouver dans le réseau d'essais ciblé stress hydriques et/ou thermiques.

Nous avons vu que l'approche de détection utilisant des modèles de GWAS multi-environnementaux semble permettre un gain de puissance de détection. Par exemple, elle permet la détection de l'effet systémique du QTL qui co-localise avec *Ppd-B1* sur la précocité. Il s'agit d'un QTL qui a largement été décrit dans la littérature mais que nous n'avons détecté que dans un type d'environnement dans le chapitre III. La suite de l'approche, qui a été présentée dans cette étude, s'est employée à expliquer l'évolution de l'effet des QTL au regard de covariables environnementales. Plusieurs corrélations fortes et significatives ont été identifiées, par exemple, le QTL localisé sur le chromosome 4B qui semble permettre un gain de $4\text{kg}\cdot\text{ha}^{-1}$ pour chaque $^{\circ}\text{C}$ au-dessus de 25°C entre la date d'épiaison et la récolte. Cependant, à ce stade, des études complémentaires semblent nécessaires afin de conclure plus précisément sur les relations de cause à effet entre covariables environnementales et l'effet de ce QTL.

Une prospection plus large des relations entre les composantes du rendement et le rendement pourrait être réalisée via la mise en place d'équations structurées prenant en compte, dans un modèle multi-trait, les relations de cause à effet entre les différentes composantes.

Discussion Générale et Perspectives

Discussion Générale et Perspectives

Cette thèse a été réalisée en partenariat entre Arvalis, Biogemma et l'INRA et elle s'inscrit dans le cadre de la création et de la préconisation variétale dans un contexte de changement climatique. L'identification de déterminants génétiques, liés à la tolérance aux stress hydriques et thermiques, est un des outils qui devrait permettre d'améliorer les connaissances sur les performances des variétés dans des conditions de cultures stressantes et permettre aussi de mettre à disposition des sélectionneurs des outils pour augmenter les performances des nouvelles variétés. En introduction, nous avons présenté les différents objectifs qui avaient été fixés lors de la rédaction du projet de thèse. La question que nous allons traiter dans cette partie est de savoir dans quelle mesure nous avons atteint ces objectifs, et de discuter des limites et perspectives de ce travail.

Comment le matériel qui est à notre disposition peut-il répondre à nos objectifs ?

Le deuxième chapitre de cette thèse présente le panel de variétés élites et les ressources moléculaires qui sont à notre disposition. Ce chapitre n'avait pas vocation à faire une étude complète de la diversité du panel élite car ceci a largement été réalisé dans des études précédentes sur un panel équivalent (Le Couviour et al., 2011; Cormier et al., 2014). La question que nous devons nous poser est de savoir s'il est vraiment intéressant de chercher à identifier des QTL majeurs dans un panel élite ?

Connaissant les processus de sélection et d'inscription des variétés et la variabilité interannuelle des événements climatiques comme la sécheresse et le stress thermique, il semblerait normal que les variétés extrêmement sensibles à ces stress ne soient pas présentes dans notre panel. En effet, nous pouvons penser qu'au cours du processus de sélection et d'inscription elles aient été confrontées à ces stress et donc éliminées. Ainsi, contrairement à des populations bi-parentales où les parents ont été choisis spécifiquement pour leur tolérance ou sensibilité aux traits étudiés, il n'est pas certain qu'une grande gamme de comportements existe dans du matériel inscrit. Les variétés qui sont à notre disposition sont déjà des variétés très performantes, il est donc attendu qu'aucun QTL fort ne soit détecté, c'est-à-dire des QTL qui expliquent une part importante de la variance phénotypique (plusieurs dizaines de pourcents). Cependant, aucune étude n'avait encore exploré la variabilité de ce panel face aux stress hydriques et/ou thermiques, et une partie du travail qui a été engagé dans cette thèse doit permettre de répondre à cette question.

De plus, les travaux qui ont été réalisés dans cette thèse s'inscrivent dans un programme de recherche bien plus large et les QTL qui ont été identifiés (même s'ils ne permettraient pas un bond dans l'amélioration variétale) pourront être valorisés au regard d'autres QTL identifiés dans la littérature ou dans d'autres projets. Par exemple, dans le cadre du projet BREEDWHEAT, le volet 3 a sélectionné un panel de diversité avec des variétés de type hiver issues de la core-collection développée par l'INRA. Le but de ce panel est d'identifier de nouvelles sources de QTL qui ne sont pas présentes dans le matériel européen et qui pourront être introgressées. Ces résultats confrontés aux nôtres seront sans

aucun doute utiles pour guider l'introggression de nouveaux variants alléliques dans le germoplasme élite.

Les ressources moléculaires sont-elles suffisantes ?

La disponibilité des marqueurs moléculaires, via l'essor de la découverte de variants alléliques (nouvelles techniques de séquençage) et une réduction des coûts de génotypage des SNP est l'une des grandes avancées de la recherche en génétique végétale ces 15 dernières années. Aujourd'hui, notre étude a jouté de l'accès à plusieurs centaines de milliers de marqueurs moléculaires (Rimbert et al., 2018), via l'utilisation d'une puce Affimetrix Axiom. Il est sans doute nécessaire de se demander si ce nombre de marqueurs et leur répartition sur le génome sont suffisants pour mener une analyse d'association efficace. Dans le chapitre II nous avons vu que les marqueurs recouvraient la quasi-totalité du génome. Il faut aussi noter que tous les marqueurs moléculaires disponibles n'ont pas été pris en compte dans les études de GWAS, seulement les marqueurs polymorphes de bonne qualité (Polymorphic High Resolution). Deux raisons majeures sont à l'origine de cette décision : (1) seuls les marqueurs bi-alléliques avec un nombre limité de données manquantes ont été utilisés, cette décision est largement guidée par l'utilisation d'outils qui ne traitent que ce type de marqueurs, par exemple, pour l'imputation des données manquantes et l'utilisation d'outils statistiques qui ne prennent en compte que des marqueurs bi-alléliques (pour des raisons de temps de calculs et de comparaison de résultats entre marqueurs moléculaires). (2) la seconde raison est liée à la redondance d'information générée par les 180,000 marqueurs sur un panel de 220 variétés. En effet, l'étude du LD qui a été réalisée a démontré qu'environ 20,000 marqueurs (nombre de clusters de SNP avec une distance < 0.1 dont la distance $= 1 - R^2$) suffisent pour couvrir chaque bloc haplotypique avec au moins un marqueur.

De mon point de vue, l'utilisation de toutes les ressources moléculaires disponibles (280,000 marqueurs initialement) n'est pas vraiment nécessaire pour une première prospection de GWAS. En effet, les 180,000 marqueurs SNP génotypés sur nos 220 variétés, finalement très proches génétiquement, car il s'agit d'un panel de diversité européen, se sont avérés largement suffisants pour une première exploration de QTL. Cependant, ce point de vue n'est pas toujours partagé, notamment par d'autres études et projets sur ce même panel qui visent à réaliser un enrichissement des régions en marqueurs moléculaires pour tenter de réduire l'intervalle du QTL et à terme d'identifier le marqueur causal. Etant donné la densité de marqueurs déjà disponible (plusieurs SNP par bloc haplotypique), il me semble moins pertinent de densifier des zones du génome associées en vue de réduire la taille des QTL. En tous cas, avec les ressources que nous avons utilisées et notre façon de définir nos QTL, augmenter le nombre de marqueurs génotypés sur le panel ne me semble intéressant que si cela permet de décrire une diversité qui est différente de celle qui est déjà à notre disposition, comme par exemple des marqueurs PAV (Presence Absence Variation) et CNV (Copy Number Variation). Pour illustrer cette remarque, il est intéressant de regarder l'étude de Cormier et al., (2014) qui a estimé que densifier des QTL en marqueurs et déterminer le gène causal comme étant le gène qui a le plus de

SNP associés ou les plus forte P-value ne marche que deux fois sur trois. Ce qui est déjà très bien, cependant, de nombreuses ressources sont mises en œuvre pour arriver à ces résultats et nous pouvons nous demander si d'autres méthodes ne sont pas plus pertinentes ?

Il est aussi possible d'illustrer cette remarque avec les résultats d'association qui ont été obtenus entre le gène *Ppd-D1* et la précocité du panel. En effet, en connaissant la position du gène *Ppd-D1* sur le génome et en regardant les SNP qui sont le plus fortement associés avec les dates de floraison ou d'épiaison dans notre panel, il apparaît que les deux SNP sont en total LD. L'un des SNP est dans le gène *Ppd-D1*, alors que le second est positionné à plus de 30 gènes de *Ppd-D1* (0.7Mb). Il est donc clair que si ce gène n'avait pas été clairement identifié dans la bibliographie, il aurait été difficile de l'identifier grâce aux méthodes de GWAS que nous avons utilisé dans cette étude. Même si elle peut largement y contribuer, l'identification de gènes candidats en se basant uniquement sur des résultats d'association n'est donc pas suffisante. Le recours à des populations biparentales semble très complémentaire pour identifier des recombinants et réduire la taille des intervalles de confiance des QTL. Enfin, la création de matériel isogénique, l'utilisation de la transgénése, ou de mutagénèse induite (Tilling, CRISPR) sont des approches nécessaires pour valider des gènes candidats.

Les résultats sont-ils en accord avec les objectifs qui avaient été déterminés ?

Le stress hydrique

Le premier article de cette thèse (Chapitre III) a été réalisé pour avoir un point de vue global de l'impact d'évènements de stress hydrique sur le réseau d'expérimentation qui était à notre disposition. Pour réaliser cette caractérisation environnementale, le modèle de culture CHN a été utilisé. Il nous a aussi permis de simuler les fréquences des stress hydriques à l'échelle de la France.

Les résultats de ce chapitre peuvent être résumés en quatre parties.

- (1) Comme il a été montré dans plusieurs études (Chapman, 2008; Chenu et al., 2011), l'utilisation de modèles de culture et le regroupement d'environnements en fonction du scénario hydrique subi par les plantes aident à expliquer une partie significative de l'interaction génotype environnement (20%). De plus, nous avons vu que les scénarios hydriques dans le réseau expérimental étaient assez représentatifs des scénarios de stress qui ont pu impacter les cultures de blé à l'échelle de la France ces vingt dernières années. Pour rappel, sur l'ensemble du territoire français, il y a environ une chance sur deux (moyenne sur la simulation des 20 dernières années) que les cultures de blé subissent un stress hydrique qui impacterait le rendement ([Annexe III.3D](#)).
- (2) Classification des variétés au regard de leur performance dans différents scénarios de stress hydriques.
- (3) La composition de notre panel de variétés nous a aussi permis d'évaluer le gain génétique dans différents scénarios hydriques. Il apparaît que le gain génétique le plus important est

trouvé dans les scénarios hydriques non stressants. Connaissant l'organisation de la sélection des variétés de blé en France, il n'est pas surprenant que le gain génétique le plus grand soit observé dans un scénario hydrique optimal. En effet, les stations de sélection sont basées majoritairement dans la moitié nord de la France (où les stress hydriques sont peu fréquents) et les environnements ciblés par les sélectionneurs sont les grands bassins de production de blé en France (exemple avec la Beauce) où les stress hydriques sont aussi peu fréquents. Nos résultats semblent donc être en accord avec la pression de sélection exercée en France. Cependant, nous avons aussi observé que le progrès génétique était moins important (voire non significatif pour le scénario de stress tardif) dans les environnements qui présentent des scénarios hydriques stressants. Il semble donc que la sélection indirecte (c'est-à-dire la sélection basée sur le rendement dans un réseau d'essais multi-locaux) ne soit pas optimale pour augmenter le gain génétique dans des scénarios de stress hydriques.) ne soit pas optimale pour augmenter le gain génétique dans des scénarios de stress hydriques.

- (4) L'identification de QTL ayant un effet positif sur le rendement dans les scénarios de stress environnementaux pourrait permettre à la sélection d'augmenter le progrès génétique dans les environnements correspondant.

Si les résultats présentés dans le chapitre III nous permettent de montrer l'intérêt d'améliorer les variétés pour leur tolérance aux stress hydriques et nous montrent quelques pistes pour y parvenir (QTL), certaines limites de notre jeu de données et de notre stratégie sont à souligner :

- Comme il a été largement montré, les modèles de culture permettent une finesse dans le calcul des stress hydriques et permettent de prendre en compte une part importante de l'interaction GxE (Chapman, 2008; Chenu et al., 2011). Cependant, l'une des limites majeures à cette stratégie est le manque de données qui auraient pu nous aider à bien calibrer le modèle, par exemple des biomasses à floraison et une caractérisation précise des sols (profondeur, données tensiométriques...). Ce type de calibration a été réalisé sur certains lieux lorsque nous avons les données (notamment les essais réalisés par Arvalis et à Walla Walla), mais pour la plupart des essais, cela n'a pas été possible.
- La réalisation d'une étude des fréquences de stress en France nous a permis de replacer notre étude dans le contexte et de conclure à une bonne représentativité de notre jeu de données. Cependant, l'analyse a été réalisée en considérant de façon équivalente chaque combinaison station météo/sol, alors qu'en réalité chaque site ne représente pas les mêmes volumes de production de blé. Pour réellement représenter l'occurrence des stress hydriques à l'échelle de la France, il serait intéressant de pondérer chaque station par rapport au volume de production de blé qu'elle représente (Chenu et al., 2011).
- Le panel élite présente une variabilité pour la date de floraison qui complique énormément l'interprétation des résultats, car les variétés n'ont pas été confrontées tout à fait au même

stress en fonction de leur précocité. Nous avons essayé de prendre en compte ces écarts de précocité en utilisant différentes stratégies (colocalisation QTL de précocité, gènes majeurs liés à la précocité utilisée en covariables ...), cependant, certaines variétés ont pu plus ou moins échapper au stress.

- Le panel n'est pas spécifiquement adapté à l'étude l'estimation du progrès génétique. Suivant les propositions de Cormier et al., (2013), la précocité du matériel a été corrigée pour estimer le progrès génétique dans chaque type de scénario hydrique. La présence d'un progrès génétique avec une pente quasiment aussi forte en condition optimale qu'en condition très fortement stressée n'était pas vraiment attendue. Il peut y avoir un biais dans l'analyse, peut-être parce que la précision d'estimation des valeurs génétiques est un peu plus faible en condition stressante par rapport aux conditions optimales (cf. héritabilité).

Hormis les QTL de floraison (notamment *Ppd-D1*), peu de QTL à fort effet ont été observés dans cette partie de l'étude. Ceci m'a conduit à faire une prospection large de QTL, en acceptant le risque d'identifier des faux positifs. Cette stratégie peut être largement remise en question si l'on ne regarde que le chapitre III, mais étant donné la difficulté d'identifier des QTL à fort effets dans cette étude, il me semble important d'explorer un maximum d'informations pour les confronter aux résultats des chapitres suivants. En effet, le but de cette stratégie était de poser dans ce chapitre des bases larges sur lesquelles éventuellement s'appuyer pour valider, ou non, les travaux des chapitres suivants. De plus, un faible effet sur un trait peut avoir un effet fort sur une autre composante. Il nous semblait important d'essayer de mettre en relation les QTL détectés sur les composantes du rendement avec les QTL de rendement. Nous voulions aussi éviter d'éliminer un QTL qui aurait un effet faible sur la floraison, mais un effet fort sur le rendement. Grâce à l'analyse des colocations avec un seuil faible, nous pouvons identifier des QTL qui serait lié à la précocité et donc plus des QTL d'échappement que des QTL de tolérance.

Le stress thermique

Le Chapitre IV est une approche diamétralement différente de celle mise en place dans le Chapitre III. En effet, dans le Chapitre III, certaines expérimentations n'étaient pas dédiées spécifiquement au stress hydrique et l'objectif était de prospecter un maximum d'environnements, pour être le plus possible représentatif des fréquences de stress hydriques en France.

Les expériences et le plan d'expérimentation du Chapitre IV ont, quant à eux, été spécifiquement réalisés pour identifier les déterminants génétiques de la mise en place du poids des grains en conditions de stress thermique post-floraison et sans stress hydrique. La comparaison de matériel en présence de stress thermique et en absence de stress hydrique thermique semble difficile (voire impossible) à réaliser au champ. Dans ce sens, cette étude est l'une des toutes premières réalisée sur

du matériel européen. Celle-ci nous a permis d'explorer la réponse de déterminants génétiques à un stress thermique spécifique.

La méthodologie statistique de détection de QTL qui a été mise en place est strictement adaptée à cette expérience, de façon à conserver un maximum de puissance de détection de QTL. Le but du modèle est de tester les effets principaux des SNP et interactions SNP x traitement. De plus, un suivi de l'évolution de la sénescence a été réalisé. En effet, des températures hautes peuvent avoir un effet délétère sur la chlorophylle et la sénescence, il est donc important de se munir d'outils physiologiques pouvant déterminer les causes de diminution du remplissage du grain. Un atout essentiel de cette expérience est que contrairement au Chapitre III, chaque plante a été exposée au même stress au même stade, et il n'y a donc pas de phénomène d'échappement. En contrepartie, nous ne sommes pas sûrs que les plantes se comportent de façon strictement équivalente à une expérimentation au champ (différence de densité, différence d'architecture racinaire car les plantes sont en pot de 1L). Plusieurs résultats, qui pourront être exploités par la suite, ont été produits dans ce chapitre.

- (1) L'exploration de la diversité génétique de réponse au stress thermique nous a permis d'établir pour chaque variété un index de tolérance au stress, donc une classification des variétés au regard de leur tolérance au stress thermique post-floraison.
- (2) Le suivi de la sénescence des plantes nous a permis de voir qu'il existe une variabilité génétique au sein du panel élite pour le début et la vitesse de sénescence. De plus, certains QTL qui impactent significativement ces paramètres de sénescence ont été détectés.
- (3) Plusieurs QTL ont été identifiés en lien avec la réponse au stress thermique sur le remplissage du grain. Ils sont positionnés sur les chromosomes 4B, 6B et 3B.

Cependant, plusieurs prospections restent à venir pour essayer de construire un schéma des déterminants génétiques de la tolérance au stress thermique. Dans ce sens, une approche gène candidat a été réalisée, mais non présentée dans ce manuscrit. Cette approche a été conduite en deux étapes, la première étape étant la réalisation d'une étude transcriptomique de l'expression des gènes à différents moments du remplissage en conditions stressante et en conditions non-stressante. Malheureusement, aucun des gènes différentiellement exprimés en réponse à l'élévation de la température ne co-localise avec les QTL que nous avons identifiés. Plusieurs raisons peuvent être avancées, d'une part l'étude transcriptomique a été réalisée uniquement sur quatre variétés, qui sont donc loin de représenter la diversité de réponse qui peut être identifiée dans le panel élite. D'autre part le niveau de transcription des gènes n'est pas toujours représentatif de l'effet des gènes sur le phénotype, du fait de possibles modifications post-transcriptionnelles et d'interactions complexes protéine-protéine. La seconde étape a été de confronter les gènes décrits sous les QTL, grâce aux annotations qui ont été réalisées dans le cadre de la publication de la séquence de référence du blé tendre (IWGSC et al. 2018), à la littérature en cherchant des orthologues, paralogues et homéologues. Le nombre de gènes sous les QTL de cette

étude varient d'une dizaine à plus de soixante-dix ce qui peut compliquer l'identification de gènes candidats. Aucun gène, dont la fonction est associée dans la littérature à des réponses au stress thermique, n'a été identifié sous les QTL 4B et 6B (QTL pour lesquels le nombre de gènes candidats est le plus petit).

Certains de ces QTL ont aussi été identifiés dans le chapitre III en situation de stress hydrique, ce qui nous conforte dans le choix de réaliser une prospection large de QTL sur des essais de plein champ (Chapitre III) et une détection précise de QTL dans des conditions stressantes contrôlées en serre (Chapitre IV). Cependant, nous pouvons nous poser la question de la pertinence de l'utilisation de ces QTL dans un programme de sélection, dans l'état actuel de nos connaissances. Il était donc nécessaire d'aller plus loin dans la prospection des réponses des QTL à des stress.

La prospection multi-stress

Après avoir tenté d'explorer le comportement du panel élite face à un stress hydrique et un stress thermique indépendamment, nous nous sommes donc employés à faire une description multi-stress dans un contexte d'essais multi-environnementaux. Ce chapitre est bien plus proche de la réalité, dans le sens où au champ, les plantes sont confrontées à de multiples stress. Il est donc pertinent d'explorer cette piste, même si d'un point de vue statistique les analyses sont plus complexes. De la même manière que dans les chapitres précédents, plusieurs résultats sont à souligner :

- (1) La première étape de ce chapitre a été de sélectionner un sous-set d'environnements qui représenterait un maximum de scénarios hydriques et thermiques, ceci sans sur-représenter le scénario optimal comme c'est le cas dans le chapitre III. Comme nous avons vu grâce aux relations génétiques entre environnements, le sous-set d'environnements qui a été sélectionné est complémentaire du réseau d'expérimentation qui a été utilisé dans l'étude de Ly et al. (2018). Cette sélection d'environnements apporte donc une diversité de conditions stressantes qui n'avait pas été explorée jusqu'à maintenant.
- (2) Les analyses statistiques qui ont été mises en œuvre pour la détection de QTLxE nous ont permis d'augmenter la puissance de détections des QTL. Le fait de prendre en compte plusieurs environnements dans le même modèle, tout en prenant en compte la relation entre environnements, nous permet d'augmenter considérablement le nombre de degrés de liberté et donc permet un gain de puissance.
- (3) Plusieurs QTL avec des effets contrastés entre environnements ont été détectés et nous avons pu relier ces effets à l'évolution de variables environnementales. Par exemple, le QTL localisé sur le chromosome 4B a montré des effets nuls ou positifs pour le rendement. Ces effets sont d'autant plus favorables pour le rendement que les températures stressantes entre épiaison et récolte sont élevées.

Cependant, si des QTL répondent différemment entre les environnements, il est compliqué de relier cette évolution des variations environnementales et de facteurs de stress. De plus, il est compliqué de mettre en relation les effets des QTL sur le rendement avec les composantes de rendement. Les résultats de ce chapitre sont quand même très encourageants, cependant, ils n'ont pas pu être analysés en profondeur dans les temps impartis. Même si ce chapitre montre des limites, il ouvre des portes à de nouvelles ressources et méthodes d'analyses, dont des analyses multi-traits et multi-environnements.

Réponses aux objectifs

Finalement, cette thèse a-t-elle répondu aux objectifs qui avaient été proposés à la rédaction du projet ?

(1) *Une meilleure compréhension de la diversité variétale chez le blé pour les mécanismes de réponse au stress hydrique.*

La variabilité du panel de variétés a été explorée pour plusieurs types de stress (hydrique, thermique) et une prospection multi-stress a permis de mettre en évidence une nouvelle ressource disponible pour la sélection.

(2) *Des zones du génome associées avec des caractères d'intérêts liés à la tolérance à la sécheresse (approche GWAS) et des gènes candidats présents dans les zones.*

Les mécanismes de réponses ont aussi été étudiés *via* la découverte de QTL de réponse à ces stress, avec plus ou moins de succès comme nous avons pu le voir. La détection de gènes candidats dans les zones du génome a été prospectée pour les QTL qui présentaient le moins de gènes, mais il est évident qu'une charge de travail non négligeable reste à produire pour faire une analyse complète. De plus, comme nous avons pu le voir, il semble que le panel ne soit pas le plus adapté en terme de puissance pour espérer identifier des QTL très résolutif avec un nombre restreint de gènes candidats.

(3) *Des nouvelles méthodes d'analyses multivariées permettant de mieux comprendre et décrire la variabilité génétique pour les caractères étudiés.*

Tout au long de cette thèse, nous nous sommes employés à mettre en place des méthodes d'analyses spécifiques des données qui étaient à notre disposition. Le chapitre V présente une analyse multi-environnementale qui correspond à nos attentes en termes de gain de puissance et d'analyse a posteriori (relation entre effet des SNP et variables environnementales). Une analyse multi-trait a aussi été réalisée et présentée au congrès scientifique Eucarpia Cereal Section 2018, mais les résultats n'ont pas été présentés dans cette thèse. Pour valider clairement ces travaux, une simulation de données phénotypique avec des relations de cause à effet entre traits semble nécessaire et n'a pas été réalisée dans ces travaux.

(4) *De nouvelles méthodes de phénotypage (densité stomatique, racines, température épi/canopée), ou d'analyses de phénotypes (cinétique d'accumulation de biomasse ou de sénescence).*

Des méthodes de détermination de la cinétique de sénescence ont été proposées pour suivre l'impact du stress thermique sur les variétés. Ces mêmes méthodes ont été appliquées à des essais aux champs, notamment, les essais réalisés à Walla Walla. L'ensemble des données a été produit, de l'ajustement des courbes de sénescence à l'identification de QTL pour ces essais, mais n'ont pas été présentées dans cette thèse. Les nouvelles méthodes de phénotypage (*densité stomatique, racines, température épi/canopée*) n'ont pas été utilisées dans cette thèse par manque de temps, ce qui limite toutefois les possibilités d'identification des mécanismes écophysologiques qui sont impactés par les QTL qui ont été identifiés.

En dehors des objectifs fixés au départ, d'autres objectifs sont apparus au cours du travail, notamment une des questions qui est souvent posée dans la recherche végétale, notamment pour des sujets comme la tolérance à des stress abiotiques, est de savoir s'il vaut mieux réaliser de nombreux essais dans plusieurs lieux et environnements, en prospectant un maximum de types de stress et de scénarios, ou s'il est préférable de contrôler un maximum les conditions de culture pour se focaliser sur un seul type de stress. Cette question s'est imposée au cours de cette thèse et l'organisation des chapitres répond en elle-même à cette question. En effet, au début de cette thèse nous avons choisi de ne pas analyser les essais avec des modalités irriguées vs non-irriguées, car les scénarios de stress peuvent être très différents d'une année à l'autre et nous avons vu que des scénarios irrigués pouvaient néanmoins présenter des stress hydriques. Donc à l'échelle d'un réseau d'essais, une classification par type de stress observé semble bien plus pertinente qu'une classification par modalité mise en place. Par contre, nous avons aussi vu que dans les réseaux d'essais il est difficile, voire impossible de décorrélérer certaines variables environnementales (par exemple stress hydrique et stress thermique). Il est donc extrêmement difficile de postuler sur les traits physiologiques ou morphologiques qui sont impactés par les QTL et donc ceux qui jouent un rôle dans les performances des variétés en cas de stress. Il semble donc que des analyses qui se focalisent sur un seul type de stress (Chapitre IV), quitte à être moins représentatives des conditions de cultures au champ, sont d'une aide précieuse et complémentaire pour disséquer finement les mécanismes physiologiques et génétiques d'un type de stress donné.

Au final, nous espérons que cette thèse contribuera à enrichir les connaissances à la fois sur la caractérisation des environnements stressés en France, sur la variabilité génétique existante dans le matériel élite et sur le déterminisme génétique (QTL) des caractères étudiés. De plus, nous espérons que ce manuscrit constituera un bon appui pour confronter et/ou valider les résultats d'actuels ou de futurs projets sur la tolérance de variétés européennes aux stress thermiques et hydriques.

Comment répondre aux objectifs qui n'ont pas été atteints ?

Si cette thèse semble avoir rempli la majorité des objectifs qui avaient été établis à l'écriture du projet, plusieurs améliorations auraient pu et peuvent être réalisées. Il faut notamment souligner le manque de relation entre nos travaux et les mécanismes physiologiques de tolérance à la sécheresse qui ont pu

être décrits dans le chapitre I. De nombreuses expérimentations et projets en lien avec notre travail sur les stress hydriques et thermiques ont été réalisés ou le seront très prochainement. On peut prendre l'exemple des plateformes Pheno3C ou Phenofield, qui comportent des abris roulants permettant de contrôler les précipitations sur des expérimentations au champ. Deux essais sur ces plateformes ont été analysés dans cette thèse et intégrés au réseau multi-environnemental, mais un intérêt important de ce type de plate-forme réside dans l'utilisation de nombreux outils de phénotypage hauts débits qui peuvent amener des informations sur la physiologie des plantes de façon complémentaire (surface foliaire, teneur en chlorophylle, température de canopée...) et de façon cinétique au cours de la culture. La phénomobile et les drones sont des outils aussi utilisables dans des champs « classiques » qui devraient permettre de faire plus largement le lien entre les détections de régions génomiques et leur rôle dans la mise en place des stratégies de tolérance au stress hydrique et thermique que nous avons vues précédemment. Autre exemple, le projet FSOV ArchiRac, projet en cours de réalisation, qui vise à caractériser finement l'architecture racinaire du blé à des stades juvéniles en conditions contrôlées. Dans ce projet, le panel élite a été caractérisé et une analyse des déterminants génétiques sera réalisée et pourra être confrontée aux résultats de notre études.

En ce qui concerne la sélection de nouvelles variétés de blé en condition de stress hydrique, des études de sélection génomique sont aussi prospectées : Rincet et al. (en prep). En effet, vu les faibles effets des QTL détectés dans notre étude, il pourrait être intéressant de travailler sur tous les marqueurs simultanément grâce à des approches de prédiction génomique, de manière à travailler sur une part plus grande de la variabilité génétique. Le principe de cette approche, qui est utilisée par la plupart des entreprises de sélection, est d'estimer l'effet de tous les SNP simultanément en utilisant un jeu de calibration, ce qui permet ensuite de prédire et sélectionner des individus génotypés mais aussi non phénotypés. Ceci est particulièrement intéressant quand le coût de phénotypage est élevé comme dans le cas présent. Cette approche nécessite néanmoins des développements supplémentaires pour qu'elle puisse également prendre en compte les GxE de manière efficace. Une solution peut être de définir des groupes d'environnements similaires ou d'introduire des covariables environnementales dans le modèle statistique (Heslot et al., 2014; Jarquín et al., 2014; Ly et al., 2018).

Références

Références

- Aastveit, A.H., Martens, H., 1986. ANOVA Interactions Interpreted by Partial Least Squares Regression. *Biometrics* 42, 829-844. <https://doi.org/10.2307/2530697>
- Abbate, P.E., Andrade, F.H., Culot, J.P., 1995. The effects of radiation and nitrogen on number of grains in wheat. *J. Agric. Sci.* 124, 351-360. <https://doi.org/10.1017/S0021859600073317>
- Acevedo, E., Silva, P., Silva, H., 2009. *Wheat Growth and Physiology*, F.A.O. Corporate Repository, pp. 1-24.
- Acevedo, E., Silva, P., Silva, H., 2006. Growth and wheat physiology, development 47.
- Acuña-Galindo, M.A., Mason, R.E., Subramanian, N.K., Hays, D.B., 2015. Meta-Analysis of Wheat QTL Regions Associated with Adaptation to Drought and Heat Stress. *Crop Sci.* 55, 477-492. <https://doi.org/10.2135/cropsci2013.11.0793>
- Ain, Q., Rasheed, A., Anwar, A., Mahmood, T., Imtiaz, M., Xia, X., He, Z., Quraishi, U.M., 2015. Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Front. Plant Sci.* 6, 743. <https://doi.org/10.3389/fpls.2015.00743>
- Alexieva, V., Sergiev, I., Mapelli, S., Karanov, E., 2001. The effect of drought and ultraviolet radiation on growth and stress markers in pea and wheat. *Plant Cell Environ.* 24, 1337-1344. <https://doi.org/10.1046/j.1365-3040.2001.00778.x>
- Al-Khatib, K., Paulsen, G.M., 1999. High-Temperature Effects on Photosynthetic Processes in Temperate and Tropical Cereals. *Crop Sci.* 39, 119-125. <https://doi.org/10.2135/cropsci1999.0011183X003900010019x>
- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration -guidelines for computing crop water requirements. Food Agric. Organ. FAO Irrigation and drainage.
- Allison, J.C.S., Daynard, T.B., 1976. Effect of photoperiod on development and number of spikelets of a temperate and some low-latitude wheats. *Ann. Appl. Biol.* 83, 93-102. <https://doi.org/10.1111/j.1744-7348.1976.tb01698.x>
- AL-Shalaldehy, G., Duwayri, M.A., 1986. Inheritance of morphologie-physiological characters and grain yield in durum wheat crosses. *Rachis* 5, 37-42.
- Anjum, F., Wahid, A., Javed, F., Arshad, M., 2008. Influence of Foliar Applied Thiourea on Flag Leaf Gas Exchange and Yield Parameters of Bread Wheat (*Triticum aestivum*) Cultivars under Salinity and Heat Stresses. *International Journal of Agriculture & Biology* 10, 619-26.
- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., The International Wheat Genome Sequencing Consortium (IWGSC), 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361. <https://doi.org/10.1126/science.aar7191>
- Auriau, P., Doussinault, G., Jahier, J., Lecomte, C., Pierre, J., Rousset, M., Saur, L., Trotter, M., 1992. Le blé tendre, in: *Amélioration Des Espèces Végétales Cultivées: Objectifs et Critères de Sélection*. Gallais, A., Bannerot, H., Anglade, P., Paris, pp. 22-38.
- Ayers, K.L., Cordell, H.J., 2010. SNP Selection in genome-wide and candidate gene studies via penalized logistic regression. *Genet. Epidemiol.* 34, 879-891. <https://doi.org/10.1002/gepi.20543>
- Baillet, N., Girusse, C., Allard, V., Piquet-Pissaloux, A., Le Gouis, J., 2018. Different grain-filling rates explain grain-weight differences along the wheat ear. *PLOS ONE* 13, e0209597. <https://doi.org/10.1371/journal.pone.0209597>
- Baker, C.K., Gallagher, J.N., 1983. The development of winter wheat in the field. 1. Relation between apical development and plant morphology within and between seasons. *J. Agric. Sci.* 101, 327-335. <https://doi.org/10.1017/S0021859600037631>
- Bancal, M.-O., Brisson, N., Gate, P., 2010. Changement climatique et cultures de blé et de colza : l'essentiel des impacts. *Présent. Méthodes Résultats Proj. Clim. Clim. 2010 Versailles FRA 2010-06-17 - 2010-06-18* 32-33.
- Barber, H.M., Lukac, M., Simmonds, J., Semenov, M.A., Gooding, M.J., 2017. Temporally and Genetically Discrete Periods of Wheat Sensitivity to High Temperature. *Front. Plant Sci.* 8. <https://doi.org/10.3389/fpls.2017.00051>
- Baret, F., Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161-173. [https://doi.org/10.1016/0034-4257\(91\)90009-U](https://doi.org/10.1016/0034-4257(91)90009-U)
- Barlow, E., Donovan, G., Lee, J., 1983. Water Relations and Composition of Wheat Ears Grown in Liquid Culture: Effect of Carbon and Nitrogen. *Funct. Plant Biol.* 10, 99. <https://doi.org/10.1071/PP9830099>
- Barnabás, B., Jäger, K., Fehér, A., 2007. The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* -. <https://doi.org/10.1111/j.1365-3040.2007.01727.x>

- Bates, D.M., Chambers, J.M., 1992. Nonlinear models, in: Chapter 10 of Statistical Models in S Eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole.
- Bates, D.M., Watts, D.G., 1988. Nonlinear Regression Analysis and Its Applications. Wiley.
- Ben-Ari, T., Boé, J., Ciais, P., Lecerf, R., Van der Velde, M., Makowski, D., 2018. Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. *Nat. Commun.* 9. <https://doi.org/10.1038/s41467-018-04087-x>
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bennett, D., Izanloo, A., Edwards, J., Kuchel, H., Chalmers, K., Tester, M., Reynolds, M., Schnurbusch, T., Langridge, P., 2012. Identification of novel quantitative trait loci for days to ear emergence and flag leaf glaucousness in a bread wheat (*Triticum aestivum* L.) population adapted to southern Australian conditions. *Theor. Appl. Genet.* 124, 697–711. <https://doi.org/10.1007/s00122-011-1740-3>
- Bentley, A.R., Horsnell, R., Werner, C.P., Turner, A.S., Rose, G.A., Bedard, C., Howell, P., Wilhelm, E.P., Mackay, I.J., Howells, R.M., Greenland, A., Laurie, D.A., Gosman, N., 2013. Short, natural, and extended photoperiod response in BC2F4 lines of bread wheat with different Photoperiod-1 (Ppd-1) alleles. *J. Exp. Bot.* 64, 1783–1793. <https://doi.org/10.1093/jxb/ert038>
- Bernardo, R., 2001. What If We Knew All the Genes for a Quantitative Trait in Hybrid Crops? *Crop Sci.* 41, 1. <https://doi.org/10.2135/cropsci2001.4111>
- Bhullar, S., Jenner, C., 1985. Differential Responses to High Temperatures of Starch and Nitrogen Accumulation in the Grain of Four Cultivars of Wheat. *Funct. Plant Biol.* 12, 363. <https://doi.org/10.1071/PP9850363>
- Blum, A., 1997. Improving wheat grain filling under stress by stem reserve mobilisation, in: Braun, H.-J., Altay, F., Kronstad, W.E., Beniwal, S.P.S., McNab, A. (Eds.), *Wheat: Prospects for Global Improvement*. Springer Netherlands, Dordrecht, pp. 135–141. https://doi.org/10.1007/978-94-011-4896-2_19
- Boer, M.P., Wright, D., Feng, L., Podlich, D.W., Luo, L., Cooper, M., van Eeuwijk, F.A., 2007. A Mixed-Model Quantitative Trait Loci (QTL) Analysis for Multiple-Environment Trial Data Using Environmental Covariables for QTL-by-Environment Interactions, With an Example in Maize. *Genetics* 177, 1801–1813. <https://doi.org/10.1534/genetics.107.071068>
- Bogard, M., Jourdan, M., Allard, V., Martre, P., Perretant, M.R., Ravel, C., Heumez, E., Orford, S., Snape, J., Griffiths, S., Gaju, O., Foulkes, J., Le Gouis, J., 2011. Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *J. Exp. Bot.* 62, 3621–3636. <https://doi.org/10.1093/jxb/err061>
- Bogard, M., Ravel, C., Paux, E., Bordes, J., Balfourier, F., Chapman, S.C., Le Gouis, J., Allard, V., 2014. Predictions of heading date in bread wheat (*Triticum aestivum* L.) using QTL-based parameters of an ecophysiological model. *J. Exp. Bot.* 65, 5849–5865. <https://doi.org/10.1093/jxb/eru328>
- Bonneau, J., Taylor, J., Parent, B., Bennett, D., Reynolds, M., Feuillet, C., Langridge, P., Mather, D., 2013. Multi-environment analysis and improved mapping of a yield-related QTL on chromosome 3B of wheat. *Theor. Appl. Genet.* 126, 747–761. <https://doi.org/10.1007/s00122-012-2015-3>
- Borojevic, Katarina, Borojevic, Ksenija, 2005. The Transfer and History of “Reduced Height Genes” (Rht) in Wheat from Japan to Europe. *J. Hered.* 96, 455–459. <https://doi.org/10.1093/jhered/esi060>
- Borrell, A., Hammer, G., Douglas, A., 2000. Does Maintaining Green Leaf Area in Sorghum Improve Yield under Drought? I. Leaf Growth and Senescence. *Crop Sci* 40.
- Borrell, A.K., Incoll, L.D., Dalling, M.J., 1991. The Influence of the Rht1 and Rht2 Alleles on the Growth of Wheat Stems and Ears. *Ann. Bot.* 67, 103–110. <https://doi.org/10.1093/oxfordjournals.aob.a088108>
- Bouffier, B., 2014. Genetic and ecophysiological dissection of tolerance to drought and heat stress in bread wheat: from environmental characterization to QTL detection. Université Blaise Pascal-Clermont-Ferrand II.
- Bouffier, B., Derory, J., Murigneux, A., Reynolds, M., Le Gouis, J., 2015. Clustering of Environmental Parameters Discriminates Drought and Heat Stress Bread Wheat Trials. *Agron. J.* 107, 1489. <https://doi.org/10.2134/agronj14.0452>
- Brancourt-Hulmel, M., 1999. Crop diagnosis and probe genotypes for interpreting genotype environment interaction in winter wheat trials. *Theor. Appl. Genet.* 99, 1018–1030.
- Brancourt-Hulmel, M., Denis, J.-B., Lecomte, C., 2000. Determining environmental covariates which explain genotype environment interaction in winter wheat through probe genotypes and biadditive factorial regression. *Theor. Appl. Genet.* 100, 285–298.

- Brancourt-Hulmel, M., Doussinault, G., Lecomte, C., Bérard, P., Le Buanec, B., Trottet, M., 2003. Genetic Improvement of Agronomic Traits of Winter Wheat Cultivars Released in France from 1946 to 1992. *CROP Sci.* 43, 37–45.
- Brancourt-Hulmel, M., Lecomte, C., Meynard, J.-M., 1999. A diagnosis of yield-limiting factors on probe genotypes for characterizing environments in winter wheat trials.
- Braun, H.J., Atlin, G., Payne, T., 2010. Multi-location testing as a tool to identify plant response to global climate change., in: Reynolds, M.P. (Ed.), *Climate Change and Crop Production*. CABI, Wallingford, pp. 115–138. <https://doi.org/10.1079/9781845936334.0115>
- Braun, H.-J., Rajaram, S., 1996. CIMMYT's approach to breeding for wide adaptation. *Euphytica* 175–183.
- Breseghele, F., 2005. Association Mapping of Kernel Size and Milling Quality in Wheat (*Triticum aestivum* L.) Cultivars. *Genetics* 172, 1165–1177. <https://doi.org/10.1534/genetics.105.044586>
- Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X., Huard, F., 2010. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Res.* 119, 201–212. <https://doi.org/10.1016/j.fcr.2010.07.012>
- Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoulaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.-M., Meynard, J.M., Delécolle, R., 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie* 18, 311–346. <https://doi.org/10.1051/agro:19980501>
- Brocklehurst, P.A., Moss, J.P., Williams, W., 1978. Effects of irradiance and water supply on grain development in wheat. *Ann. Appl. Biol.* 90, 265–276. <https://doi.org/10.1111/j.1744-7348.1978.tb02635.x>
- Brooks, A., Jenner, C., Aspinall, D., 1982. Effects of Water Deficit on Endosperm Starch Granules and on Grain Physiology of Wheat and Barley. *Funct. Plant Biol.* 9, 423. <https://doi.org/10.1071/PP9820423>
- Browning, B.L., Browning, S.R., 2016. Genotype Imputation with Millions of Reference Samples. *Am. J. Hum. Genet.* 98, 116–126. <https://doi.org/10.1016/j.ajhg.2015.11.020>
- Budak, H., Hussain, B., Khan, Z., Ozturk, N.Z., Ullah, N., 2015. From Genetics to Functional Genomics: Improvement in Drought Signaling and Tolerance in Wheat. *Front. Plant Sci.* 6. <https://doi.org/10.3389/fpls.2015.01012>
- Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J., 2009. *ASReml-R reference manual*. State Qld. Dep. Prim. Ind. Fish. Brisb.
- Butler, J.D., Byrne, P.F., Mohammadi, V., Chapman, P.L., Haley, S.D., 2005. Agronomic Performance of Rht Alleles in a Spring Wheat Population across a Range of Moisture Levels. *Crop Sci.* 45, 939–947. <https://doi.org/10.2135/cropsci2004.0323>
- Calderini, D.F., Abeledo, L.G., Savin, R., Slafer, G.A., 1999. Effect of temperature and carpel size during pre-anthesis on potential grain weight in wheat. *J. Agric. Sci.* 132, 453–459.
- Calderini, D.F., Slafer, G.A., 1998. Changes in yield and yield stability in wheat during the 20th century. *Field Crops Res.* 57, 335–347.
- Cassman, K.G., 1999. Ecological Intensification of Cereal Production Systems: Yield Potential, Soil Quality, and Precision Agriculture. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5952–5959.
- Chapman, S.C., 2008. Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. *Euphytica* 161, 195–208. <https://doi.org/10.1007/s10681-007-9623-z>
- Chatelin, M.H., Aubry, C., Poussin, J.C., Meynard, J.M., Massé, J., Verjux, N., Gate, P., Le Bris, X., 2005. DéciBlé, a software package for wheat crop management simulation. *Agric. Syst.* 83, 77–99. <https://doi.org/10.1016/j.agsy.2004.03.003>
- Chen, F.H., Dong, G.H., Zhang, D.J., Liu, X.Y., Jia, X., An, C.B., Ma, M.M., Xie, Y.W., Barton, L., Ren, X.Y., Zhao, Z.J., Wu, X.H., Jones, M.K., 2015. Agriculture facilitated permanent human occupation of the Tibetan Plateau after 3600 B.P. *Science* 347, 248. <https://doi.org/10.1126/science.1259172>
- Chen, J.M., Black, T.A., 1992. Defining leaf area index for non-flat leaves. *Plant Cell Environ.* 15, 421–429. <https://doi.org/10.1111/j.1365-3040.1992.tb00992.x>
- Chenu, K., 2015. Characterizing the crop environment – nature, significance and applications, in: *Crop Physiology*. Elsevier, pp. 321–348. <https://doi.org/10.1016/B978-0-12-417104-6.00013-3>
- Chenu, K., Cooper, M., Hammer, G.L., Mathews, K.L., Dreccer, M.F., Chapman, S.C., 2011. Environment characterization as an aid to wheat improvement: interpreting genotype-environment interactions by

- modelling water-deficit patterns in North-Eastern Australia. *J. Exp. Bot.* 62, 1743–1755. <https://doi.org/10.1093/jxb/erq459>
- Chenu, K., Deihimfard, R., Chapman, S.C., 2013. Large-scale characterization of drought pattern: a continent-wide modelling approach applied to the Australian wheatbelt - spatial and temporal trends. *New Phytol.* 198, 801–820. <https://doi.org/10.1111/nph.12192>
- Cho, S., Kim, K., Kim, Y.J., Lee, J.-K., Cho, Y.S., Lee, J.-Y., Han, B.-G., Kim, H., Ott, J., Park, T., 2010. Joint Identification of Multiple Genetic Variants via Elastic-Net Variable Selection in a Genome-Wide Association Analysis: Identifying multiple variants via EN. *Ann. Hum. Genet.* 74, 416–428. <https://doi.org/10.1111/j.1469-1809.2010.00597.x>
- Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L., Sourdille, P., Couloux, A., Paux, E., Leroy, P., Mangenot, S., Guilhot, N., Le Gouis, J., Balfourier, F., Alaux, M., Jamilloux, V., Poulain, J., Durand, C., Bellec, A., Gaspin, C., Safar, J., Dolezel, J., Rogers, J., Vandepoele, K., Aury, J.-M., Mayer, K., Berges, H., Quesneville, H., Wincker, P., Feuillet, C., 2014. Structural and functional partitioning of bread wheat chromosome 3B. *Science* 345, 1249721–1249721. <https://doi.org/10.1126/science.1249721>
- Christopher, J.T., Christopher, M.J., Borrell, A.K., Fletcher, S., Chenu, K., 2016. Stay-green traits to improve wheat adaptation in well-watered and water-limited environments. *J. Exp. Bot.* 67, 5159–5172. <https://doi.org/10.1093/jxb/erw276>
- Christopher, J.T., Manschadi, A.M., Hammer, G.L., Borrell, A.K., 2008. Developmental and physiological traits associated with high yield and stay-green phenotype in wheat. *Aust. J. Agric. Res.* 59, 354. <https://doi.org/10.1071/AR07193>
- Christopher, J.T., Veyradier, M., Borrell, A.K., Harvey, G., Fletcher, S., Chenu, K., 2014. Phenotyping novel stay-green traits to capture genetic variation in senescence dynamics. *Funct. Plant Biol.* 41, 1035. <https://doi.org/10.1071/FP14052>
- Christopher, M., Chenu, K., Jennings, R., Fletcher, S., Butler, D., Borrell, A., Christopher, J., 2018. QTL for stay-green traits in wheat in well-watered and water-limited environments. *Field Crops Res.* 217, 32–44. <https://doi.org/10.1016/j.fcr.2017.11.003>
- Commuri, P.D., Jones, R.J., 2001. High Temperatures during Endosperm Cell Division in Maize: A Genotypic Comparison under In Vitro and Field Conditions. *CROP Sci.* 41, 9.
- Condon, A.G., 2004. Breeding for high water-use efficiency. *J. Exp. Bot.* 55, 2447–2460. <https://doi.org/10.1093/jxb/erh277>
- Cooper, M., Fox, P.N., 1996. Environmental characterization based on probe and reference genotypes. Cooper M Hammer GL Eds *Plant Adapt. Crop Improv.* Wallingford UK 529–547.
- Cooper, M., Woodruff, D.R., Eisemann, R.L., Brennan, P.S., DeLacy, I.H., 1995. A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat: managed-environments for selection among genotypes. *Theor. Appl. Genet.* 90. <https://doi.org/10.1007/BF00221995>
- Cordell, H.J., Clayton, D.G., 2002. A Unified Stepwise Regression Procedure for Evaluating the Relative Effects of Polymorphisms within a Gene Using Case/Control or Family Data: Application to HLA in Type 1 Diabetes. *Am. J. Hum. Genet.* 70, 124–141. <https://doi.org/10.1086/338007>
- Cormier, F., 2015. Nitrogen use efficiency in wheat in bread wheat (*T. aestivum* L.): breeding & gene discovery. Université Blaise Pascal-Clermont-Ferrand II.
- Cormier, F., Faure, S., Dubreuil, P., Heumez, E., Beauchêne, K., Lafarge, S., Praud, S., Le Gouis, J., 2013. A multi-environmental study of recent breeding progress on nitrogen use efficiency in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 126, 3035–3048. <https://doi.org/10.1007/s00122-013-2191-9>
- Cormier, F., Le Gouis, J., Dubreuil, P., Lafarge, S., Praud, S., 2014. A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 127, 2679–2693. <https://doi.org/10.1007/s00122-014-2407-7>
- Cossani, C.M., Reynolds, M.P., 2012. Physiological Traits for Improving Heat Tolerance in Wheat. *PLANT Physiol.* 160, 1710–1718. <https://doi.org/10.1104/pp.112.207753>
- Cox, D.R., Snell, E.J., 1989. *Analysis of Binary Data*, Second Edition. Chapman and Hall/CRC.
- Cullis, B.R., Smith, A.B., Coombes, N.E., 2006. On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11, 381–393. <https://doi.org/10.1198/108571106X154443>
- Darwinkel, A., 1978. Patterns of tillering and grain production of winter wheat at a wide range of plant densities. *Neth. J. Agric. Sci.* 26, 383–398.
- de Wit, C.T., 1965. Photosynthesis of Leaf Canopies. *Agric. Res. Rep.* 663 Wagening. Cent. Agric. Publ. Doc.

- Deak, K.I., Malamy, J., 2005. Osmotic regulation of root system architecture: Osmotic regulation of lateral root formation. *Plant J.* 43, 17–28. <https://doi.org/10.1111/j.1365-313X.2005.02425.x>
- Demotes-Mainard, S., Jeuffroy, M.-H., 2004. Effects of nitrogen and radiation on dry matter and nitrogen accumulation in the spike of winter wheat. *Field Crops Res.* 87, 221–233. <https://doi.org/10.1016/j.fcr.2003.11.014>
- Denis, J. baptiste, 1988. Two way analysis using covariates. *Statistics* 19, 123–132. <https://doi.org/10.1080/02331888808802080>
- Denyer, K., Hylton, C., Smith, A., 1994. The Effect of High Temperature on Starch Synthesis and the Activity of Starch Synthase. *Funct. Plant Biol.* 21, 783. <https://doi.org/10.1071/PP9940783>
- Desclaux, D., 1996. De l'intérêt de génotypes révélateurs de facteurs limitants dans l'analyse des interactions génotype x milieu chez le soja (*Glycine max.* L. Merrill). Institut national polytechnique de Toulouse, Toulouse, France.
- Deswarte, J.-C., Beauchêne, K., Arjaure, G., Moittie, T., Bedard, C., Duchalais, L., Derory, J., Zimmerli, C., Laurent, V., Debiton, C., Gakiere, B., Baret, F., 2017. Identification des traits de tolérance à la sécheresse et élaboration des outils d'aide à leur évaluation (Synthèse des programmes de recherche 2012), FSOV-2012-K.
- Dias, A.S., Lidon, F.C., 2009. Evaluation of Grain Filling Rate and Duration in Bread and Durum Wheat, under Heat Stress after Anthesis. *J. Agron. Crop Sci.* 195, 137–147. <https://doi.org/10.1111/j.1439-037X.2008.00347.x>
- Dunn, O.J., 1964. Multiple Comparisons Using Rank Sums. *Technometrics* 6, 241–252. <https://doi.org/10.1080/00401706.1964.10490181>
- Edae, E.A., Byrne, P.F., Haley, S.D., Lopes, M.S., Reynolds, M.P., 2014. Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor. Appl. Genet.* 127, 791–807. <https://doi.org/10.1007/s00122-013-2257-8>
- Efron, B., 1978. Regression and ANOVA with Zero-One Data: Measures of Residual Variation. *J. Am. Stat. Assoc.* 73, 113–121. <https://doi.org/10.2307/2286531>
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- FAO, 2018. FAOSTAT.
- Farooq, M., Bramley, H., Palta, J.A., Siddique, K.H.M., 2011. Heat Stress in Wheat during Reproductive and Grain-Filling Phases. *Crit. Rev. Plant Sci.* 30, 491–507. <https://doi.org/10.1080/07352689.2011.615687>
- Farooq, M., Hussain, M., Siddique, K.H.M., 2014. Drought Stress in Wheat during Flowering and Grain-filling Periods. *Crit. Rev. Plant Sci.* 33, 331–349. <https://doi.org/10.1080/07352689.2014.875291>
- Farquhar, G., Richards, R., 1984. Isotopic Composition of Plant Carbon Correlates With Water-Use Efficiency of Wheat Genotypes. *Aust. J. Plant Physiol.* 11, 539. <https://doi.org/10.1071/PP9840539>
- Ferris, R., 1998. Effect of High Temperature Stress at Anthesis on Grain Yield and Biomass of Field-grown Crops of Wheat. *Ann. Bot.* 82, 631–639. <https://doi.org/10.1006/anbo.1998.0740>
- Finger, R., 2010. Evidence of slowing yield growth – The example of Swiss cereal yields. *Food Policy* 35, 175–182. <https://doi.org/10.1016/j.foodpol.2009.11.004>
- Finlay, K.W., Wilkinson, G.N., 1963. The analysis of adaptation in a plant-breeding programme. *Crop Pasture Sci.* 14, 742–754.
- Fischer, R., Maurer, R., 1978. Drought resistance in spring wheat cultivars. I. Grain yield responses. *Aust. J. Agric. Res.* 29, 897. <https://doi.org/10.1071/AR9780897>
- Flint-Garcia, S.A., Thornsberry, J.M., Buckler, E.S., 2003. Structure of Linkage Disequilibrium in Plants. *Annu. Rev. Plant Biol.* 54, 357–374. <https://doi.org/10.1146/annurev.arplant.54.031902.134907>
- Flom, P.L., Development, N., Institutes, R., York, N., 2007. Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use 7.
- FranceAgriMer, 2017. Les variétés de céréales à paille préférées des agriculteurs. Terre-Net Média.
- FranceAgriMer [WWW Document], 2019. . FranceAgriMer Établ. National Prod. Agric. Mer. URL <https://www.franceagrimer.fr/filiere-grandes-cultures/Cereales>
- Frensch, J., Hsiao, T.C., Steudle, E., 1996. Water and solute transport along developing maize roots. *Planta* 198, 348–355. <https://doi.org/10.1007/BF00620050>

- Frensham, A., Cullis, B., Verbyla, A., 1997. Genotype by Environment Variance Heterogeneity in a Two-Stage Analysis. *Biometrics* 53, 1373–1383. <https://doi.org/10.2307/2533504>
- Gahlaut, 2012. Genetic dissection of water stress tolerance in bread wheat.
- Gallo, K., Daughtry, C., Wiegand, C., 1993. Errors in Measuring Absorbed Radiation and Computing Crop Radiation Use Efficiency. *Agron. J.* 85, 1222–1228.
- Gallo, K.P., Daughtry, C., 1986. Techniques for measuring intercepted and absorbed photosynthetically active radiation in corn canopies. *Agron. J.* 78, 752–756.
- Gao, X., Becker, L.C., Becker, D.M., Starmer, J.D., Province, M.A., 2009. Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet. Epidemiol.* n/a-n/a. <https://doi.org/10.1002/gepi.20430>
- Gaspar, T., Franck, T., Bisbis, B., Kevers, C., Laurent, J., Hausman, J.F., Dommès, J., 2002. Concepts in plant stress physiology. Application to plant tissue cultures. *Plant Growth Regul.* 37, 263–285. <https://doi.org/10.1023/A:1020835304842>
- Gate, P., 2007. Evolution du climat et production du blé. *Perspect Agric* 336, 24–31.
- Gate, P., 1995. *Ecophysiologie du blé: de la plante à la culture*. Lavoisier, France.
- Gate, P., Blondlot, A., Gouache, D., Deudon, O., Vignier, L., 2008. Impacts du changement climatique sur la croissance et le développement du blé en France. Quelles solutions et quelles actions à développer ? *OCL - Ol. Corps Gras Lipides* 15, 332.
- Gate, P., Brisson, N., Gouache, D., 2010. Les causes du plafonnement du rendement du blé en France : d'abord une origine climatique. *Comptes-Rendus Académie Agric. Fr.* 96, 17.
- Gelman, A., 2005. Analysis of variance? why it is more important than ever. *Ann. Stat.* 33, 1–53. <https://doi.org/10.1214/009053604000001048>
- Genard, M., Reich, M., Lobit, P., Besset, J., 1999. Correlations between sugar and acid content and peach growth. *J. Hortic. Sci. Biotechnol.* 74, 772–776. <https://doi.org/10.1080/14620316.1999.11511187>
- Gervois, S., Ciais, P., de Noblet-Ducoudré, N., Brisson, N., Vuichard, N., Viovy, N., 2008. Carbon and water balance of European croplands throughout the 20th century: carbon balance of european croplands. *Glob. Biogeochem. Cycles* 22, n/a-n/a. <https://doi.org/10.1029/2007GB003018>
- Gibson, L.R., Paulsen, G.M., 1999. Yield Components of Wheat Grown under High Temperature Stress during Reproductive Growth. *Crop Sci.* 39, 1841–1846. <https://doi.org/10.2135/cropsci1999.3961841x>
- Giroux, M.J., Morris, C.F., 1998. Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci.* 95, 6262–6266. <https://doi.org/10.1073/pnas.95.11.6262>
- Gómez-Macpherson, H., 1998. Growth of Near-isogenic Wheat Lines Differing in Development—Plants in a Simulated Canopy. *Ann. Bot.* 82, 323–330. <https://doi.org/10.1006/anbo.1998.0692>
- Goncharov, N.P., Golovkina, K.A., Kondratenko, E.Y., 2009. Taxonomy and molecular phylogeny of natural and artificial wheat species. *Breed. Sci.* 59, 492–498. <https://doi.org/10.1270/jsbbs.59.492>
- Górny, A.G., Garczyński, S., 2005. Genotypic and nutrition-dependent variation in water use efficiency and photosynthetic activity of leaves in winter wheat (*Triticum aestivum* L.). *J. Appli Genet* 43, 145–160.
- Gouache, D., Bouchon, A.-S., Jouanneau, E., Le Bris, X., 2015. Agrometeorological analysis and prediction of wheat yield at the departmental level in France. *Agric. For. Meteorol.* 209–210, 1–10. <https://doi.org/10.1016/j.agrformet.2015.04.027>
- Gouache, D., Le Bris, X., Bogard, M., Deudon, O., Pagé, C., Gate, P., 2012. Evaluating agronomic adaptation options to increasing heat stress under climate change during wheat grain filling in France. *Eur. J. Agron.* 39, 62–70. <https://doi.org/10.1016/j.eja.2012.01.009>
- Graybosch, R.A., Peterson, C.J., 2010. Genetic Improvement in Winter Wheat Yields in the Great Plains of North America, 1959–2008. *Crop Sci.* 50, 1882. <https://doi.org/10.2135/cropsci2009.11.0685>
- Gupta, P., Balyan, H., Gahlaut, V., 2017. QTL Analysis for Drought Tolerance in Wheat: Present Status and Future Possibilities. *Agronomy* 7, 5. <https://doi.org/10.3390/agronomy7010005>
- Hall, A.E., 2012. Heat Stress, in: *Plant Stress Physiology*, CAB International. S. Shabala, pp. 118–131.
- Hammer, G., Jordan, D., 2007. An integrated systems approach to crop improvement. Spiertz JHJ Struik PC Van Laar HH Eds *Scale Complex. Plant Syst. Res. Gene–plant–crop Relat. Wagening. UR–Frontis Ser. No 21 Springer Dordecht Neth.* Pp 45–61.
- Harding, S.A., Guikema, J.A., Paulsen, G.M., 1990. Photosynthetic Decline from High Temperature Stress during Maturation of Wheat: I. Interaction with Senescence Processes. *PLANT Physiol.* 92, 648–653. <https://doi.org/10.1104/pp.92.3.648>

- Hay, R., Kirby, E., 1991. Convergence and synchrony—a review of the coordination of development in wheat. *Aust. J. Agric. Res.* 42, 661. <https://doi.org/10.1071/AR9910661>
- Hay, R.K.M., 1999. Physiological Control of Growth and Yield in Wheat: Analysis and Synthesis, in: Smith, D.L., Hamel, C. (Eds.), *Crop Yield*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–38. https://doi.org/10.1007/978-3-642-58554-8_1
- Hays, D.B., Do, J.H., Mason, R.E., Morgan, G., Finlayson, S.A., 2007. Heat stress induced ethylene production in developing wheat grains induces kernel abortion and increased maturation in a susceptible cultivar. *Plant Sci.* 172, 1113–1123. <https://doi.org/10.1016/j.plantsci.2007.03.004>
- Hellebusi, J.A., 1976. Osmoregulation. *Annu. Rev. Plant Physiol. Plant Physiol* 27, 485–505.
- Heslot, N., Akdemir, D., Sorrells, M.E., Jannink, J.-L., 2014. Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127, 463–480. <https://doi.org/10.1007/s00122-013-2231-5>
- Hey, J., Machado, C.A., 2003. The study of structured populations — new hope for a difficult and divided science. *Nat. Rev. Genet.* 4, 535–543. <https://doi.org/10.1038/nrg1112>
- Hoggart, C.J., Whittaker, J.C., De Iorio, M., Balding, D.J., 2008. Simultaneous Analysis of All SNPs in Genome-Wide and Re-Sequencing Association Studies. *PLoS Genet.* 4, e1000130. <https://doi.org/10.1371/journal.pgen.1000130>
- Holmes, M.G., Smith, H., 1977. The Function of phytochrome in the natural environment-II. The influence of vegetation canopies on the spectral energy distribution of natural daylight. *Photochem. Photobiol.* 25, 539–545. <https://doi.org/10.1111/j.1751-1097.1977.tb09125.x>
- Holworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014a. APSIM – Evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.* 62, 327–350. <https://doi.org/10.1016/j.envsoft.2014.07.009>
- Holworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014b. APSIM – Evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.* 62, 327–350. <https://doi.org/10.1016/j.envsoft.2014.07.009>
- Hunt, L.A., Poorten, G. van der, Pararajasingham, S., 1991. Postanthesis temperature effects on duration and rate of grain filling in some winter and spring wheats. *Can. J. Plant Sci.* 71, 609–617. <https://doi.org/10.4141/cjps91-092>
- Jamieson, P.D., Semenov, M.A., Brooking, I.R., Francis, G.S., 1998. Sirius: a mechanistic model of wheat response to environmental variation. *Eur. J. Agron.* 8, 161–179. [https://doi.org/10.1016/S1161-0301\(98\)00020-3](https://doi.org/10.1016/S1161-0301(98)00020-3)
- Jamil, M., Ali, A., Ghafoor, A., Akbar, K.F., Aziz, A., Naveed, N.H., Yasin, N.A., Gul, A., Mujeeb-Kazi, A., 2017. Digital image analysis of seed shape influenced by heat stress in diverse bread wheat germplasm. *Pak J Bot* 1279–1284.
- Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., de los Campos, G., 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127, 595–607. <https://doi.org/10.1007/s00122-013-2243-1>
- Jenner, C., 1994. Starch Synthesis in the Kernel of Wheat Under High Temperature Conditions. *Funct. Plant Biol.* 21, 791. <https://doi.org/10.1071/PP9940791>
- Johnson, I.R., Thornley, J.H.M., 1984. Temperature Dependence of Plant and Crop Processes. *Ann. Bot.* 55, 1–24. <https://doi.org/doi.org/10.1093/oxfordjournals.aob.a086868>
- Jolly, C.J., Glenn, G.M., Rahman, S., 1996. GSP-1 genes are linked to the grain hardness locus (Ha) on wheat chromosome 5D. *Proc. Natl. Acad. Sci. U. S. A.* 93, 2408–2413.

- Jones, H.E., Lukac, M., Brak, B., Martinez-Eixarch, M., Alhomedhi, A., Gooding, M.J., Wingen, L.U., Griffiths, S., 2017. Photoperiod sensitivity affects flowering duration in wheat. *J. Agric. Sci.* 155, 32–43. <https://doi.org/10.1017/S0021859616000125>
- Jones, R.J., Roessler, J., Ouattar, S., 1985. Thermal Environment During Endosperm Cell Division in Maize: Effects on Number of Endosperm Cells and Starch Granules. *Crop Sci.* 25, 830. <https://doi.org/10.2135/cropsci1985.0011183X002500050025x>
- Jones, R.M., Mather, K., 1958. Interaction of Genotype and Environment in Continuous Variation: II. Analysis. *Biometrics* 14, 489. <https://doi.org/10.2307/2527515>
- Josse, J., Chavent, M., Liquet, B., Husson, F., 2012. Handling Missing Values with Regularized Iterative Multiple Correspondence Analysis. *J. Classif.* 29, 91–116. <https://doi.org/10.1007/s00357-012-9097-0>
- Josse, J., Husson, F., 2016. missMDA : A Package for Handling Missing Values in Multivariate Data Analysis. *J. Stat. Softw.* 70. <https://doi.org/10.18637/jss.v070.i01>
- Kendall, M.G., Stuart, A., 1979. *The Advanced Theory of Statistics*, 4th ed. Griffin, London.
- Kirby, E.J., Appleyard, M., National Agricultural Centre (Great Britain), 1987. *Cereal development guide* (2nd ed), 2de ed, Arable Unit. Arable Unit, National, Agricultural Centre, Kenilworth, England.
- Kirby, E.J.M., 1988. Analysis of leaf, stem and ear growth in wheat from terminal spikelet stage to anthesis. *Field Crops Res.* 18, 127–140. [https://doi.org/10.1016/0378-4290\(88\)90004-4](https://doi.org/10.1016/0378-4290(88)90004-4)
- Klepper, B., Belford, R.K., Rickman, R.W., 1984. Root and Shoot Development in Winter Wheat. *Agron. J.* 76, 117. <https://doi.org/10.2134/agronj1984.00021962007600010029x>
- Kobata, T., Palta, J.A., Turner, N.C., 1992. Rate of Development of Postanthesis Water Deficits and Grain Filling of Spring Wheat. *Crop Sci.* 32, 1238. <https://doi.org/10.2135/cropsci1992.0011183X003200050035x>
- Korte, A., Vilhjálmsson, B.J., Segura, V., Platt, A., Long, Q., Nordborg, M., 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat. Genet.* 44, 1066–1071. <https://doi.org/10.1038/ng.2376>
- Kumar Tewari, A., Charan Tripathy, B., 1998. Temperature-Stress-Induced Impairment of Chlorophyll Biosynthetic Reactions in Cucumber and Wheat. *Plant Physiol.* 117, 851–858. <https://doi.org/10.1104/pp.117.3.851>
- Lacaze, X., Hayes, P.M., Korol, A., 2009. Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. *Heredity* 102, 163–173.
- Lacaze, X., Roumet, P., 2004. Environment characterisation for the interpretation of environmental effect and genotype x environment interaction. *Theor. Appl. Genet.* 109, 1632–1640. <https://doi.org/10.1007/s00122-004-1786-6>
- Langer, S.M., Longin, C.F.H., Wärschum, T., 2014. Flowering time control in European winter wheat. *Front. Plant Sci.* 5. <https://doi.org/10.3389/fpls.2014.00537>
- Lantican, M.A., Pingali, P.L., Rajaram, S., 2001. Growth in Wheat Yield Potential in Marginal Environments. Spons. CIMMYT Wheat Program Or. State Univ. USAID AgriPro 73–79.
- Le Bris, X., Soenen, B., Laberdesque, M., Maunas, M., Gouache, D., Lorgeou, J., Cohan, J., Laurent, F., Bouthier, A., Garcia, C., 2016. “CHN”, a crop model to add value to phenotyping and approach genetic variation for RUE and WUE. Presented at the International Symposium on “Crop Modelling for Agriculture and Food Security under Global Change,” Berlin, Germany.
- Le Couviour, F., Faure, S., Poupard, B., Flodrops, Y., Dubreuil, P., Praud, S., 2011. Analysis of genetic structure in a panel of elite wheat varieties and relevance for association mapping. *Theor. Appl. Genet.* 123, 715–727. <https://doi.org/10.1007/s00122-011-1621-9>
- Le Gouis, J., Béghin, D., Heumez, E., Pluchard, P., 2000. Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. *Eur. J. Agron.* 12, 163–173. [https://doi.org/10.1016/S1161-0301\(00\)00045-9](https://doi.org/10.1016/S1161-0301(00)00045-9)
- Le Gouis, J., Bordes, J., Ravel, C., Heumez, E., Faure, S., Praud, S., Galic, N., Remoué, C., Balfourier, F., Allard, V., Rousset, M., 2012. Genome-wide association analysis to identify chromosomal regions determining components of earliness in wheat. *Theor. Appl. Genet.* 124, 597–611. <https://doi.org/10.1007/s00122-011-1732-3>
- Lebon, E., Dumas, V., Pieri, P., Schultz, H., 2003. Modelling the seasonal dynamics of the soil water balance of vineyards. *Funct. Plant Biol.* 30, 699–710.
- Lecomte, C., 2005. L'évaluation expérimentale des innovations variétales. Proposition d'outils d'analyse de l'interaction génotype-milieu adaptés à la diversité des besoins et des contraintes des acteurs de la filière semences. INAPG (AgroParisTech).

- Levitt, J., 1980. Responses of plants to environmental stresses. Vol. 2: Water, radiation, salt, and other stresses, 2. ed. ed, Physiological ecology. Academic Press, New York, NY.
- Li, H., Su, G., Jiang, L., Bao, Z., 2017. An efficient unified model for genome-wide association studies and genomic selection. *Genet. Sel. Evol.* 49. <https://doi.org/10.1186/s12711-017-0338-x>
- Lim, P.O., Kim, H.J., Gil Nam, H., 2007a. Leaf Senescence. *Annu. Rev. Plant Biol.* 58, 115–136. <https://doi.org/10.1146/annurev.arplant.57.032905.105316>
- Lim, P.O., Kim, H.J., Gil Nam, H., 2007b. Leaf Senescence. *Annu. Rev. Plant Biol.* 58, 115–136. <https://doi.org/10.1146/annurev.arplant.57.032905.105316>
- Lippert, C., Listgarten, J., Liu, Y., Kadie, C.M., Davidson, R.I., Heckerman, D., 2011a. FaST linear mixed models for genome-wide association studies. *Nat. Methods* 8, 833–835. <https://doi.org/10.1038/nmeth.1681>
- Lippert, C., Listgarten, J., Liu, Y., Kadie, C.M., Davidson, R.I., Heckerman, D., 2011b. FaST linear mixed models for genome-wide association studies. *Nat. Methods* 8, 833–835. <https://doi.org/10.1038/nmeth.1681>
- Listgarten, J., Lippert, C., Heckerman, D., 2013. FaST-LMM-Select for addressing confounding from spatial structure and rare variants. *Nat. Genet.* 45, 470.
- Liu, B., Asseng, S., Müller, C., Ewert, F., Elliott, J., Lobell, D.B., Martre, P., Ruane, A.C., Wallach, D., Jones, J.W., Rosenzweig, C., Aggarwal, P.K., Alderman, P.D., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A., Deryng, D., Sanctis, G.D., Doltra, J., Fereres, E., Folberth, C., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Kimball, B.A., Koehler, A.-K., Kumar, S.N., Nendel, C., O’Leary, G.J., Olesen, J.E., Ottman, M.J., Palosuo, T., Prasad, P.V.V., Priesack, E., Pugh, T.A.M., Reynolds, M., Rezaei, E.E., Rötter, R.P., Schmid, E., Semenov, M.A., Shcherbak, I., Stehfest, E., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wall, G.W., Wang, E., White, J.W., Wolf, J., Zhao, Z., Zhu, Y., 2016. Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nat. Clim. Change* 6, 1130–1136. <https://doi.org/10.1038/nclimate3115>
- Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011. Climate Trends and Global Crop Production Since 1980. *Science* 333, 616–620. <https://doi.org/10.1126/science.1204531>
- Löffler, C.M., Wei, J., Fast, T., Gogerty, J., Langton, S., Bergman, M., Merrill, B., Cooper, M., 2005. Classification of Maize Environments Using Crop Simulation and Geographic Information Systems. *Crop Sci.* 45, 1708. <https://doi.org/10.2135/cropsci2004.0370>
- Long, S.P., Ainsworth, E.A., Rogers, A., Ort, D.R., 2004. Rising atmospheric carbon dioxide: Plants face the future. *Annu. Rev. Plant Biol.* 55, 591–628. <https://doi.org/10.1146/annurev.arplant.55.031903.141610>
- Lopes, M.S., Reynolds, M.P., 2012. Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. *J. Exp. Bot.* 63, 3789–3798. <https://doi.org/10.1093/jxb/ers071>
- Lucas, M.E., Hoad, S.P., Russell, G., Bingham, I.J., 2000. Management of cereal root systems (No. 43).
- Ly, D., Huet, S., Gauffreteau, A., Rincet, R., Touzy, G., Mini, A., Jannink, J.-L., Cormier, F., Paux, E., Lafarge, S., Le Gouis, J., Charmet, G., 2018. Whole-genome prediction of reaction norms to environmental stress in bread wheat (*Triticum aestivum* L.) by genomic random regression. *Field Crops Res.* 216, 32–41. <https://doi.org/10.1016/j.fcr.2017.08.020>
- Malo, N., Libiger, O., Schork, N.J., 2008. Accommodating Linkage Disequilibrium in Genetic-Association Analyses via Ridge Regression. *Am. J. Hum. Genet.* 82, 375–385. <https://doi.org/10.1016/j.ajhg.2007.10.012>
- Malosetti, M., Ribaut, J.-M., van Eeuwijk, F.A., 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* 4. <https://doi.org/10.3389/fphys.2013.00044>
- Malosetti, M., Ribaut, J.M., Vargas, M., Crossa, J., van Eeuwijk, F.A., 2008. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* 161, 241–257. <https://doi.org/10.1007/s10681-007-9594-0>
- Malosetti, M., Voltas, J., Romagosa, I., Ullrich, S.E., Van Eeuwijk, F.A., 2004. Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* 137, 139–145.
- Manès, Y., Gomez, H.F., Puhl, L., Reynolds, M., Braun, H.J., Trethowan, R., 2012. Genetic Yield Gains of the CIMMYT International Semi-Arid Wheat Yield Trials from 1994 to 2010. *Crop Sci.* 52, 1543. <https://doi.org/10.2135/cropsci2011.10.0574>
- Manschadi, A., Manske, G., Vlek, P., 2013. Root Architecture and Resource Acquisition: Wheat as a Model Plant, in: Beeckman, T. (Ed.), *Plant Roots*. CRC Press, pp. 22-1-22–18. <https://doi.org/10.1201/b14550-27>

- Martin, J.H., Waldren, R.P., Stamp, D.L., 2006. Principles of field crop production, 4. ed. ed. Pearson/Prentice Hall, Upper Saddle River, NJ.
- Martre, P., Jamieson, P.D., Semenov, M.A., Zyskowski, R.F., Porter, J.R., Triboi, E., 2006. Modelling protein content and composition in relation to crop nitrogen dynamics for wheat. *Eur. J. Agron.* 138–154.
- Mascarenhas, J.P., Crone, D.E., 1996. Pollen and the heat shock response. *Sex. Plant Reprod.* 9, 370–374. <https://doi.org/10.1007/BF02441959>
- Mathews, K.L., Chapman, S.C., Trethowan, R., Singh, R.P., Crossa, J., Pfeiffer, W., van Ginkel, M., DeLacy, I., 2006. Global Adaptation of Spring Bread and Durum Wheat Lines Near-Isogenic for Major Reduced Height Genes. *Crop Sci.* 46, 603–613. <https://doi.org/10.2135/cropsci2005.05-0056>
- Mathews, K.L., Malosetti, M., Chapman, S., McIntyre, L., Reynolds, M., Shorter, R., van Eeuwijk, F., 2008. Multi-environment QTL mixed models for drought stress adaptation in wheat. *Theor. Appl. Genet.* 117, 1077–1091. <https://doi.org/10.1007/s00122-008-0846-8>
- Meyer, K., 2009. Factor-analytic models for genotype \times environment type problems and structured covariance matrices. *Genet. Sel. Evol.* 41, 21. <https://doi.org/10.1186/1297-9686-41-21>
- Meynard, J.-M., 1997. Which crop models for decision support in crop management? Example of the DECIBLE system. In: Proceeding of the INRA-KCW Workshop on DSS, Laon.
- Meynard, J.M., Sebillotte, M., 1994. L’elaboration du rendement du ble, base pour l’etude des autres cereales a talles, in: L’elaboration Du Rendement Des Principales Cultures Annuelles. L. Combe (Editeur), D. Picard (Editeur), Paris, FRA, pp. 31–53.
- Mezmouk, S., Dubreuil, P., Bosio, M., Décousset, L., Charcosset, A., Praud, S., Mangin, B., 2011. Effect of population structure corrections on the results of association mapping tests in complex maize diversity panels. *Theor. Appl. Genet.* 122, 1149–1160. <https://doi.org/10.1007/s00122-010-1519-y>
- Millet, E., Welcker, C., Kruijer, W., Negro, S., Nicolas, S., Praud, S., Ranc, N., Presterl, T., Tuberosa, R., Bedo, Z., Draye, X., Usadel, B., Charcosset, A., van Eeuwijk, F., Tardieu, F., Coupel-Ledru, A., Bauland, C., 2016. Genome-wide analysis of yield in Europe: allelic effects as functions of drought and heat scenarios. *Plant Physiol.* pp.00621.2016. <https://doi.org/10.1104/pp.16.00621>
- Miralles, D.J., Katz, S.D., Colloca, A., Slafer, G.A., 1998. Floret development in near isogenic wheat lines differing in plant height. *Field Crops Res.* 59, 21–30. [https://doi.org/10.1016/S0378-4290\(98\)00103-8](https://doi.org/10.1016/S0378-4290(98)00103-8)
- Miralles, D.J., Slafer, G.A., Richards, R.A., 2003. Influence of historic photoperiod during stem elongation on the number of fertile florets in wheat. *J. Agric. Sci.* 141, 155–158. <https://doi.org/10.1017/S0021859603003551>
- Möhring, J., Piepho, H.-P., 2009. Comparison of Weighting in Two-Stage Analysis of Plant Breeding Trials. *Crop Sci.* 49, 1977. <https://doi.org/10.2135/cropsci2009.02.0083>
- Monteith, J.L., 1977. Climate and the Efficiency of Crop Production in Britain [and Discussion]. *Philos. Trans. R. Soc. B Biol. Sci.* 281, 277–294. <https://doi.org/10.1098/rstb.1977.0140>
- Morgan, J., 1991. A Gene Controlling Differences in Osmoregulation in Wheat. *Funct. Plant Biol.* 18, 249. <https://doi.org/10.1071/PP9910249>
- Moss, G.I., Downey, L.A., 1971. Influence of Drought Stress on Female Gametophyte Development in Corn (*Zea mays* L.) and Subsequent Grain Yield1. *Crop Sci.* 11, 368. <https://doi.org/10.2135/cropsci1971.0011183X001100030017x>
- Muchow, R.C., Cooper, M., Hammer, G.L., 1996. Characterizing environmental challenges using models. *Plant Adapt. Crop Improv.* CAB International, Wallingford, 349–364.
- Muchow, R.C., Sinclair, T.R., 1986. Water and nitrogen limitations in soybean grain production II. Field and model analyses. *Field Crops Res.* 15, 143–156. [https://doi.org/10.1016/0378-4290\(86\)90083-3](https://doi.org/10.1016/0378-4290(86)90083-3)
- Munns, R., Weir, R., 1981. Contribution of Sugars to Osmotic Adjustment in Elongating and Expanded Zones of Wheat Leaves During Moderate Water Deficits at Two Light Levels. *Funct. Plant Biol.* 8, 93. <https://doi.org/10.1071/PP9810093>
- Murtagh, F., Legendre, P., 2014. Ward’s Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm. *J. Classif.* 31, 274–295. <https://doi.org/10.1007/s00357-014-9161-z>
- Mwadingeni, L., Shimelis, H., Rees, D.J.G., Tsilo, T.J., 2017. Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. *PLOS ONE* 12, e0171692. <https://doi.org/10.1371/journal.pone.0171692>
- Nicolas, M., Gleadow, R., Dalling, M., 1984. Effects of Drought and High Temperature on Grain Growth in Wheat. *Funct. Plant Biol.* 11, 553. <https://doi.org/10.1071/PP9840553>

- Nicolas, M.E., Gleadow, R.M., Dalling, M.J., 1985. Effect of Post-anthesis Drought on Cell Division and Starch Accumulation in Developing Wheat Grains. *Ann. Bot.* 55, 433–444. <https://doi.org/10.1093/oxfordjournals.aob.a086922>
- Nooden, L.D., 2014. *Senescence and Aging in Plants*. Elsevier Science, Saint Louis.
- Ober, E.S., Clark, C.J.A., Perry, A., 2010. Traits related to genotypic differences in effective water use and drought tolerance in UK winter wheat. *Asp. Appl. Biol.* 105, 13–22.
- Ober, E.S., Setter, T.L., Madison, J.T., Thompson, J.F., Shapiro, P.S., 1991. Influence of Water Deficit on Maize Endosperm Development: Enzyme Activities and RNA Transcripts of Starch and Zein Synthesis, Abscisic Acid, and Cell Division. *Plant Physiol.* 97, 154–164.
- Ogren, W.L., 1984. Photorespiration: Pathways, Regulation, and Modification. *Annu. Rev. Plant Physiol.* 35, 415–442. <https://doi.org/10.1146/annurev.pp.35.060184.002215>
- Olivares-Villegas, J.J., Reynolds, M.P., McDonald, G.K., 2007. Drought-adaptive attributes in the Seri/Babax hexaploid wheat population. *Funct. Plant Biol.* 34, 189. <https://doi.org/10.1071/FP06148>
- Oury, F.-X., Godin, C., Mailliard, A., Chassin, A., Gardet, O., Giraud, A., Heumez, E., Morlais, J.-Y., Rolland, B., Rousset, M., Trottet, M., Charmet, G., 2012. A study of genetic progress due to selection reveals a negative effect of climate change on bread wheat yield in France. *Eur. J. Agron.* 40, 28–38. <https://doi.org/10.1016/j.eja.2012.02.007>
- Pachauri, R.K., Mayer, L., Intergovernmental Panel on Climate Change (Eds.), 2015. *Climate change 2014: synthesis report*. Intergovernmental Panel on Climate Change, Geneva, Switzerland.
- Palta, J.A., Kobata, T., Turner, N.C., Fillery, I.R., 1994. Remobilization of Carbon and Nitrogen in Wheat as Influenced by Postanthesis Water Deficits. *Crop Sci.* 34, 118. <https://doi.org/10.2135/cropsci1994.0011183X003400010021x>
- Passioura, J.B., 1983. Roots and drought resistance. *Agric. Water Manag.* 7, 265–280. [https://doi.org/10.1016/0378-3774\(83\)90089-6](https://doi.org/10.1016/0378-3774(83)90089-6)
- Passioura, J.B., 1977. Grain yield, harvest index, and water use of wheat. *J Aust Inst Agric Sci* 43, 117–120.
- Patterson, N., Price, A.L., Reich, D., 2006. Population Structure and Eigenanalysis. *PLOS Genet.* 2, 1–20. <https://doi.org/10.1371/journal.pgen.0020190>
- Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Comput. Geo- Sci.* 683–691.
- Peltonen-Sainio, Pirjo, Jauhiainen, L., Laurila, I.P., 2009. Cereal yield trends in northern European conditions: Changes in yield potential and its realisation. *Field Crops Res.* 110, 85–90. <https://doi.org/10.1016/j.fcr.2008.07.007>
- Peltonen-Sainio, P., Jauhiainen, L., Laurila, I.P., 2009. Cereal yield trends in northern European conditions: changes in yield potential and its realisation. *Field Crops Res.* 110, 85–90.
- Peng, J., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flinham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D., Harberd, N.P., 1999. ‘Green revolution’ genes encode mutant gibberellin response modulators. *Nature* 400, 256–261. <https://doi.org/10.1038/22307>
- Peterson, R.F., 1965. *Wheat: botany, cultivation and utilisation*, Leonard Hill. ed. London.
- Piepho, H.P., 1998. Empirical best linear unbiased prediction in cultivar trials using factor analytic variance-covariance structures. *Theor Appl Genet* 195–201.
- Piepho, H.-P., 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* 761–766.
- Pinto, R.S., Lopes, M.S., Collins, N.C., Reynolds, M.P., 2016. Modelling and genetic dissection of staygreen under heat stress. *Theor. Appl. Genet.* 129, 2055–2074. <https://doi.org/10.1007/s00122-016-2757-4>
- Pinto, R.S., Reynolds, M.P., 2015. Common genetic basis for canopy temperature depression under heat and drought stress associated with optimized root distribution in bread wheat. *Theor. Appl. Genet.* 128, 575–585. <https://doi.org/10.1007/s00122-015-2453-9>
- Pinto, R.S., Reynolds, M.P., Mathews, K.L., McIntyre, C.L., Olivares-Villegas, J.-J., Chapman, S.C., 2010. Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. *Theor. Appl. Genet.* 121, 1001–1021. <https://doi.org/10.1007/s00122-010-1351-4>
- Porter, J., Gawith, M., 1999. Temperatures and the growth and development of wheat: A review. *Eur. J. Agron.* 23–36. [https://doi.org/10.1016/S1161-0301\(98\)00047-1](https://doi.org/10.1016/S1161-0301(98)00047-1)

- Postic, F., 2016. Test et apports d'outils de phénotypage racinaires directs (imagerie des racines) et indirects (méthode électrique capacitive) pour une utilisation en sélection variétale au champ: application au blé. Avignon.
- Prasad, P.V.V., Pisipati, S.R., Momčilović, I., Ristic, Z., 2011. Independent and Combined Effects of High Temperature and Drought Stress During Grain Filling on Plant Yield and Chloroplast EF-Tu Expression in Spring Wheat: Effects of Temperature and Drought on Wheat Plants. *J. Agron. Crop Sci.* 197, 430–441. <https://doi.org/10.1111/j.1439-037X.2011.00477.x>
- Prasad, P.V.V., Pisipati, S.R., Ristic, Z., Bukovnik, U., Fritz, A.K., 2008. Impact of Nighttime Temperature on Physiology and Growth of Spring Wheat. *Crop Sci.* 48, 2372. <https://doi.org/10.2135/cropsci2007.12.0717>
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155, 945–959.
- Pritchard, J.K., Wen, X., Falush, D., 2010. Documentation for structure software: Version 2.3 39.
- R Development Core Team, 2011. , R: A Language and Environment for Statistical Computing. Vienna, Austria : the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at <http://www.R-project.org/>.
- R. Wheeler, T., D. Hong, T., H. Ellis, R., R. Batts, G., Morison, J., Hadley, P., 1996. The duration and rate of grain growth, and harvest index, of wheat (*Triticum aestivum* L.) in response to temperature and CO₂. *J. Exp. Bot. - J EXP BOT* 47, 623–630.
- Rajaram, S., van Ginkel, M., Fischer, R.A., Post, A., Juarez, C., 1995. CIMMYT's Wheat Breeding Mega-Environments (ME) 7.
- Rawson, H., Bagga, A., 1979. Influence of Temperature Between Floral Initiation and Flag Leaf Emergence on Grain Number in Wheat. *Funct. Plant Biol.* 6, 391. <https://doi.org/10.1071/PP9790391>
- Rebetzke, G.J., Condon, A.G., Richards, R.A., Farquhar, G.D., 2002. Selection for Reduced Carbon Isotope Discrimination Increases Aerial Biomass and Grain Yield of Rainfed Bread Wheat. *Crop Sci.* 42, 739. <https://doi.org/10.2135/cropsci2002.0739>
- Reynolds, Matthew, Foulkes, M.J., Slafer, G.A., Berry, P., Parry, M.A.J., Snape, J.W., Angus, W.J., 2009. Raising yield potential in wheat. *J. Exp. Bot.* 60, 1899–1918. <https://doi.org/10.1093/jxb/erp016>
- Reynolds, M., Manes, Y., Izanloo, A., Langridge, P., 2009. Phenotyping approaches for physiological breeding and gene discovery in wheat. *Ann. Appl. Biol.* 155, 309–320. <https://doi.org/10.1111/j.1744-7348.2009.00351.x>
- Reynolds, M.P., Rebetzke, G.J., Pellegrinesci, A., Trethowan, R., 2006. Drought adaptation in wheat. In *Drought Tolerance in Cereals*, in: *Drought Tolerance in Cereals*. J.M. Ribaut, New York, pp. 402–436.
- Reynolds, M.P., Trethowan, R.M., 2007. Physiological Interventions in Breeding for Adaptation to Abiotic Stress, in: Spiertz, J.H.J., Struik, P.C., Laar, H.H.V. (Eds.), *Scale and Complexity in Plant Systems Research*. Springer Netherlands, Dordrecht, pp. 129–146. https://doi.org/10.1007/1-4020-5906-X_11
- Richard, C., Christopher, J., Chenu, K., Borrell, A., Christopher, M., Hickey, L., 2018. Selection in Early Generations to Shift Allele Frequency for Seminal Root Angle in Wheat. *Plant Genome* 11, 0. <https://doi.org/10.3835/plantgenome2017.08.0071>
- Richards, R., Rawson, H., Johnson, D., 1986. Glaucousness in Wheat: Its Development and Effect on Water-use Efficiency, Gas Exchange and Photosynthetic Tissue Temperatures*. *Funct. Plant Biol.* 13, 465. <https://doi.org/10.1071/PP9860465>
- Rimbert, H., Darrier, B., Navarro, J., Kitt, J., Choulet, F., Leveugle, M., Duarte, J., Rivière, N., Eversole, K., on behalf of The International Wheat Genome Sequencing Consortium, Le Gouis, J., on behalf The BreedWheat Consortium, Davassi, A., Balfourier, F., Le Paslier, M.-C., Berard, A., Brunel, D., Feuillet, C., Poncet, C., Sourdille, P., Paux, E., 2018. High throughput SNP discovery and genotyping in hexaploid wheat. *PLOS ONE* 13, e0186329. <https://doi.org/10.1371/journal.pone.0186329>
- Rincent, R., Moreau, L., Monod, H., Kuhn, E., Melchinger, A.E., Malvar, R.A., Moreno-Gonzalez, J., Nicolas, S., Madur, D., Combes, V., Dumas, F., Altmann, T., Brunel, D., Ouzunova, M., Flament, P., Dubreuil, P., Charcosset, A., Mary-Huard, T., 2014. Recovering Power in Association Mapping Panels with Variable Levels of Linkage Disequilibrium. *Genetics* 197, 375–387. <https://doi.org/10.1534/genetics.113.159731>
- Rodríguez-Álvarez, M.X., Boer, M.P., van Eeuwijk, F.A., Eilers, P.H., 2016. Spatial Models for Field Trials. *ArXiv Prepr. ArXiv160708255*.
- Rodriguez-Alvarez, M.X., Boer, M.P., van Eeuwijk, F.A., Eilers, P.H.C., 2018. Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spat. Stat.* 23, 52–71.

- Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Gonzalez Parez, L., Crossa, J., Reynolds, M., Singh, R., 2016. Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. *G3* 58 GenesGenomesGenetics. <https://doi.org/10.1534/g3.116.032888>
- Sadras, V.O., Hayman, P.T., Rodriguez, D., Monjardino, M., Bielich, M., Unkovich, M., Mudge, B., Wang, E., 2016. Interactions between water and nitrogen in Australian cropping systems: physiological, agronomic, economic, breeding and modelling perspectives. *Crop Pasture Sci.* <https://doi.org/10.1071/CP16027>
- Sadras, V.O., Lake, L., Chenu, K., McMurray, L.S., Leonforte, A., 2012a. Water and thermal regimes for field pea in Australia and their implications for breeding. *Crop Pasture Sci.* 63, 33. <https://doi.org/10.1071/CP11321>
- Sadras, V.O., Lake, L., Chenu, K., McMurray, L.S., Leonforte, A., 2012b. Water and thermal regimes for field pea in Australia and their implications for breeding 63, 33–44. <https://doi.org/10.1071/cp11321>
- Saini, H.S., Aspinall, D., 1982. Abnormal Sporogenesis in Wheat (*Triticum aestivum* L.) Induced by Short Periods of High Temperature. *Ann. Bot.* 49, 835–846. <https://doi.org/10.1093/oxfordjournals.aob.a086310>
- Saini, H.S., Aspinall, D., 1981. Effect of Water Deficit on Sporogenesis in Wheat (*Triticum aestivum* L.). *Ann. Bot.* 48, 623–633. <https://doi.org/10.1093/oxfordjournals.aob.a086170>
- Savin, R., Slafer, G.A., Cossani, C.M., Abeledo, L.G., Sadras, V.O., 2015. Cereal yield in Mediterranean-type environments: challenging the paradigms on terminal drought, the adaptability of barley vs wheat and the role of nitrogen fertilization, in: *Crop Physiology*. Elsevier, pp. 141–158. <https://doi.org/10.1016/B978-0-12-417104-6.00007-8>
- Scott, R.A., Milliken, G.A., 1993. A SAS Program for Analyzing Augmented Randomized Complete-Block Designs. *Crop Sci.* 33, 865. <https://doi.org/10.2135/cropsci1993.0011183X003300040046x>
- Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q., Nordborg, M., 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat. Genet.* 44, 825–830. <https://doi.org/10.1038/ng.2314>
- Semenov, M.A., Shewry, P.R., 2011. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. *Sci. Rep.* 1. <https://doi.org/10.1038/srep00066>
- Sharkey, T.D., 1997. The Physiology of Plants Under Stress: Abiotic Factors. Erik T. Nilsen , David M. Orcutt. *Q. Rev. Biol.* 72, 476–477. <https://doi.org/10.1086/419999>
- Sheoran, I.S., Saini, H.S., 1996. Drought-induced male sterility in rice: Changes in carbohydrate levels and enzyme activities associated with the inhibition of starch accumulation in pollen. *Sex. Plant Reprod.* 9, 161–169. <https://doi.org/10.1007/BF02221396>
- Sheoran, S., Malik, R., Narwal, S., Tyagi, B., Mittal, V., Kharub, A.S., 2015. Genetic and molecular dissection of drought tolerance in wheat and barley. *J. Wheat Res.* 14.
- Shewry, P.R., 2009. Wheat. *J. Exp. Bot.* 60, 1537–1553. <https://doi.org/10.1093/jxb/erp058>
- Shirdelmoghanloo, H., Taylor, J.D., Lohraseb, I., Rabie, H., Brien, C., Timmins, A., Martin, P., Mather, D.E., Emebiri, L., Collins, N.C., 2016. A QTL on the short arm of wheat (*Triticum aestivum* L.) chromosome 3B affects the stability of grain weight in plants exposed to a brief heat shock early in grain filling. *BMC Plant Biol.* 16. <https://doi.org/10.1186/s12870-016-0784-6>
- Simmons, S.R., Crookston, R.K., 1979. Rate and Duration of Growth of Kernels Formed at Specific Florets in Spikelets of Spring Wheat1. *Crop Sci.* 19, 690. <https://doi.org/10.2135/cropsci1979.0011183X001900050036x>
- Simon, J.C., Lemaire, G., 1987. Tillering and leaf area index in grasses in the vegetative phase. *Grass Forage Sci.* 42, 373–380. <https://doi.org/10.1111/j.1365-2494.1987.tb02127.x>
- Slafer, G.A., Andrade, F.H., 1993. Physiological attributes related to the generation of grain yield in bread wheat cultivars released at different eras. *Field Crops Res.* 31, 351–367. [https://doi.org/10.1016/0378-4290\(93\)90073-V](https://doi.org/10.1016/0378-4290(93)90073-V)
- Smith, A., Cullis, B., Thompson, R., 2001. Analyzing Variety by Environment Data Using Multiplicative Mixed Models and Adjustments for Spatial Field Trend. *Biometrics* 57, 1138–1147.
- Smith, A.B., Cullis, B.R., Thompson, R., 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.* 143, 449. <https://doi.org/10.1017/S0021859605005587>
- Snape, J., Butterworth, K., Whitechurch, E., Worland, A.J., 2001. Waiting for Fine Times: Genetics of Flowering Time in Wheat, in: Bedö, Z., Láng, L. (Eds.), *Wheat in a Global Environment*. Springer Netherlands, Dordrecht, pp. 67–74. https://doi.org/10.1007/978-94-017-3674-9_7

- Sofield, I., Evans, L., Cook, M., Wardlaw, I., 1977. Factors Influencing the Rate and Duration of Grain Filling in Wheat. *Aust. J. Plant Physiol.* 4, 785. <https://doi.org/10.1071/PP9770785>
- Sonder, K., 2018. Global map of wheat mega-environments.
- Spiertz, J.H.J., Hamer, R.J., Xu, H., Primo-Martin, C., Don, C., van der Putten, P.E.L., 2006. Heat stress in wheat (*Triticum aestivum* L.): Effects on grain growth and quality traits. *Eur. J. Agron.* 25, 89–95. <https://doi.org/10.1016/j.eja.2006.04.012>
- Stephens, M., 2013. A Unified Framework for Association Analysis with Multiple Related Phenotypes. *PLoS ONE* 8, e65245. <https://doi.org/10.1371/journal.pone.0065245>
- Steudle, E., Frensch, J., 1996. Water transport in plants: Role of the apoplast. *Plant Soil* 187, 67–79. <https://doi.org/10.1007/BF00011658>
- Stocker, T.F., Qin, D., Plattner, G.-K., et al. (eds), 2013. *Climate Change 2013: The Physical Science Basis*. IPCC, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Stone, P., Nicolas, M., 1994. Wheat Cultivars Vary Widely in Their Responses of Grain Yield and Quality to Short Periods of Post-Anthesis Heat Stress. *Funct. Plant Biol.* 21, 887–900. <https://doi.org/10.1071/PP9940887>
- Stone, P.J., Nicolas, M.E., 1994. Wheat cultivars vary widely in their responses of grain yield and quality to short periods of post-anthesis heat stress. *Aust. J. Plant Physiol.* 887–900.
- Streck, N.A., 2005. Climate change and agroecosystems: the effect of elevated atmospheric CO₂ and temperature on crop growth, development, and yield. *Ciênc. Rural* 35, 730–740. <https://doi.org/10.1590/S0103-84782005000300041>
- Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P., Reynolds, M.P., 2015. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor. Appl. Genet.* 128, 353–363. <https://doi.org/10.1007/s00122-014-2435-3>
- Sukumaran, S., Lopes, M., Dreisigacker, S., Reynolds, M., 2018. Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theor. Appl. Genet.* 131, 985–998. <https://doi.org/10.1007/s00122-017-3037-7>
- Suzuki, N., Rivero, R.M., Shulaev, V., Blumwald, E., Mittler, R., 2014. Abiotic and biotic stress combinations. *New Phytol.* 203, 32–43. <https://doi.org/10.1111/nph.12797>
- Suzuki, R., Shimodaira, H., 2006. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22, 1540–1542. <https://doi.org/10.1093/bioinformatics/btl117>
- Tahir, I.S.A., Nakata, N., 2005. Remobilization of Nitrogen and Carbohydrate from Stems of Bread Wheat in Response to Heat Stress during Grain Filling. *J. Agron. Crop Sci.* 191, 106–115. <https://doi.org/10.1111/j.1439-037X.2004.00127.x>
- Taiz, L., Zeiger, E., 2002. *Plant physiology*, 3rd ed. ed. Sinauer Associates, Sunderland, Mass.
- Tardieu, F., 2013. Plant response to environmental conditions: assessing potential production, water demand, and negative effects of water deficit. *Front. Physiol.* 4. <https://doi.org/10.3389/fphys.2013.00017>
- Tardieu, F., 2012. Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. *J. Exp. Bot.* 63, 25–31. <https://doi.org/10.1093/jxb/err269>
- Tardieu, F., 2006. Leaf growth under water-limited conditions, in: *Drought Adaptation in Cereals*, Crop Science. Food Products Press, New York. Ribaut, J.-M., pp. 145–169.
- Tashiro, T., Wardlaw, I.F., 1990. The Effect of High Temperature at Different Stages of Ripening on Grain Set, Grain Weight and Grain Dimensions in the Semi-dwarf Wheat “Banks.” *Ann. Bot.* 65, 51–61.
- Thomas, H., Howarth, C.J., 2000. Five ways to stay green. *J. Exp. Bot.* 51, 329–337. https://doi.org/10.1093/jexbot/51.suppl_1.329
- Tian, F., Bradbury, P.J., Brown, P.J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T.R., McMullen, M.D., Holland, J.B., Buckler, E.S., 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43, 159–162. <https://doi.org/10.1038/ng.746>
- Tottman, D.R., 1987. The decimal code for the growth stages of cereals, with illustrations. *Ann. Appl. Biol.* 110, 441–454. <https://doi.org/10.1111/j.1744-7348.1987.tb03275.x>
- Trethowan, R.M., van Ginkel, M., Rajaram, S., 2002. Progress in Breeding Wheat for Yield and Adaptation in Global Drought Affected Environments. *Crop Sci.* 42, 1441. <https://doi.org/10.2135/cropsci2002.1441>
- Tricker, P.J., ElHabi, A., Schmidt, J., Fleury, D., 2018. The physiological and genetic basis of combined drought and heat tolerance in wheat. *J. Exp. Bot.* <https://doi.org/10.1093/jxb/ery081>

- Tuberosa, R., 2002. Mapping QTLs Regulating Morpho-physiological Traits and Yield: Case Studies, Shortcomings and Perspectives in Drought-stressed Maize. *Ann. Bot.* 89, 941–963. <https://doi.org/10.1093/aob/mcf134>
- Van Der Loo, M.P.J., 2010. Distribution based outlier detection for univariate data. *Discuss. Pap. 10003 Stat. Neth. Hagu.*
- van Eeuwijk, F.A., 1992. Interpreting genotype-by-environment interaction using redundancy analysis. *Theor. Appl. Genet.* 85, 89–100. <https://doi.org/10.1007/BF00223849>
- van Eeuwijk, F.A., Bink, M.C., Chenu, K., Chapman, S.C., 2010. Detection and use of QTL for complex traits in multiple environments. *Curr. Opin. Plant Biol.* 13, 193–205. <https://doi.org/10.1016/j.pbi.2010.01.001>
- van Eeuwijk, F.A., Bustos-Korts, D.V., Malosetti, M., 2016. What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype \times Environment Interactions? *Crop Sci.* 56, 2119. <https://doi.org/10.2135/cropsci2015.06.0375>
- Van Eeuwijk, F.A., Kang, M.S., Denis, J.B., 1996. Incorporating additional information on genotypes and environments in models for twoway genotype by environment tables. In: Gauch HG, Kang M (eds) *Genotype-by-environment interaction*. CRC Press, Boca Raton, pp 15–49.
- VanRaden, P.M., 2008. Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* 91, 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Vargas, M., Crossa, J., van Eeuwijk, F.A., Ramírez, M.E., Sayre, K., 1999. Using Partial Least Squares Regression, Factorial Regression, and AMMI Models for Interpreting Genotype \times Environment Interaction. *Crop Sci.* 39, 955. <https://doi.org/10.2135/cropsci1999.0011183X003900040002x>
- Vargas, M., Glaz, B., Alvarado, G., Pietragalla, J., Morgounov, A., Zelenskiy, Y., Crossa, J., 2015. Analysis and Interpretation of Interactions in Agricultural Research. *Agron. J.* 107, 748. <https://doi.org/10.2134/agronj13.0405>
- Varlet-Grancher, C., Chartier, M., Gosse, G., Bonhomme, R., 1982. Rayonnement utile pour la photosynthèse des végétaux en conditions naturelles. *Oecol Plant* 2, 189–202.
- Varshney, R.K., Bansal, K.C., Aggarwal, P.K., Datta, S.K., Craufurd, P.Q., 2011. Agricultural biotechnology for crop improvement in a variable climate: hope or hype? *Trends Plant Sci.* 16, 363–371. <https://doi.org/10.1016/j.tplants.2011.03.004>
- Verbeke, G., Molenberghs, G., 2011. *Linear mixed models for longitudinal data*, 6. [Nachdr.]. ed, Springer series in statistics. Springer, New York, NY.
- Verma, V., Foulkes, M.J., Worland, A.J., Sylvester-Bradley, R., Caligari, P.D.S., Snape, J.W., 2004. Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. *Euphytica* 135, 255–263. <https://doi.org/10.1023/B:EUPH.0000013255.31618.14>
- Wahid, A., Gelani, S., Ashraf, M., Foolad, M., 2007. Heat tolerance in plants: An overview. *Environ. Exp. Bot.* 61, 199–223. <https://doi.org/10.1016/j.envexpbot.2007.05.011>
- Wardlaw, I.F., 1974. Temperature Control of Translocation, in: In: R. L. Bielske, A. R. Ferguson and M. M. Cresswell, Eds., *Mechanism of Regulation of Plant Growth*, Royal Society New Zealand, Wellington. pp. 533–538.
- Wasson, A.P., Richards, R.A., Chatrath, R., Misra, S.C., Prasad, S.V.S., Rebetzke, G.J., Kirkegaard, J.A., Christopher, J., Watt, M., 2012. Traits and selection strategies to improve root systems and water uptake in water-limited wheat crops. *J. Exp. Bot.* 63, 3485–3498. <https://doi.org/10.1093/jxb/ers111>
- Weigand, C., 2011. *Wheat import projections towards 2050US Wheat Associates, USA.*
- Westgate, M.E., Boyer, J.S., 1985. Osmotic adjustment and the inhibition of leaf, root, stem and silk growth at low water potentials in maize. *Planta* 164, 540–549. <https://doi.org/10.1007/BF00395973>
- Wiegand, C.L., Cuellar, J.A., 1981. Duration of Grain Filling and Kernel Weight of Wheat as Affected by Temperature 1. *Crop Sci.* 21, 95. <https://doi.org/10.2135/cropsci1981.0011183X001100010027x>
- Wingler, A., Purdy, S.J., Edwards, S.-A., Chardon, F., Masclaux-Daubresse, C., 2010. QTL analysis for sugar-regulated leaf senescence supports flowering-dependent and -independent senescence pathways. *New Phytol.* 185, 420–433. <https://doi.org/10.1111/j.1469-8137.2009.03072.x>
- Xu, H., Jiang, B., Cao, Y., Zhang, Y., Zhan, X., Shen, X., Cheng, S., Lou, X., Cao, L., 2015. Detection of Epistatic and Gene-Environment Interactions Underlying Three Quality Traits in Rice Using High-Throughput Genome-Wide Data. *BioMed Res. Int.* 2015, 1–7. <https://doi.org/10.1155/2015/135782>

- Xu, Q., Paulsen, A.Q., Guikema, J.A., Paulsen, G.M., 1995. Functional and ultrastructural injury to photosynthesis in wheat by high temperature during maturation. *Environ. Exp. Bot.* 35, 43–54. [https://doi.org/10.1016/0098-8472\(94\)00030-9](https://doi.org/10.1016/0098-8472(94)00030-9)
- Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., Dubcovsky, J., 2003. Positional cloning of the wheat vernalization gene VRN1. *Proc. Natl. Acad. Sci.* 100, 6263–6268. <https://doi.org/10.1073/pnas.0937399100>
- Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S., Buckler, E.S., 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203.
- Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421.
- Zanke, C.D., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., Argillier, O., Stiewe, G., Hinze, M., Neumann, F., Eichhorn, A., Polley, A., Jaenecke, C., Ganal, M.W., Röder, M.S., 2015. Analysis of main effect QTL for thousand grain weight in European winter wheat (*Triticum aestivum* L.) by genome-wide association mapping. *Front. Plant Sci.* 6. <https://doi.org/10.3389/fpls.2015.00644>
- Zhang, J., Gizaw, S.A., Bossolini, E., Hegarty, J., Howell, T., Carter, A.H., Akhunov, E., Dubcovsky, J., 2018. Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. *Theor. Appl. Genet.* <https://doi.org/10.1007/s00122-018-3111-9>
- Zhang, L.-Y., Liu, D.-C., Guo, X.-L., Yang, W.-L., Sun, J.-Z., Wang, D.-W., Zhang, A., 2010. Genomic Distribution of Quantitative Trait Loci for Yield and Yield-related Traits in Common Wheat: Meta-QTL Analysis for Its Related Traits in Wheat. *J. Integr. Plant Biol.* 52, 996–1007. <https://doi.org/10.1111/j.1744-7909.2010.00967.x>
- Zhao, C.-X., He, M.-R., Wang, Z.-L., Wang, Y.-F., Lin, Q., 2009. Effects of different water availability at post-anthesis stage on grain nutrition and quality in strong-gluten winter wheat. *C. R. Biol.* 332, 759–764. <https://doi.org/10.1016/j.crv.2009.03.003>
- Zhao, H., Dai, T., Jiang, D., Cao, W., 2008. Effects of High Temperature on Key Enzymes Involved in Starch and Protein Formation in Grains of Two Wheat Cultivars. *J. Agron. Crop Sci.* 194, 47–54. <https://doi.org/10.1111/j.1439-037X.2007.00283.x>
- Zhao, H., Dai, T., Jing, Q., Jiang, D., Cao, W., 2007. Leaf senescence and grain filling affected by post-anthesis high temperatures in two different wheat cultivars. *Plant Growth Regul.* 51, 149–158. <https://doi.org/10.1007/s10725-006-9157-8>
- Zhao, K., Aranzana, M.J., Kim, S., Lister, C., Shindo, C., Tang, C., Toomajian, C., Zheng, H., Dean, C., Marjoram, P., Nordborg, M., 2007. An Arabidopsis Example of Association Mapping in Structured Samples. *PLoS Genet.* 3, e4. <https://doi.org/10.1371/journal.pgen.0030004>
- Zhao, K., Tung, C.-W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., McClung, A.M., Bustamante, C.D., McCouch, S.R., 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2. <https://doi.org/10.1038/ncomms1467>
- Zheng, B.S., Le Gouis, J., Leflon, M., Rong, W.Y., Laperche, A., Brancourt-Hulmel, M., 2010. Using probe genotypes to dissect QTL × environment interactions for grain yield components in winter wheat. *Theor. Appl. Genet.* 121, 1501–1517. <https://doi.org/10.1007/s00122-010-1406-6>
- Zhou, X., Stephens, M., 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* 11, 407–409. <https://doi.org/10.1038/nmeth.2848>

Annexes I

Annexe I.1 : Récapitulative de la localisation des QTL identifier dans la bibliographie

ANT : Anthesis (date d'épiaison); BM: Biomass (Biomasse); CHL: Chlorophyll (Chlorophylle); CID: Carbon isotopic discrimination (Discrimination Isotopique du Carbone); CL: Coleoptile length (Longueur du Coléoptile); CT: canopy temperature (Température de la Canopée); DSI: Drought stress index (Indice de Stress Hydrique) ; EV: early vigor (Vigueur précoce); FLG: Flag leaf glaucousness (Glaucéscence de la Feuille Drapeau); FLL: Flag leaf length (Longueur de la Feuille Drapeau); FLR: Flag leaf rolling (Enroulement de la Feuille Drapeau); FLW: Flag leaf width (Largeur de la Feuille Drapeau); GFL: Grain filling length ; GFr: grain filling rate (Taux de Remplissage du Grain); HD: Heading (Date de Floraison); HI: Harvest index (Indice de Récolte); HSI: Heat stress index (Indice de Stress Thermique); KN: kernel number (per m², plants, or spike) (Nombre de Grain par m², par plante ou par épi); KS: kernel size (Taille du grain); LA: Leaf area (Aire de la feuille); MS: Membrane stability (Stabilité de la membrane); NC: Nitrogen content (Contenu en Azote); NDVI: Normalized difference vegetative index (Indice de Végétation par Différence Normalisée); OA: osmotic adjustment (Ajustement Osmotique); Pheno : Phenology traits (traits phénologiques); PH: Plant height (Hauteur de la Plante); PL: Peduncle length (Longueur du Pédoncule); PM: Physiological maturity (Maturité Physiologique); POR: leaf porosity (Porosité de la Feuille); PTS: photosynthesis (Photosynthèse); QLTY: grain quality traits (Traits de Qualité du Grain); RT: Root architectural traits (Traits d'architecture Racinaire); SEN: senescence (Sénescence); SN: Spike number (per m² or plants) (Nombre d'épi par m² ou par plante); SPL: Spike length (Longueur de l'épi); SPN: Spikelet number (per plants or spike) (Nombre d'épillets par plante ou par épi); STL: Stem length (Longueur de la tige); TKW: Thousand kernel weight (Poids de Mille Grains) ; WSC: Water soluble carbohydrate (Glucides Solubles dans l'eau); GY: grain yield (Rendement en grains).

Type de traits = soit le trait est un organe puit, un organe source ou alors un trait qui fait la balance entre les deux types d'organes. Traits = différents traits qui sont regroupé ici, avec

Type de traits	Trait	Environnement	Chromosomes	ID
Puits	GFr	DR	- - - - - 4A - - - - -	14
		IR, DR	- 1B - 2A - - - - - 4B - 5A - - - 6A 6B - - 7B -	37
		IR, DR	- 1B - 2A - - - - - 4B - 5A - - - 6A 6B - - 7B -	37
HI	DR	DR	- 1B - 2A - - - - 3B 3D - - 4D - - - - - 7A - -	4
		DR	1A 1B - 2A - - - 3A - - 4A - 4D - - - 6A - - - - -	56
		DR	- - - 2A 2B - - - - - - - - - - - 7B -	38
		HI	- - - - - - - - - - - 5B - - - - - 7D	19
		IR, DR	1A 1B - - - - - - - 4A - - 5A 5B - - - - -	11
		IR, DR	- 1B - 2A 2B - - 3B - 4A - - 5A 5B - 6A 6B - - 7B -	25
		DR, HI	- 1B - - - - - 3B - 4A - 4D 5A 5B - - 6B - 7A - -	38
		IR, HI	- - - - 2B - 3A 3B - 4A 4B - - - - 6A - - - - -	16
KN	AMC	AMC	- - - - - - - - - - 5B - - - - - 7D	15
		DR	1A 1B - - 2B - 3A 3B - - 4B 4D - - 6A - - 6D 7A 7B -	4
		DR	- - - - 2B - - - - - - - - - 5D - - 6D - - -	8
		DR	- - - - - 4A - - - - - - - - - - -	14
		DR	- 1B - - 2B - - - - - - - 5A - - - 6B - - 7B 7D	24
		DR	- 1B - - - - - - - - - 5B - - - - - 7D	44
		DR	1A 1B - 2A - - 3A 3B - - - 5A - 5D 6A 6B - - -	56
		DR	- - - - - - - - - - - - - - - 7A - - -	39
		DR	- - - - - - - - - 4B - 5A - - - - - - -	39
		DR	- - - - - 2D - - - - - - - - - - -	51
		DR	- - - - - - - 3B 3D - 4B - - 5B - - - - -	54
		DR	- 1B - 2A - - - - - - - 4B 4D - 5B - 6A - - - - 7D	49
		DR	1A - - - 2B - - - - - - - - - - - - -	38
		HI	1A - 1D 2A 2B - - 3B 3D 4A 4B 4D 5A 5B 5D 6A 6B 6D - 7B -	10
		HI	- - - - - - - - - 4B - - - - - - - 7A - 7D	19
		HI	- - - - 2B - - - - - - - - - - 6B - - - -	53
		HI	- - - - 2B - 3A - - - - - 5A - - - - -	50
		HI	- - - - 2B - - - - - - - - - 6A - - - - -	55
		IR, DR	1A 1B 1D - - - 3B 3D - 4B - - 5B - - - 7A 7B -	11
		DR, HI	- 1B - 2A 2B 2D - 3B - 4A 4B 4D 5A 5B 5D 6A - - 7A 7B -	38
IR, DR, HI	- - - - - - - 3B 3D - - - 5A 5B - - - 7A - -	5		
IR, DR, HI	- 1B - - - - - 3B - 4A - - - 5B - - 6B - - - 26			
IR, HI	1A 1B - 2A 2B - - 3B 3D - 4B - - - - 7A - -	16		
NEC, GH	- - 1D 2A - - - 3D - - - - - 6A - - 7A - 7D	13		
KS	HI	HI	- - - - - - - - - 5A - - - - -	19
		IR, DR	- 1B - - - - 3A 3B 3D 4A - - - - - 6B - - 7B 7D	11
TKW	AMC	AMC	- - - - - - - - - - 6A - - - - 7D	15
		DR	- - 1D - 2B - - - 3D 4A - - - 5B - 6A 6B - 7A - 7D	4
		DR	- 1B - - 2B - 3A 3B - 4A 4B - - - - -	24
		DR	- - - - - - - - - - 6A - 6D - - 7D	44
DR	- 1B 1D - 2B - 3A 3B - - - - - 5B - - 6B - - 7B -	56		

PTS	DR	- 1B - 2A	- - - - 4A 4B	- - - - 6B	- - - - 8	
	HI	- - - - -	- - - - -	- - - - 7A	- - 36	
	IR	- - - - -	- - 3A 3B	- - - - 7A	- - 21	
	IR, DR, HI	- 1B - 2A	- - - 3B - 4B	- - 5B - 6A	- - 7B - 9	
RA	IR, DR	1A - - - 2B - 3A	- - - - -	- - - - 7D	18	
RL	IR, DR	- - - - 2B - - 3B 3D	- - - - 5A	- - - - 7A	- - 18	
RML	IR, DR	- 1B - - - - 3A	- - - - -	- 5B 5D - - - - 7B	- 18	
RPA	IR, DR	- - - - 2B	- - - - -	- - - - -	- - - - 18	
	IR, DR	- - - - 2B	- - 3B - 4A 4B	- - 5B - - - - -	- - 7D	18
RSA	IR, DR	- - - - 2B	- - 3B - 4A 4B	- - 5B - - - - -	- - 7D	18
RTL	IR, DR	1A 1B - - - - - 3B	- - - 4B	- - 5B 5D - - - - -	- - 7D	18
SEN	DR	- - - - 2B	- - - - -	- - - - -	- - - - -	- - 35
	HI	- - - 2A	- - 3A 3B - 4A 4B	- 5A - 5D 6A 6B - 7A 7B 7D	- 36	
	IR, DR	- - - - -	- - - - -	- - - - 6B	- - - - -	11

Source-Puit

ANT	DR	- 1B 1D - 2B	- - - - 4A - 4D 5A	- - - - -	- 7A 7B	- 20
	IR, DR, HI	- - 1D - - - - - - -	- - - - 4D 5A	- - - - -	- - - - -	- 26
DSI	DR	- - - - -	- - - 4A - - - - -	- - - - -	- - - - -	- 14
	IR, DR	- - - - -	- - - 4A - - - - -	- - - - -	- 7A 7B	- 11
	IR, DR, HI	- - - 2A	- - - - -	- 4B - - - 5B - 6A	- - - - -	- 9
GFL	HI	- - 1D - 2B - 3A 3B	- 4A 4B 4D - 5B 5D 6A - 6D	- - - - -	- 7D	10
	HI	- - - - -	- - - 5A - - - - -	- - - - -	- 7A	- 19
	HI	1A - - - - -	- - 3B - 4A 4B - 5A	- - - - -	- - - 7B	- 43
	HI	- 1B - - - - -	- - - - -	- - - - -	- - - - -	- 41
	HI	- 1B - - - - -	- - - - -	- - - 5A - - - - -	- - - - -	- 53
	DR	- - - 2A	- - - - -	- - - - -	- - - - -	- 38
	DR, HI	- 1B 1D 2A	- - - - -	- 4A 4B - 5A - - 6A 6B 6D	- 7B	- 38
	IR, DR	1A 1B - - - - -	- - 3B 3D - - - 5B	- - - - -	- 7A	- 11
	IR, DR	- 1B - - - 2B	- - - - -	- 4A 4B - 5A 5B - - - -	- 7A 7B	- 25
HD	AMC	1A - - - 2B	- - - - -	- 4B - 5A 5B	- - - - -	- 7A 7B - 3
	HI	- - - - -	- - - - -	- - - 5B - - - - -	- 7A - 7D	19
	IR	- - - - -	- - - 4A - - - - -	- - - - -	- - - - -	- 1
	DR	- - 1D - - - - -	- - - - -	- 4B - - - - -	- - - - -	- 7D 44
	DR	- - - - -	- - - - -	- - - - -	- - - 7A	- 0
	DR	- - - - -	- - - - -	- - - 5A 5B - - - 6B	- - - - -	- 0
	DR	1A - - - - 2D - 3B	- - - - -	- - - - -	- - - - -	- 0
	DR	- - - - -	- - - - -	- - - - -	- - - 7B	- 0
	DR	- 1B - - - - 3B	- - - - -	- - - - -	- 6A - - - -	- 7D 0
	HI	- 1B - - - 2B	- - - - -	- - - - -	- 7A 7B	- 40
	HI	- - - - -	- - - - -	- - - 5A - - - 6A 6B - 7A	- - - 41	
	DR, HI	- 1B - - - 2B	- - - - -	- - - - -	- 7A 7B	- 40
	DR, HI	- - - - -	- - - - -	- - - - -	- - - - -	- 41
	DR, HI	- - 1D - - - - -	- - - - -	- 4A - - - - -	- - - - -	- 7D 44
	DR, HI	- 1B - - - 2B - - 3B - 4A	- 4D 5A 5B - - - - -	- - - - -	- - - - -	- 38
	IR, DR	- - 1D 2A 2B - 3A 3B	- - - 4B - - - - -	- - - - -	- - - - -	- 7D 11
IR, DR	- 1B - - - 2B - 3A	- - - - -	- 4B - 5A - - - - -	- - - 7B	- 25	
IR, DR, HI	- - - - 2B	- - - - -	- - - 5A 5B - - - -	- 7A 7B	- 5	
NEC, GH	- - - - -	- - 3A - - - 4A	- - - - -	- - - - -	- 7A - 7D 13	
HSI	HI	- - - - -	- 3B 3D - - - - -	- 5D - - - 7A	- - 19	
OA	AMC	- - - - -	- - - - -	- - - - -	- 7A	- 23
	DR	- - - - -	- - - - -	- - - - -	- 7A	- 22
	IR, DR	- - - 2A 2B - 3A 3B	- - 4B - 5A 5B - 6A 6B	- - - - -	- - - - -	- 25
	IR, DR, HI	- - - 2A	- - - - -	- - - - -	- - - - -	- 9
PH	DR	- - 1D - - - 3A	- - - - -	- 5A - - - - -	- - - - -	- 4
	DR	- 1B 1D - 2B - 3A	- - - 4A 4B - 5A 5B	- - 6B - - - -	- - -	- 20
	DR	1A - 1D - 2B - - - -	- 4A 4B - - - 5D	- - - - -	- - - 7D	24
	DR	- - - - 2B - - - -	- 4A 4B - - - - -	- - - - -	- - - 7D	44
	DR	- 1B 1D - 2B 2D 3A	- - - - -	- - - - -	- 6A - - - 7A 7B	- 56
	DR	- - - - -	- - - - -	- - - - -	- 6A - - - -	- 48
	DR	- 1B - - - 2D	- - - - -	- - - - -	- - - - -	- 51
	DR	- - - - -	- - 3B 3D - 4B - - - -	- - - 6B - - - -	- - - -	- 54
	DR	- - - - -	- - - - -	- 4D - - - 6A 6B	- - - - -	- 49
	DR	- - - - 2B - - - - 3D	- - - - -	- 5B - - - - -	- - - - -	- 38
	DR	- - - 2A	- - - - -	- - - - -	- - - - -	- 46
	DR	- 1B - - - 2B - - 3B - 4A 4B 4D	- - - - -	- 6B - - - 7B	- 47	
	HI	- - - - -	- - - - -	- 4B - - - - -	- - - - -	- 19
	HI	- - - - -	- - - - -	- 5A - - - 6A 6B - 7A	- - - 41	
	HI	- - - - -	- - - - -	- 4A 4B - - - - -	- - - - -	- 44
	HI	- - - - -	- - - - -	- - - 4D - - - - -	- - - - -	- 38
	IR	- - - - -	- - - - -	- 4A - - - - -	- - - - -	- 1
	DR, HI	1A - - - 2B	- - - - -	- 4A 4B - - - - -	- - - - -	- 44

	DR, HI	-	1B	1D	-	2B	2D	-	3B	-	4A	4B	-	5A	-	-	-	6B	6D	-	-	-	38
	IR, DR	-	-	-	-	-	-	-	3B	-	-	-	-	-	5B	-	6A	-	-	7A	7B	-	11
	IR, DR, HI	-	-	-	-	-	-	-	-	-	-	4B	-	-	-	-	-	-	-	-	-	-	26
	NEC, GH	-	-	-	-	-	-	-	-	3D	-	-	-	5A	5B	-	-	6B	6D	7A	7B	-	6
	NEC, GH	1A	-	1D	-	-	-	3A	3B	-	4A	4B	-	5A	5B	-	6A	-	6D	7A	-	7D	13
Pheno	AMC	1A	-	-	-	2B	-	-	-	-	4A	-	-	5A	5B	-	-	-	-	7A	7B	-	3
	DR	-	-	-	2A	2B	-	3A	3B	-	-	-	-	5A	5B	5D	6A	6B	-	7A	7B	-	34
	NEC, GH	-	-	-	-	2B	-	-	-	-	-	-	-	5A	5B	5D	-	-	6D	7A	-	7D	12
	DR	-	1B	-	-	2B	-	3A	3B	-	-	-	-	-	-	-	-	-	-	7A	-	-	39
	HI	-	-	-	-	2B	-	-	-	-	-	-	-	-	-	-	-	-	-	7A	-	-	41
	IR, DR, HI	-	-	-	-	-	-	3A	3B	-	-	4B	-	5A	-	-	-	-	-	-	-	-	5
PM	HI	-	-	-	-	-	-	-	-	-	-	-	-	5B	-	-	-	-	7A	-	7D	19	
	IR, DR	-	-	-	2A	-	-	-	-	-	-	-	-	-	-	-	-	-	7A	-	-	11	
	IR, DR, HI	-	1B	-	2A	-	-	3A	-	-	4B	-	5A	-	6A	6B	6D	-	7B	-	-	5	
	IR, DR, HI	1A	-	1D	-	-	-	-	-	-	-	4D	5A	-	-	-	-	-	-	7B	-	26	
STL	IR, DR	-	-	-	-	2B	-	-	3B	-	4A	4B	-	-	-	-	-	6B	-	7A	7B	25	
WSC	DR	-	-	-	-	-	-	3A	-	-	-	-	-	-	-	-	6A	-	-	-	-	4	
	DR	1A	-	-	2A	2B	2D	-	-	-	-	-	-	-	5D	-	6B	-	-	-	7D	38	
	IR, DR	1A	1B	1D	2A	2B	-	3A	3B	3D	4A	4B	4D	5A	5B	5D	-	6B	6D	7A	7B	7D	32
	IR, DR	1A	-	1D	2A	-	-	3B	-	4A	4B	-	5A	-	-	-	-	6B	-	7A	7B	7D	37
	DR, HI	-	1B	1D	-	-	2D	-	3B	-	4A	4B	-	5A	-	-	-	6B	-	7A	7B	7D	38
	IR, DR, HI	-	1B	-	-	-	-	3A	3B	-	-	-	-	-	5B	-	-	-	-	7B	-	5	

Annexes III

Supplementary data III.1: wheat varieties tested, year of release (YR) and geographical origin.

Variety	YR	Origin	Variety	YR	Origin	Variety	YR	Origin
ACCOR	2007	France	CROUSTY	1995	France	ORCAS	2010	Germany
ACCROC	2010	France	DIALOG	2007	France	OREGRAIN	2011	France
ACIENDA	2004	France	DINOSOR	2005	France	ORVANTIS	2000	France
ACOUSTIC	2010	France	EINSTEIN	2002	UK	OXEBO	2010	France
ADAGIO	2009	France	EPHOROS	2004	France	PAJERO	1995	Belgium
ADHOC	2010	France	EPIDOC	2006	France	PAKITO	2010	France
AEROBIC	2009	France	ESPERIA	2002	Italy	PALEDOR	2005	France
ALCAZAR	2004	France	EUCLIDE	2007	France	PAPAGENO	2007	Austria
ALCHEMY	2005	UK	EUREKA	1992	France	PARADOR	2000	France
ALDRIC	2007	France	EXELCIOR	2007	France	PEPIDOR	2007	France
ALIGATOR	2010	France	EXOTIC	2005	France	PERFECTOR	2004	France
ALIXAN	2005	France	EXPERT	2007	France	PHARE	2007	France
ALIZEO	2010	France	FAIRPLAY	2011	France	PIERROT	2009	Danemark
ALLEZ_Y	2010	France	FARINELLI	2010	France	PIRENEO	2005	Austria
ALTAMIRA	2009	France	FIGARO	2011	France	PLAYER	2010	France
ALTIGO	2007	France	FLAMENKO	2010	France	PR22R20	2002	France
ALTRIA	1996	France	FLAUBERT	2010	France	PR22R58	2002	Italy
AMADOR	2010	France	FLUOR	2010	France	PREMIO	2007	France
AMBELLO	2010	France	FOLKLOR	2010	France	PREVERT	2010	France
AMBITION	2004	Danemark	FORBLANC	2009	France	PRIMO	2009	Italy
AMUNDSEN	2007	France	FRELON	2000	France	QUALITY	2002	Italy
ANDALOU	2001	France	GALACTIC	2007	France	RACINE	2010	France
ANTONIUS	2006	France	GALLANT	2007	UK	RAZZANO	2010	France
APACHE	1998	France	GALOPAIN	2009	France	RECITAL	1986	France
APRILIO	2010	France	GALPINO	2010	Spain	RENAN	1989	France
ARAMIS	2009	France	GALVANO	2008	Danemark	RIMBAUD	2010	France
AREZZO	2007	France	GARANTUS	2007	Poland	ROBIGUS	2002	UK
ARISTOTE	2010	France	GARCIA	2006	France	RONCARD	2011	France
ARKEOS	2010	France	GLADIATOR	2002	UK	ROSARIO	2004	France
ARLEQUIN	2007	France	GLASGOW	2004	UK	ROYSSAC	2002	France
ARTEDECO	2011	France	GONCOURT	2009	France	RUBISKO	2011	France
ASCOTT	2011	France	GRAINDOR	2006	France	RUSTIC	2005	Belgium
ATHLON	2010	France	HAUSSMANN	2006	France	SAINT_EX	2010	France
ATTITUDE	2009	France	HEKTO	2009	France	SAMURAI	2005	France
ATTLASS	2005	France	HEREWARD	1992	France	SANKARA	2004	France
AUBUSSON	2001	France	HISSEO	2010	Italy	SCENARIO	2010	France
AUTAN	2000	France	ILLICO	2010	France	SCOR	2009	France
AVENUE	2011	Serbia	INNOV	2005	France	SEBASTO	2007	France
AZIMUT	2004	France	INSTINCT	2006	France	SELEKT	2007	France
AZZERTI	2010	France	INTERET	2007	France	SEYRAC	2006	France
BAGOU	2007	France	IRIDIUM	2007	France	SIRTAKI	2008	Italy
BAROK	2009	France	ISENGRAIN	1997	France	SOBBEL	2009	France
BASMATI	2011	Italy	ISTABRAQ	2003	UK	SOGOOD	2006	France
BASTIDE	2002	France	JAGUAR	2009	Italy	SOISSONS	1988	France
BERGAMO	2011	France	JB_DIEGO	2006	UK	SOKAL	2010	France
BERMUDE	2007	France	KALAHARI	2010	Germany	SOLEDAD	2008	Italy
BIANCOR	2010	France	KALANGO	2002	Italy	SOLEHIO	2009	France
BOISSWATER	2007	France	KALYSTAR	2010	France	SOLLARIO	2007	France
BOKARO	2002	Italy	KARILLON	2010	France	SOLUTION	2007	France
BOLOGNA	2002	Spain	KOMETUS	2011	Germany	SOPHYTRA	2007	France
BOREGAR	2007	France	KORELI	2007	France	SORRIAL	2009	France
BUENNO	2007	France	KWS_PODIUM	2010	UK	SPONSOR	1995	France
CAMP_REMY	1980	France	LASER	2010	Italy	SWEET	2010	France
CAMPERO	2006	France	LAURIER	2011	France	SWINGGY	2009	France
CAPHORN	2000	France	LEAR	2007	UK	SY_ALTEO	2010	France
CATALAN	2002	France	LIMES	2002	France	SY_EPSON	2011	France
CCB_INGENIO	2006	Spain	LORD	2009	France	SY_MATTIS	2010	France
CELESTIN	2010	France	MANAGER	2006	France	SY_MOISSON	2011	France
CELLULE	2011	France	MARCELIN	2009	France	SY_TOLBIAC	2011	France
CENTENAIRE	2002	Belgium	MARKSMAN	2006	UK	TAPIDOR	2001	France
CEZANNE	1998	France	MERCATO	2005	France	TIMING	2010	France
CH_NARA	2009	Switzerland	MIROIR	2010	France	TOISONDOR	2004	France
CHARGER	1997	France	MOSKITO	2011	France	TRAPEZ	2009	France
CHEVALIER	2006	Austria	MUSIK	2010	France	TREMIÉ	1992	France
CLAIRE	1997	UK	NIRVANA	2001	France	TULIP	2010	France
COMPIL	2010	France	NOGAL	2006	Spain	USKI	2009	France
CORDIALE	2005	France	NUCLEO	2010	France	VALODOR	2007	France
CORVUS	1998	Belgium	OAKLEY	2006	UK	VISCOUNT	2007	UK
COURTOT	1974	France	ODYSSEE	2011	France	WARRIOR	2009	UK
CROISADE	2010	France	ORATORIO	1996	France	XI19	2001	UK

Supplementary data III.2: Trial description, crop management and weather related data

III.2A. Experiment description, location, sowing date, density and harvest date.

<i>Trial name</i>	<i>Years</i>	<i>Treatment</i>	<i>Location</i>	<i>Lat.</i>	<i>Long.</i>	<i>Alt.</i>	<i>Sowing date</i>	<i>Sowing density</i>	<i>Harvest date</i>	<i>Plot Size</i>
<i>Gre12</i>	2012	True	Greoux-Les-Bains	43.75	5.86	300	17/11/2011	350	18/07/2012	11.2
<i>Coi12</i>	2012	True	Coins	46.86	1.75	166	17/10/2011	300	25/07/2012	6
<i>Cha12</i>	2012	False	Chalon	48.95	4.42	102	22/10/2011	375	01/08/2012	10
<i>Rec12</i>	2012	False	Reclainville	48.34	1.75	151	28/10/2011	318	01/08/2012	6.75
<i>Cap12</i>	2012	False	Cappelle-en-pévèle	50.30	3.10	42	28/10/2011	280	11/08/2012	6.5
<i>Mon12</i>	2012	False	Mons	49.88	3.01	85	20/10/2011	225	26/07/2012	5.2
<i>Lou12</i>	2012	False	Louvilles	48.32	1.82	144	17/10/2011	250	04/08/2012	8.4
<i>Mau12</i>	2012	False	Maulle	48.91	1.85	54	14/10/2011	280	31/07/2012	9.1
<i>Gre13</i>	2013	False	Greoux-Les-Bains	43.75	5.86	300	08/11/2012	330	26/07/2013	11.2
<i>Vra13</i>	2013	False	Vraux	48.95	4.42	102	23/10/2012	375	12/08/2013	10
<i>Rea13</i>	2013	False	Realville	44.14	1.50	100	14/11/2012	250	31/07/2013	6.75
<i>Cle13</i>	2013	False	Clermont-Ferrand	45.79	3.15	331	08/11/2012	290	29/07/2013	7.5
<i>Mon13</i>	2013	False	Mons	49.88	3.01	85	24/10/2012	225	31/07/2013	5.2
<i>Ver13</i>	2013	False	Verneuil-l'Étan	48.38	2.50	91	29/10/2012	250	29/07/2013	10
<i>And13</i>	2013	False	Andelu	48.89	1.80	97	27/10/2012	280	31/07/2013	4.9
<i>Gre14</i>	2014	True	Greoux-Les-Bains	43.75	5.86	300	07/11/2013	330	21/07/2014	11.2
<i>Wal14</i>	2014	True	Walla Walla	48.11	-118.29	293	04/11/2013	269	15/07/2016	5.5
<i>Cha14</i>	2014	True	Champigny	46.67	0.20	135	22/10/2013	300	17/07/2014	8.4
<i>All14</i>	2014	False	Allones	48.34	1.65	150	16/10/2013	300	10/08/2014	6
<i>Wal15</i>	2015	True	Walla Walla	46.11	-118.29	293	08/11/2014	304	15/07/2015	5.5
<i>Cle16</i>	2016	True	Clermont-Ferrand	45.78	3.15	332	12/11/2015	250	11/07/2016	3.5
<i>Cas16</i>	2016	True	Castelnaudary	43.30	1.96	160	05/11/2015	275	22/07/2016	6.4
<i>Wal16</i>	2016	True	Walla Walla	46.11	-118.29	293	23/10/2015	304	13/07/2016	5.5
<i>Olm16</i>	2016	False	Ouziouers-Le-Marche	47.88	1.52	125	26/10/2015	250	19/07/2016	6.37
<i>Wal17</i>	2017	True	Walla Walla	46.11	-118.29	293	09/11/2016	300	18/07/2017	5.5

III.2B. Technical report, nitrogen fertilization and irrigation

TRIAL	DATE	INTERVENTION	PRODUCT	QUANTITIES	UNITS
ALL14RF	24/02/2014	Fertilization	Ammonitrate	51	kg/ha
ALL14RF	27/03/2014	Fertilization	Ammonitrate	78	kg/ha
ALL14RF	26/02/2014	Fertilization	Ammonitrate	16	kg/ha
ALL14RF	06/05/2014	Fertilization	Ammonitrate	11.7	kg/ha
ALL14RF	26/02/2014	Fertilization	P ₂ O ₅	42	kg/ha
ALL14RF	06/03/2014	Fertilization	SO ₃	35	kg/ha
AND13RF	17/02/2013	Fertilization	Ammonitrate	80	kg/ha
AND13RF	21/03/2013	Fertilization	Ammonitrate	80	kg/ha
AND13RF	24/04/2013	Fertilization	Ammonitrate	80	kg/ha
CAP12RF	01/02/2012	Fertilization	Ammonitrate	48	kg/ha
CAP12RF	20/03/2012	Fertilization	Ammonitrate	130	kg/ha
CAP12RF	07/05/2012	Fertilization	Ammonitrate	50	kg/ha
CAS16IR	25/01/2016	Fertilization	Ammonitrate	60	kg/ha
CAS16IR	20/02/2016	Fertilization	Ammonitrate	70	kg/ha
CAS16IR	14/04/2016	Fertilization	Ammonitrate	60	kg/ha
CAS16IR	04/05/2016	Irrigation	Water	25	mm
CAS16IR	24/05/2016	Irrigation	Water	40	mm
CAS16IR	06/06/2016	Irrigation	Water	30	mm
CAS16RF	25/01/2016	Fertilization	Ammonitrate	60	kg/ha
CAS16RF	20/02/2016	Fertilization	Ammonitrate	70	kg/ha
CAS16RF	14/04/2016	Fertilization	Ammonitrate	60	kg/ha
CHA12RF	02/03/2012	Fertilization	Ammonitrate	50	kg/ha
CHA12RF	11/04/2012	Fertilization	Ammonitrate	110	kg/ha
CHA12RF	21/05/2012	Fertilization	Ammonitrate	40	kg/ha
CHA14IR	06/02/2014	Fertilization	Ammonitrate	30	kg/ha
CHA14IR	25/03/2014	Fertilization	Ammonitrate	62	kg/ha
CHA14IR	17/04/2014	Fertilization	Ammonitrate	90	kg/ha
CHA14IR	21/02/2014	Fertilization	P ₂ O ₅	61	kg/ha
CHA14IR	15/04/2014	Irrigation	Water	35	mm
CHA14RF	06/02/2014	Fertilization	Ammonitrate	30	kg/ha
CHA14RF	25/03/2014	Fertilization	Ammonitrate	62	kg/ha
CHA14RF	17/04/2014	Fertilization	Ammonitrate	90	kg/ha
CHA14RF	21/02/2014	Fertilization	P ₂ O ₅	61	kg/ha
CLE13RF	21/03/2013	Fertilization	Ammonitrate	50	kg/ha
CLE13RF	02/05/2013	Fertilization	Ammonitrate	75	kg/ha
CLE13RF	13/05/2013	Fertilization	Ammonitrate	75	kg/ha
CLE13RF	03/06/2013	Fertilization	Ammonitrate	50	kg/ha
CLE16RF	10/03/2016	Fertilization	Ammonitrate	50	kg/ha
CLE16RF	27/04/2016	Fertilization	Ammonitrate	55	kg/ha
CLE16RO	10/03/2016	Fertilization	Ammonitrate	50	kg/ha
CLE16RO	27/04/2016	Fertilization	Ammonitrate	55	kg/ha
CLE16RF	02/03/2016-24/05/2016	Shelter closures	Water		
CLE16RF	10/06/2016-24/06/2016	Shelter closures	Water		
COI12IR	16/02/2012	Fertilization	Ammonitrate	70	kg/ha
COI12IR	11/03/2012	Fertilization	Ammonitrate	100	kg/ha
COI12IR	11/04/2012	Fertilization	Ammonitrate	40	kg/ha
COI12IR	27/02/2012	Fertilization	P ₂ O ₅	47	kg/ha
COI12IR	27/02/2012	Fertilization	SO ₃	43	kg/ha
COI12IR	12/04/2012	Irrigation	Water	15	mm
COI12RF	16/02/2012	Fertilization	Ammonitrate	70	kg/ha
COI12RF	11/03/2012	Fertilization	Ammonitrate	100	kg/ha
COI12RF	11/04/2012	Fertilization	Ammonitrate	40	kg/ha
COI12RF	27/02/2012	Fertilization	P ₂ O ₅	47	kg/ha
COI12RF	27/02/2012	Fertilization	SO ₃	43	kg/ha
GRE12IR	21/03/2012	Fertilization	Ammonitrate	60	kg/ha
GRE12IR	12/04/2012	Fertilization	Ammonitrate	50	kg/ha
GRE12IR	27/04/2012	Fertilization	Ammonitrate	100	kg/ha
GRE12IR	23/01/2012	Fertilization	P ₂ O ₅	48	kg/ha
GRE12IR	23/01/2012	Fertilization	K ₂ O	24	kg/ha
GRE12IR	23/01/2012	Fertilization	SO ₃	43	kg/ha
GRE12IR	28/03/2012	Irrigation	Water	30	mm
GRE12IR	17/05/2012	Irrigation	Water	35	mm
GRE12IR	08/06/2012	Irrigation	Water	40	mm
GRE12IR	15/06/2012	Irrigation	Water	30	mm
GRE12RF	21/03/2012	Fertilization	Ammonitrate	60	kg/ha
GRE12RF	12/04/2012	Fertilization	Ammonitrate	50	kg/ha

GRE12RF	27/04/2012	Fertilization	Ammonitrate	100	kg/ha
GRE12RF	23/01/2012	Fertilization	P ₂ O ₅	48	kg/ha
GRE12RF	23/01/2012	Fertilization	K ₂ O	24	kg/ha
GRE12RF	23/01/2012	Fertilization	SO ₃	43	kg/ha
GRE12RF	28/03/2012	Irrigation	Water	30	mm
GRE13RF	23/11/2012	Fertilization	Ammonitrate	35	kg/ha
GRE13RF	22/02/2013	Fertilization	Ammonitrate	60	kg/ha
GRE13RF	04/04/2013	Fertilization	Ammonitrate	80	kg/ha
GRE13RF	07/05/2013	Fertilization	Ammonitrate	50	kg/ha
GRE13RF	23/11/2012	Fertilization	P ₂ O ₅	40	kg/ha
GRE13RF	23/11/2012	Fertilization	K ₂ O	25	kg/ha
GRE13RF	23/11/2012	Fertilization	SO ₃	50	kg/ha
GRE14IR	18/02/2014	Fertilization	Ammonitrate	50	kg/ha
GRE14IR	21/03/2014	Fertilization	Ammonitrate	80	kg/ha
GRE14IR	18/04/2014	Fertilization	Ammonitrate	50	kg/ha
GRE14IR	06/05/2014	Fertilization	Ammonitrate	80	kg/ha
GRE14IR	29/11/2013	Fertilization	P ₂ O ₅	20	kg/ha
GRE14IR	29/11/2013	Fertilization	K ₂ O	10	kg/ha
GRE14IR	08/11/2013	Irrigation	Water	15	mm
GRE14IR	14/04/2014	Irrigation	Water	55	mm
GRE14IR	08/05/2014	Irrigation	Water	30	mm
GRE14IR	16/05/2014	Irrigation	Water	35	mm
GRE14IR	27/05/2014	Irrigation	Water	33	mm
GRE14IR	05/06/2014	Irrigation	Water	40	mm
GRE14IR	12/06/2014	Irrigation	Water	25	mm
GRE14RF	18/02/2014	Fertilization	Ammonitrate	50	kg/ha
GRE14RF	21/03/2014	Fertilization	Ammonitrate	80	kg/ha
GRE14RF	18/04/2014	Fertilization	Ammonitrate	50	kg/ha
GRE14RF	06/05/2014	Fertilization	Ammonitrate	80	kg/ha
GRE14 RF	29/11/2013	Fertilization	P ₂ O ₅	20	kg/ha
GRE14RF	29/11/2013	Fertilization	K ₂ O	10	kg/ha
GRE14RF	08/11/2013	Irrigation	Water	15	mm
LOU12RF	02/03/2012	Fertilization	Ammonitrate	50	kg/ha
LOU12RF	05/04/2012	Fertilization	Ammonitrate	60	kg/ha
LOU12RF	23/04/2012	Fertilization	Ammonitrate	40	kg/ha
MAU12RF	03/03/2012	Fertilization	Ammonitrate	50	kg/ha
MAU12RF	20/03/2012	Fertilization	Ammonitrate	80	kg/ha
MAU12RF	15/05/2012	Fertilization	Ammonitrate	50	kg/ha
MON12RF	12/03/2012	Fertilization	Ammonitrate	50	kg/ha
MON12RF	29/03/2012	Fertilization	Ammonitrate	50	kg/ha
MON12RF	11/04/2012	Fertilization	Ammonitrate	50	kg/ha
MON12RF	21/05/2012	Fertilization	Ammonitrate	50	kg/ha
MON13RF	26/03/2013	Fertilization	Ammonitrate	50	kg/ha
MON13RF	26/04/2013	Fertilization	Ammonitrate	50	kg/ha
MON13RF	07/05/2013	Fertilization	Ammonitrate	50	kg/ha
MON13RF	13/06/2013	Fertilization	Ammonitrate	50	kg/ha
OLM16RO	18/03/2016	Fertilization	Ammonitrate	85	kg/ha
OLM16RO	14/04/2016	Fertilization	Ammonitrate	75	kg/ha
OLM16RO	21/03/2016	Irrigation	Water	20	mm
OLM16RO	18/04/2016	Irrigation	Water	15	mm
OLM16RO	19/01/2016-13/04/2016	Shelter closures	Water		
OLM16RO	21/04/2016-31/07/2016	Shelter closures	Water		
REC12RF	02/03/2012	Fertilization	Ammonitrate	50	kg/ha
REC12RF	29/03/2012	Fertilization	Ammonitrate	50	kg/ha
REC12RF	09/05/2012	Fertilization	Ammonitrate	50	kg/ha
REC12RF	02/03/2012	Fertilization	SO ₃	37	kg/ha
VER13RF	21/02/2013	Fertilization	Ammonitrate	50	kg/ha
VER13RF	28/03/2013	Fertilization	Ammonitrate	100	kg/ha
VER13RF	24/05/2013	Fertilization	Ammonitrate	50	kg/ha
VRA13RF	01/03/2013	Fertilization	Ammonitrate	50	kg/ha
VRA13RF	27/03/2013	Fertilization	Ammonitrate	110	kg/ha
VRA13RF	27/05/2013	Fertilization	Ammonitrate	50	kg/ha
WAL14IR	23/03/2014	Fertilization	Ammonitrate	60	kg/ha
WAL14IR	21/05/2014	Irrigation	Water	19.75	mm
WAL14IR	29/05/2014	Irrigation	Water	22	mm
WAL14IR	04/06/2014	Irrigation	Water	22.5	mm
WAL14IR	10/06/2014	Irrigation	Water	23.5	mm
WAL14IR	23/06/2014	Irrigation	Water	20.75	mm
WAL14RF	23/03/2014	Fertilization	Ammonitrate	60	kg/ha
WAL15IR	14/05/2015	Fertilization	Ammonitrate	45	kg/ha

WAL15IR	09/03/2015	Irrigation	Water	20	mm
WAL15IR	30/04/2015	Irrigation	Water	17.15	mm
WAL15IR	29/05/2015	Irrigation	Water	18	mm
WAL15IR	04/06/2015	Irrigation	Water	18	mm
WAL15IR	14/06/2015	Irrigation	Water	18	mm
WAL15IR	20/06/2015	Irrigation	Water	16	mm
WAL15RF	14/05/2015	Fertilization	Ammonitrate	45	kg/ha
WAL15RF	09/03/2015	Irrigation	Water	20	mm
WAL15RF	30/04/2015	Irrigation	Water	17.15	mm
WAL16IR	23/04/2016	Fertilization	Ammonitrate	45	kg/ha
WAL16IR	20/05/2016	Irrigation	Water	28	mm
WAL16IR	30/05/2016	Irrigation	Water	22	mm
WAL16IR	03/06/2016	Irrigation	Water	28.5	mm
WAL16IR	13/06/2016	Irrigation	Water	34	mm
WAL16RF	23/04/2016	Fertilization	Ammonitrate	45	kg/ha
WAL17IR	01/06/2017	Fertilization	Ammonitrate	44	kg/ha
WAL17IR	23/05/2017	Irrigation	Water	33	mm
WAL17IR	04/06/2017	Irrigation	Water	22	mm
WAL17IR	22/06/2017	Irrigation	Water	22	mm
WAL17IR	03/07/2017	Irrigation	Water	14	mm
WAL17RF	01/06/2017	Fertilization	Ammonitrate	44	kg/ha

III.2C. Weather characteristics of the MET

Trial name (Trial) with the first three letters referring to location, the two numbers the year and the last two letters the management. Note also that Cle16RO and Olm16RO are both rain out sheltered experiment. Days from first leaves emergence to flowering for the reference “Apache” (DTA), days from sowing to maturity of “Apache” (DTM), plant-available soil water capacity (PAWC, mm), sum of irrigation (IRR, mm), cumulative in-crop rainfall (ICR, mm), water potential available (WPA), average daily temperature from sowing to maturity (AvgTemp), cumulative radiation from sowing to maturity (CumRad, MJ.m⁻²), cumulative reference evapotranspiration from sowing to maturity (CumET0), environment cluster (EC)

<i>Trial</i>	<i>DTA</i>	<i>DTM</i>	<i>PAWC</i>	<i>IRR</i>	<i>ICR</i>	<i>WPA</i> ¹	<i>AvgTemp</i>	<i>CumRad</i>	<i>CumET0</i>	<i>ETs</i>
<i>All14RF</i>	214	252	245.61	0	489.9	735.51	9.66	248.6	449.3	OPT
<i>And13RF</i>	209	243	268.01	0	423.36	691.37	8.65	253.59	447.56	OPT
<i>Cap12RF</i>	206	246	298.46	0	522.22	820.68	8.77	234.12	410.21	OPT
<i>Cha12RF</i>	210	247	268.01	0	448.3	716.31	8.95	258.69	397.07	OPT
<i>Cha14IR</i>	193	233	104.98	35	587.89	727.87	10.17	259.88	448.55	OPT
<i>Cha14RF</i>	195	233	104.98	0	587.89	692.87	10.17	259.88	448.55	OPT
<i>Cle16RF</i>	173	210	329.99	0	156.5	486.49	10.27	227.98	478.96	OPT
<i>Coi12IR</i>	208	245	121.24	15	499.49	635.73	9.36	267.64	432.72	OPT
<i>Coi12RF</i>	207	244	121.24	0	499.49	620.73	9.33	265.79	428.43	OPT
<i>Gre12IR</i>	172	205	114.49	135	244.4	493.89	10.05	324.62	461.49	OPT
<i>Lou12RF</i>	210	248	245.61	0	494.5	740.11	8.43	229.58	358.4	OPT
<i>Mau12RF</i>	211	249	236.25	0	505.9	742.15	9.63	241.58	420.62	OPT
<i>Mon12RF</i>	212	252	236.55	0	544.5	781.05	8.59	238.57	378	OPT
<i>Mon13RF</i>	212	248	236.55	0	402.7	639.25	7.79	252	410.68	OPT
<i>Rea13RF</i>	175	213	120.1	0	694.75	814.86	9.99	247.45	431.96	OPT
<i>Rec12RF</i>	202	239	245.61	0	405.1	650.71	8.93	252.12	404.8	OPT
<i>Ver13RF</i>	205	239	236.25	0	437.27	673.52	8.41	252.17	464.36	OPT
<i>Vra13RF</i>	210	245	268.01	0	403.7	671.71	7.95	250.98	406.88	OPT
<i>Wal16IR</i>	190	235	237.8	112.5	462.4	812.7	9.03	297.92	595.07	OPT
<i>Wal17IR</i>	186	219	148.32	91	388.24	627.56	7	250.81	508.85	OPT
<i>Wal17RF</i>	185	218	148.32	0	390.78	539.1	6.93	248.13	503.04	OPT
<i>Cle13RF</i>	189	225	103.09	0	285.6	388.69	8.59	246.35	493.36	LWD
<i>Gre12RF</i>	173	206	114.49	30 ²	244.8	389.29	10.13	327.6	468.73	LWD
<i>Gre13RF</i>	186	222	114.49	0	485.8	600.29	9.28	318.5	454.83	LWD
<i>Wal14IR</i>	169	207	237.8	108.5	259.6	605.9	9.53	285.59	622.84	LWD
<i>Wal16RF</i>	190	235	237.8	0	462.4	700.2	9.03	297.92	595.07	LWD
<i>Cas16IR</i>	179	217	101.71	95	353.2	549.91	11.17	229.57	432.15	MWD
<i>Cle16RO</i>	167	205	103.09	0	155.9	258.99	10.04	215.98	455.15	MWD
<i>Gre14IR</i>	162	198	114.49	233	494.6	842.09	10.63	283.58	429.36	MWD
<i>Wal15IR</i>	159	198	237.8	107.15	251.8	596.75	11.12	297.64	624.21	MWD
<i>Cas16RF</i>	182	218	101.71	0	354.2	455.91	11.23	232.37	436.64	HWD
<i>Gre14RF</i>	162	198	114.49	15 ²	494.8	624.29	10.63	283.58	429.36	HWD
<i>Olm16RO</i>	195	236	174.72	35 ²	105.2	314.92	9.3	232.64	323.91	HWD
<i>Wal14RF</i>	169	207	237.8	0	260	497.8	9.53	285.59	622.84	HWD
<i>Wal15RF</i>	159	198	237.8	37.15 ²	251.8	526.75	11.12	297.64	624.21	HWD

¹Water potentially available (WPA) = PAW + IRR + ICR

²Field irrigation to valorize the nitrogen input

III.2D. Characterization of environments according to cumulative water deficit from April to June for each environment.

Cumulative water deficit was defined as the sum of rain and irrigation less the reference evapotranspiration.

<i>Envi.</i>	<i>April</i>	<i>May</i>	<i>June</i>	<i>Sum</i>	<i>ETs</i>
<i>All14RF</i>	-30.60	-2.40	-48.50	-81.50	OPT
<i>And13RF</i>	-45.37	9.45	-44.97	-80.89	OPT
<i>Cap12RF</i>	14.24	-71.40	-10.02	-67.18	OPT
<i>Cha12RF</i>	18.00	-73.50	-20.78	-76.28	OPT
<i>Cha14IR</i>	0.29	21.49	-84.44	-62.67	OPT
<i>Cha14RF</i>	-34.71	21.49	-84.44	-97.67	OPT
<i>Cle16RF</i>	27.68	-8.20	-46.29	-26.81	OPT
<i>Coi12IR</i>	27.14	-36.41	-33.85	-43.12	OPT
<i>Coi12RF</i>	12.14	-36.41	-33.85	-58.12	OPT
<i>Gre12IR</i>	26.83	-7.79	-52.66	-33.62	OPT
<i>Lou12RF</i>	17.50	-63.40	38.20	-7.70	OPT
<i>Mau12RF</i>	10.29	-30.61	11.50	-8.82	OPT
<i>Mon12RF</i>	17.50	-63.40	38.20	-7.70	OPT
<i>Mon13RF</i>	-55.73	16.85	-36.16	-75.04	OPT
<i>Real3RF</i>	-11.46	36.87	-40.86	-15.45	OPT
<i>Rec12RF</i>	27.40	-62.30	-17.50	-52.40	OPT
<i>Ver13RF</i>	-47.54	6.66	-15.54	-56.42	OPT
<i>Vra13RF</i>	-14.87	16.66	-47.16	-45.37	OPT
<i>Wal16IR</i>	-76.76	-34.73	-60.54	-172.03	OPT
<i>Wal17IR</i>	-28.60	-55.20	-55.25	-139.04	OPT
<i>Wal17RF</i>	-28.60	-88.20	-96.71	-213.50	OPT
<i>Cle13RF</i>	-13.00	-13.00	-106.40	-132.40	LWD
<i>Gre12RF</i>	26.83	-42.39	-129.90	-145.46	LWD
<i>Gre13RF</i>	14.64	-21.31	-149.35	-156.02	LWD
<i>Wal14IR</i>	-43.94	-84.75	-73.95	-202.64	LWD
<i>Wal16RF</i>	-76.76	-84.73	-123.04	-284.53	LWD
<i>Cas16IR</i>	-29.69	24.29	-9.47	-14.87	MWD
<i>Cle16RO</i>	-72.52	-59.30	-78.68	-210.50	MWD
<i>Gre14IR</i>	0.00	-14.78	28.58	13.80	MWD
<i>Wal15IR</i>	-65.30	-97.12	-121.00	-283.41	MWD
<i>Cas16RF</i>	-29.69	-40.51	-43.16	-113.36	HWD
<i>Gre14RF</i>	-55.00	-112.78	-36.22	-204.00	HWD
<i>Olm16RO</i>	-41.99	-86.83	-95.54	-224.36	HWD
<i>Wal14RF</i>	-43.94	-126.50	-140.30	-310.74	HWD
<i>Wal15RF</i>	-65.30	-115.12	-173.00	-353.41	HWD

- **Supplementary data III.3: Drought pattern simulation over France**

III.3A.Site simulation

A total of 48,762 simulations, 21 seasons x 258 sites x 3 soil depths x 3 sowing dates, were carried out to estimate the frequency of water stress at French scale. The 258 sites are presented in row in the following table. Depending on the region, one optimal sowing Cultivar/Date combination was selected for each site based on agronomists expertise. The optimal date was extended by +/-7 days to cover the sowing dispersion. Furthermore, three different soils were used to represent the different available water capacity specific to each region (soil 1 = low, soil 2 = medium and soil 3 =high available water capacity).

Sites	Cultivar	Date	LAT.	LONG.	ALT.	Soil_1	Soil_2	Soil_3
6701	Apache	21/10	45.977	5.330	250	RA0066501	RA0016501	RA5016501
6801	Apache	21/10	45.770	5.687	330	RA0066501	RA0016501	RA5016501
6851	Apache	21/10	46.043	5.073	280	RA0066501	RA0016501	RA5016501
6854	Apache	21/10	45.805	5.237	198	RA0066501	RA0016501	RA5016501
6875	Apache	21/10	46.047	5.073	273	RA0066501	RA0016501	RA5016501
2451	Soissons	26/10	45.974	4.819	280	RA0066501	RA0016501	RA5016501
2457	Soissons	26/10	49.817	3.203	98	PI5533002	PI5533001	PI0066001
2459	Soissons	26/10	49.590	3.607	78	PI5533002	PI5533001	PI0066001
2462	Soissons	26/10	46.166	3.400	249	AU0033500	AU0116001	AU5016501
2471	Soissons	26/10	46.544	3.706	220	AU0033500	AU0116001	AU5016501
2490	Soissons	26/10	46.718	2.947	225	AU0033500	AU0116001	AU5016501
3301	Apache	26/10	44.061	5.990	461	PR5133000	PR5516000	PR5066000
3302	Apache	26/10	43.752	5.858	296	PR5133000	PR5516000	PR5066000
3365	Apache	26/10	43.655	7.202	4	PR5133000	PR5516000	PR5066000
3382	Apache	26/10	45.246	4.668	355	RA0066501	RA0016501	RA5016501
4001	Soissons	26/10	49.515	4.765	174	CH0086001	CH0133001	CH0036002
4053	Soissons	26/10	43.003	1.107	411	MP0533001	MP00165_0008	MP0036000
4065	Soissons	26/10	43.118	1.613	285	MP0533001	MP00165_0008	MP0036000
4084	Soissons	26/10	48.505	3.750	77	CH0086001	CH0133001	CH0036002
4092	Soissons	26/10	48.328	4.023	112	CH0086001	CH0133001	CH0036002
4701	Soissons	26/10	48.178	3.795	226	CH0086001	CH0133001	CH0036002
4758	Soissons	26/10	48.047	4.658	310	CH0086001	CH0133001	CH0036002
4759	Soissons	26/10	48.347	3.638	120	CH0086001	CH0133001	CH0036002
4775	Soissons	26/10	48.235	4.046	122	CH0086001	CH0133001	CH0036002
4780	Soissons	26/10	48.365	4.473	140	CH0086001	CH0133001	CH0036002
6401	Soissons	1/11	48.248	4.336	140	CH0086001	CH0133001	CH0036002
6402	Soissons	1/11	43.200	2.316	126	LR0116001	LR5561001	LR5116001
6474	Soissons	1/11	43.306	1.951	170	LR0116001	LR5561001	LR5116001
6485	Soissons	1/11	43.958	2.885	365	MP0533001	MP00165_0008	MP0036000
6486	Soissons	1/11	44.386	2.335	406	MP0533001	MP00165_0008	MP0036000
301	Apache	21/10	44.368	2.022	345	MP0533001	MP00165_0008	MP0036000
361	Apache	21/10	43.442	5.227	5	PR5133000	PR5516000	PR5066000
386	Apache	21/10	43.526	5.427	173	PR5133000	PR5516000	PR5066000
4351	Rubisko	16/10	43.513	4.670	1	PR5133000	PR5516000	PR5066000
6301	Apache	26/10	49.183	-0.450	64	BN5531001	BN5533001	BN0066501
1401	Rubisko	21/10	48.988	-0.008	60	BN5531001	BN5533001	BN0066501
1452	Rubisko	21/10	48.850	-0.898	101	BN5531001	BN5533001	BN0066501
1471	Rubisko	21/10	45.665	-0.316	30	PC01330_0020	PC0036002	PC0036001
6101	Rubisko	21/10	45.479	-0.146	79	PC01330_0020	PC0036002	PC0036001
6151	Rubisko	21/10	45.262	0.003	81	PC01330_0020	PC0036002	PC0036001
2101	Altigo	6/10	45.632	0.106	58	PC01330_0020	PC0036002	PC0036001
5801	Soissons	11/10	46.022	0.178	118	PC01330_0020	PC0036002	PC0036001
5802	Soissons	11/10	46.153	-1.157	4	PC01330_0020	PC0036002	PC0036001
7101	Apache	16/10	45.858	-0.319	36	PC01330_0020	PC0036002	PC0036001
7153	Apache	16/10	46.018	-0.406	34	PC01330_0020	PC0036002	PC0036001
8901	Altigo	11/10	45.988	-1.027	3	PC01330_0020	PC0036002	PC0036001
2202	Bermude	1/11	45.373	-0.575	64	PC01330_0020	PC0036002	PC0036001
2290	Bermude	1/11	46.140	-0.690	50	PC01330_0020	PC0036002	PC0036001

2901	Bermude	1/11	47.066	2.366	161	CE0093501	CE5016001	CE00365_0055
2951	Bermude	1/11	47.050	2.650	176	CE0093501	CE5016001	CE00365_0055
2955	Bermude	1/11	47.498	2.414	212	CE0093501	CE5016001	CE00365_0055
2957	Bermude	1/11	46.425	2.235	462	CE0093501	CE5016001	CE00365_0055
2965	Bermude	1/11	45.242	1.740	230	LI0186501	LI0086501	LI0066501
2982	Bermude	1/11	47.275	5.092	219	BO55110_0042	BO0066501	BO0036002
3501	Altigo	26/10	48.232	-3.307	262	BR0086000	BR0036500	BR5016500
3502	Altigo	26/10	48.536	-2.852	136	BR0086000	BR0036500	BR5016500
3560	Altigo	26/10	46.238	1.468	405	LI0186501	LI0086501	LI0066501
5601	Rubisko	1/11	45.943	1.728	510	LI0186501	LI0086501	LI0066501
5603	Rubisko	1/11	45.158	0.673	215	AQ5533001	AQ00860_0032	AQ0036001
5676	Rubisko	1/11	45.025	0.718	125	AQ5533001	AQ00860_0032	AQ0036001
5683	Rubisko	1/11	44.935	1.013	60	AQ5533001	AQ00860_0032	AQ0036001
5695	Rubisko	1/11	44.851	0.023	10	AQ5533001	AQ00860_0032	AQ0036001
1801	Soissons	16/10	45.201	0.133	65	AQ5533001	AQ00860_0032	AQ0036001
1856	Soissons	16/10	44.803	0.513	33	AQ5533001	AQ00860_0032	AQ0036001
1858	Soissons	16/10	47.250	5.988	307	FC5531001	FC0016501	FC5016001
1888	Soissons	16/10	44.583	4.738	73	RA0066501	RA0016501	RA5016501
2801	Rubisko	16/10	44.717	4.990	175	RA0066501	RA0016501	RA5016501
2863	Rubisko	16/10	44.898	5.016	160	RA0066501	RA0016501	RA5016501
2873	Rubisko	16/10	44.775	5.338	385	RA0066501	RA0016501	RA5016501
3601	Soissons	16/10	49.025	1.222	138	HN0096001	HN0066501	HN0066001
3651	Soissons	16/10	48.836	0.915	177	HN0096001	HN0066501	HN0066001
3653	Soissons	16/10	49.194	0.713	59	HN0096001	HN0066501	HN0066001
3663	Soissons	16/10	49.306	1.612	90	HN0096001	HN0066501	HN0066001
3680	Soissons	16/10	48.898	1.406	125	HN0096001	HN0066501	HN0066001
3693	Soissons	16/10	48.465	1.516	155	CE0093501	CE5016001	CE00365_0055
3701	Soissons	16/10	48.062	1.377	126	CE0093501	CE5016001	CE00365_0055
3758	Soissons	16/10	48.653	1.370	151	CE0093501	CE5016001	CE00365_0055
4101	Rubisko	16/10	48.445	-4.416	94	BR0086000	BR0036500	BR5016500
4165	Rubisko	16/10	48.240	-4.062	70	BR0086000	BR0036500	BR5016500
4501	Rubisko	21/10	48.067	-3.817	249	BR0086000	BR0036500	BR5016500
4551	Rubisko	21/10	47.973	-4.170	92	BR0086000	BR0036500	BR5016500
4584	Rubisko	21/10	48.276	-3.523	153	BR0086000	BR0036500	BR5016500
858	Toisonдор	11/10	47.951	-4.123	33	BR0086000	BR0036500	BR5016500
1001	Rubisko	16/10	43.865	4.400	59	LR0116001	LR5561001	LR5116001
1002	Rubisko	16/10	43.633	1.378	151	MP0533001	MP00165_0008	MP0036000
1057	Rubisko	16/10	43.458	2.003	334	MP0533001	MP00165_0008	MP0036000
1062	Rubisko	16/10	43.423	1.655	148	MP0533001	MP00165_0008	MP0036000
1066	Rubisko	16/10	43.690	0.600	121	MP0533001	MP00165_0008	MP0036000
1071	Rubisko	16/10	43.958	0.370	160	MP0533001	MP00165_0008	MP0036000
1083	Rubisko	16/10	43.604	1.085	149	MP0533001	MP00165_0008	MP0036000
1084	Rubisko	16/10	43.610	-0.100	160	MP0533001	MP00165_0008	MP0036000
5158	Toisonдор	11/10	43.516	0.401	187	MP0533001	MP00165_0008	MP0036000
5175	Toisonдор	11/10	43.838	0.383	121	MP0533001	MP00165_0008	MP0036000
5176	Toisonдор	11/10	43.627	0.875	175	MP0533001	MP00165_0008	MP0036000
5201	Rubisko	6/10	44.833	-0.692	47	AQ5533001	AQ00860_0032	AQ0036001
5202	Rubisko	6/10	44.535	-1.133	23	AQ5533001	AQ00860_0032	AQ0036001
2501	Altigo	11/10	45.040	-0.126	21	AQ5533001	AQ00860_0032	AQ0036001
3951	Altigo	11/10	44.547	-0.323	86	AQ5533001	AQ00860_0032	AQ0036001
3952	Altigo	11/10	43.583	3.966	3	LR0116001	LR5561001	LR5116001
7001	Soissons	6/10	43.716	3.867	80	LR0116001	LR5561001	LR5116001
7053	Soissons	6/10	48.070	-1.727	36	BR0086000	BR0036500	BR5016500
2701	Bermude	11/10	48.590	-2.073	58	BR0086000	BR0036500	BR5016500
2751	Bermude	11/10	48.327	-1.597	87	BR0086000	BR0036500	BR5016500
2752	Bermude	11/10	46.860	1.722	155	CE0093501	CE5016001	CE00365_0055
2754	Bermude	11/10	46.485	2.071	302	CE0093501	CE5016001	CE00365_0055
2771	Bermude	11/10	46.634	1.063	117	CE0093501	CE5016001	CE00365_0055
7601	Rubisko	16/10	46.983	1.416	164	CE0093501	CE5016001	CE00365_0055
7655	Rubisko	16/10	46.442	1.580	272	CE0093501	CE5016001	CE00365_0055
7656	Rubisko	16/10	46.660	1.790	185	CE0093501	CE5016001	CE00365_0055
7699	Rubisko	16/10	47.443	0.746	108	CE0093501	CE5016001	CE00365_0055
7701	Rubisko	16/10	47.031	0.395	115	CE0093501	CE5016001	CE00365_0055

7781	Rubisko	16/10	45.363	5.313	384	RA0066501	RA0016501	RA5016501
7801	Rubisko	16/10	45.143	5.290	280	RA0066501	RA0016501	RA5016501
7867	Rubisko	16/10	45.339	5.053	260	RA0066501	RA0016501	RA5016501
7868	Rubisko	16/10	47.091	5.495	222	FC5531001	FC0016501	FC5016001
9151	Rubisko	16/10	46.978	5.313	185	FC5531001	FC0016501	FC5016001
9199	Rubisko	16/10	43.916	-0.500	59	AQ5533001	AQ00860_0032	AQ0036001
9553	Rubisko	16/10	43.593	-0.743	48	AQ5533001	AQ00860_0032	AQ0036001
9558	Rubisko	16/10	43.692	-1.067	31	AQ5533001	AQ00860_0032	AQ0036001
1101	Soissons	26/10	44.308	-0.778	56	AQ5533001	AQ00860_0032	AQ0036001
1151	Soissons	26/10	43.572	-1.297	37	AQ5533001	AQ00860_0032	AQ0036001
3002	Soissons	26/10	47.316	1.683	84	CE0093501	CE5016001	CE00365_0055
3401	Soissons	26/10	47.793	1.067	83	CE0093501	CE5016001	CE00365_0055
3458	Soissons	26/10	45.533	4.292	400	RA0066501	RA0016501	RA5016501
6601	Soissons	1/11	45.607	4.083	374	RA0066501	RA0016501	RA5016501
1954	Apache	21/10	45.302	3.422	435	AU0033500	AU0116001	AU5016501
2351	Soissons	21/10	47.150	-1.608	26	PL0186500	PL0563501	PL0036000
2362	Soissons	21/10	47.415	-1.513	12	PL0186500	PL0563501	PL0036000
8701	Soissons	21/10	47.693	-1.687	43	PL0186500	PL0563501	PL0036000
8767	Soissons	21/10	47.453	-0.969	71	PL0186500	PL0563501	PL0036000
5401	Bermude	1/10	47.983	1.777	125	CE0093501	CE5016001	CE00365_0055
5465	Bermude	1/10	48.159	2.210	120	CE0093501	CE5016001	CE00365_0055
5471	Bermude	1/10	47.928	1.917	125	CE0093501	CE5016001	CE00365_0055
5472	Bermude	1/10	44.747	1.397	259	MP0533001	MP00165_0008	MP0036000
5559	Bermude	1/10	44.336	1.206	192	MP0533001	MP00165_0008	MP0036000
5701	Bermude	1/10	44.172	0.593	59	AQ5533001	AQ00860_0032	AQ0036001
5702	Bermude	1/10	44.205	0.408	40	AQ5533001	AQ00860_0032	AQ0036001
8851	Altigo	1/10	44.534	0.623	145	AQ5533001	AQ00860_0032	AQ0036001
901	Apache	26/10	44.088	0.186	169	AQ5533001	AQ00860_0032	AQ0036001
954	Apache	26/10	44.419	0.285	28	AQ5533001	AQ00860_0032	AQ0036001
1251	Apache	26/10	47.483	-0.600	50	PL0186500	PL0563501	PL0036000
1280	Apache	26/10	47.256	-0.076	69	PL0186500	PL0563501	PL0036000
1284	Apache	26/10	47.133	-0.150	60	PL0186500	PL0563501	PL0036000
3101	Soissons	26/10	49.518	-1.505	61	BN5531001	BN5533001	BN0066501
3170	Soissons	26/10	49.051	-1.034	25	BN5531001	BN5533001	BN0066501
3190	Soissons	26/10	49.049	-1.443	60	BN5531001	BN5533001	BN0066501
3252	Soissons	26/10	48.566	-1.100	110	BN5531001	BN5533001	BN0066501
3254	Soissons	26/10	48.731	3.559	185	CH0086001	CH0133001	CH0036002
3259	Soissons	26/10	48.700	4.588	105	CH0086001	CH0133001	CH0036002
3265	Soissons	26/10	48.950	4.416	102	CH0086001	CH0133001	CH0036002
3268	Soissons	26/10	47.845	5.340	467	CH0086001	CH0133001	CH0036002
3276	Soissons	26/10	48.631	4.901	139	CH0086001	CH0133001	CH0036002
3281	Soissons	26/10	48.300	-0.613	135	PL0186500	PL0563501	PL0036000
4601	Soissons	26/10	48.083	-0.640	90	PL0186500	PL0563501	PL0036000
4668	Soissons	26/10	48.580	5.958	336	LO5531004	LO5016500	LO0016501
6501	Soissons	1/11	48.686	6.220	212	LO5531004	LO5016500	LO0016501
6557	Soissons	1/11	48.495	6.903	332	LO5531004	LO5016500	LO0016501
6560	Soissons	1/11	49.482	5.577	345	LO5531004	LO5016500	LO0016501
6564	Soissons	1/11	48.750	5.403	313	LO5531004	LO5016500	LO0016501
8100	Soissons	26/10	47.768	-3.442	42	BR0086000	BR0036500	BR5016500
8180	Soissons	26/10	47.950	-2.400	67	BR0086000	BR0036500	BR5016500
8199	Soissons	26/10	48.067	-2.285	86	BR0086000	BR0036500	BR5016500
8253	Soissons	26/10	48.055	-2.917	80	BR0086000	BR0036500	BR5016500
8254	Soissons	26/10	47.885	-2.728	112	BR0086000	BR0036500	BR5016500
5001	Bermude	1/11	49.072	6.125	190	LO5531004	LO5016500	LO0016501
5056	Bermude	1/11	48.767	7.305	377	LO5531004	LO5016500	LO0016501
5058	Bermude	1/11	47.078	3.933	598	BO55110_0042	BO0066501	BO0036002
5071	Bermude	1/11	47.000	3.100	175	BO55110_0042	BO0066501	BO0036002
5901	Rubisko	21/10	50.576	3.099	47	NO0133001	NO0066001	NO0036001
5980	Rubisko	21/10	51.053	2.338	11	NO0133001	NO0066001	NO0036001
6202	Bermude	16/10	49.458	2.116	89	PI5533002	PI5533001	PI0066001
6254	Bermude	16/10	49.154	2.954	100	PI5533002	PI5533001	PI0066001
6259	Bermude	16/10	49.262	2.500	88	PI5533002	PI5533001	PI0066001
4453	Soissons	21/10	49.588	2.551	92	PI5533002	PI5533001	PI0066001

4467	Soissons	21/10	49.651	2.372	94	PI5533002	PI5533001	PI0066001
4479	Soissons	21/10	49.485	3.098	160	PI5533002	PI5533001	PI0066001
4499	Soissons	21/10	48.433	0.100	144	BN5531001	BN5533001	BN0066501
4901	Soissons	21/10	48.707	-0.002	171	BN5531001	BN5533001	BN0066501
4954	Soissons	21/10	50.516	1.616	5	NO0133001	NO0066001	NO0036001
4971	Soissons	21/10	50.221	3.157	77	NO0133001	NO0066001	NO0036001
5355	Altigo	21/10	50.783	1.931	75	NO0133001	NO0066001	NO0036001
5358	Altigo	21/10	45.787	3.148	331	AU0033500	AU0116001	AU5016501
7201	Altigo	21/10	43.471	-1.533	69	AQ5533001	AQ00860_0032	AQ0036001
7270	Altigo	21/10	43.383	-0.416	183	AQ5533001	AQ00860_0032	AQ0036001
8502	Soissons	21/10	43.451	-0.104	310	AQ5533001	AQ00860_0032	AQ0036001
8504	Soissons	21/10	43.321	-0.576	241	AQ5533001	AQ00860_0032	AQ0036001
8505	Soissons	21/10	43.493	-0.538	125	AQ5533001	AQ00860_0032	AQ0036001
8507	Soissons	21/10	43.186	-0.003	360	MP0533001	MP00165_0008	MP0036000
201	Rubisko	16/10	43.466	0.033	178	MP0533001	MP00165_0008	MP0036000
278	Rubisko	16/10	43.278	0.518	297	MP0533001	MP00165_0008	MP0036000
6001	Rubisko	16/10	43.386	0.150	217	MP0533001	MP00165_0008	MP0036000
6051	Rubisko	16/10	42.738	2.873	42	LR0116001	LR5561001	LR5116001
6061	Rubisko	16/10	48.548	7.640	150	AL5066000	AL0066500	AL0091000
6062	Rubisko	16/10	47.607	7.523	263	AL5066000	AL0066500	AL0091000
6082	Rubisko	16/10	47.927	7.407	207	AL5066000	AL0066500	AL0091000
6085	Rubisko	16/10	48.063	7.330	202	AL5066000	AL0066500	AL0091000
8001	Rubisko	16/10	47.750	7.290	245	AL5066000	AL0066500	AL0091000
8076	Rubisko	16/10	45.723	4.937	198	RA0066501	RA0016501	RA5016501
8082	Rubisko	16/10	45.747	5.083	235	RA0066501	RA0016501	RA5016501
1602	Soissons	21/10	46.218	4.518	525	RA0066501	RA0016501	RA5016501
1652	Soissons	21/10	45.634	4.458	580	RA0066501	RA0016501	RA5016501
1653	Soissons	21/10	45.988	4.723	173	RA0066501	RA0016501	RA5016501
1654	Soissons	21/10	47.787	6.363	271	FC5531001	FC0016501	FC5016001
1667	Soissons	21/10	47.710	5.888	255	FC5531001	FC0016501	FC5016001
1701	Soissons	21/10	46.297	4.800	216	BO55110_0042	BO0066501	BO0036002
1757	Soissons	21/10	46.784	4.848	178	BO55110_0042	BO0066501	BO0036002
1758	Soissons	21/10	47.947	0.200	51	PL0186500	PL0563501	PL0036000
1762	Soissons	21/10	48.348	0.368	135	PL0186500	PL0563501	PL0036000
1771	Soissons	21/10	45.633	5.885	235	RA0066501	RA0016501	RA5016501
1788	Soissons	21/10	45.560	5.975	291	RA0066501	RA0016501	RA5016501
7953	Soissons	21/10	45.563	5.795	375	RA0066501	RA0016501	RA5016501
7954	Soissons	21/10	45.906	6.108	426	RA0066501	RA0016501	RA5016501
7956	Soissons	21/10	49.383	1.181	151	HN0096001	HN0066501	HN0066001
7957	Soissons	21/10	49.610	1.541	160	HN0096001	HN0066501	HN0066001
7959	Soissons	21/10	49.935	1.095	33	HN0096001	HN0066501	HN0066001
7965	Soissons	21/10	49.712	1.074	137	HN0096001	HN0066501	HN0066001
7966	Soissons	21/10	48.610	2.680	91	IF0533501	IF0066003	IF0066001
7977	Soissons	21/10	49.062	2.703	113	IF0533501	IF0066003	IF0066001
8601	Soissons	21/10	48.783	2.200	174	IF0533501	IF0066003	IF0066001
8651	Soissons	21/10	48.750	2.116	154	IF0533501	IF0066003	IF0066001
8673	Soissons	21/10	48.773	2.008	167	IF0533501	IF0066003	IF0066001
8674	Soissons	21/10	46.261	-0.211	117	PC01330_0020	PC0036002	PC0036001
8676	Soissons	21/10	46.982	-0.215	80	PC01330_0020	PC0036002	PC0036001
401	Apache	21/10	46.494	-0.058	167	PC01330_0020	PC0036002	PC0036001
451	Apache	21/10	46.316	-0.400	59	PC01330_0020	PC0036002	PC0036001
601	Soissons	26/10	46.955	-0.591	164	PC01330_0020	PC0036002	PC0036001
1302	Soissons	26/10	46.858	-0.250	110	PC01330_0020	PC0036002	PC0036001
1305	Soissons	26/10	46.265	-0.010	124	PC01330_0020	PC0036002	PC0036001
1353	Soissons	26/10	46.458	-0.460	100	PC01330_0020	PC0036002	PC0036001
8302	Soissons	26/10	50.133	1.837	70	PI5533002	PI5533001	PI0066001
8303	Soissons	26/10	49.847	2.670	91	PI5533002	PI5533001	PI0066001
8401	Soissons	26/10	49.956	1.769	120	PI5533002	PI5533001	PI0066001
8402	Soissons	26/10	43.693	1.865	131	MP0533001	MP00165_0008	MP0036000
101	Soissons	16/10	43.911	2.118	172	MP0533001	MP00165_0008	MP0036000
152	Soissons	16/10	43.869	1.854	125	MP0533001	MP00165_0008	MP0036000
157	Soissons	16/10	44.029	1.378	107	MP0533001	MP00165_0008	MP0036000
160	Soissons	16/10	44.082	1.410	87	MP0533001	MP00165_0008	MP0036000

170	Soissons	16/10	43.382	6.383	80	PR5133000	PR5516000	PR5066000
198	Soissons	16/10	43.425	6.738	2	PR5133000	PR5516000	PR5066000
751	Soissons	21/10	44.140	4.853	53	PR5133000	PR5516000	PR5066000
2602	Soissons	21/10	44.083	5.058	99	PR5133000	PR5516000	PR5066000
2653	Soissons	21/10	46.450	-1.098	8	PL0186500	PL0563501	PL0036000
2661	Soissons	21/10	46.700	-1.383	90	PL0186500	PL0563501	PL0036000
2671	Soissons	21/10	46.621	-1.633	50	PL0186500	PL0563501	PL0036000
3801	Soissons	16/10	46.406	-1.401	14	PL0186500	PL0563501	PL0036000
3811	Soissons	16/10	46.583	0.300	117	PC01330_0020	PC0036002	PC0036001
3855	Soissons	16/10	46.440	0.122	150	PC01330_0020	PC0036002	PC0036001
4201	Apache	11/10	46.416	0.855	152	PC01330_0020	PC0036002	PC0036001
4279	Apache	11/10	46.685	0.246	121	PC01330_0020	PC0036002	PC0036001
6901	Soissons	21/10	46.796	0.781	130	PC01330_0020	PC0036002	PC0036001
6903	Soissons	21/10	45.866	1.183	402	LI0186501	LI0086501	LI0066501
6951	Soissons	21/10	46.215	1.083	209	LI0186501	LI0086501	LI0066501
6972	Soissons	21/10	48.302	6.123	289	LO5531004	LO5016500	LO0016501
6976	Soissons	21/10	47.823	3.548	152	BO55110_0042	BO0066501	BO0036002
7301	Apache	16/10	48.325	2.381	115	IF0533501	IF0066003	IF0066001
7351	Apache	16/10	48.730	2.111	175	IF0533501	IF0066003	IF0066001
7372	Apache	16/10	48.966	2.424	52	IF0533501	IF0066003	IF0066001
7492	Soissons	16/10	49.018	2.533	108	IF0533501	IF0066003	IF0066001

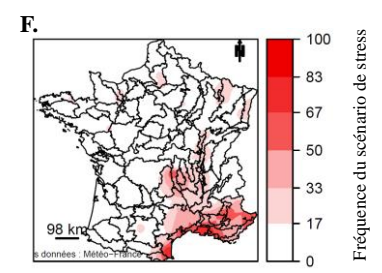
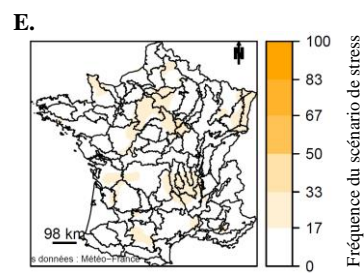
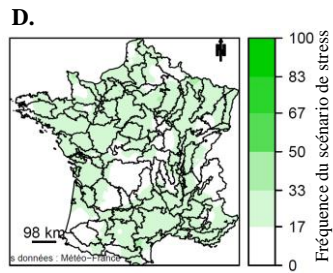
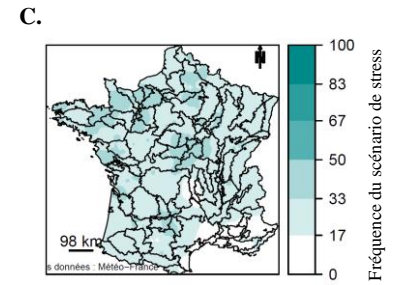
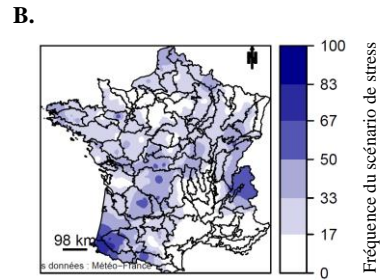
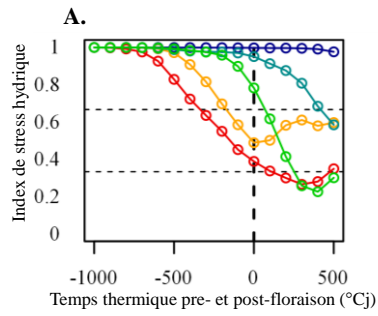
III.3B. Soil description

Soil 1, 2 or 3 have different available water capacities (AWC) which was chosen for French historical water stressed pattern simulation. Depth and AWC are presented in the following table as the physical properties of the first soil horizons. Depth: Maximal rooting depth in cm; AWC: soil available water capacity in mm; the following columns are the percentage of clay, fine silt (<0.02 mm), coarse silt (>0.02mm); fine sand (<0.2mm); coarse sand (>0.2mm); Organic matter.

SOIL	Id. Soil	Depth	AWC	Clay	Fine Silt	Coarse silt	Fine sand	Coarse sand	Organic matter
sol1	RA0066501	80	128	12	24	30	12	21	1
sol1	PI5533002	80	101	29	14	29	9	15	4
sol1	AU0033500	60	84	20	31	10	11	26	2
sol1	PR5133000	70	114	27	35	18	13	5	2
sol1	CH0086001	60	100	10	15	20	27	27	2
sol1	MP0533001	70	85	20	20	25	15	18	3
sol1	LR0116001	70	77	41	29	11	8	9	3
sol1	BN5531001	60	75	30	25	20	14	6	5
sol1	PC5133003	80	112	30	30	15	17	5	3
sol1	CE0093501	70	87	15	18	10	15	40	1
sol1	LI0186501	60	89	15	16	8	21	38	3
sol1	BO55110_0042	60	66	38	33	16	2	7	3
sol1	BR0086000	60	94	11	11	26	24	24	4
sol1	AQ5533001	80	103	30	25	20	14	6	5
sol1	FC5531001	60	78	33	24	36	3	2	3
sol1	HN0096001	60	61	4	2	3	20	70	1
sol1	PL0186500	60	81	15	16	8	21	38	2
sol1	LO5531004	45	53	34	37	10	4	11	5
sol1	NO0133001	90	129	24	23	23	14	14	2
sol1	AL5066000	120	186	23	31	41	2	1	2
sol1	IF0533501	70	90	18	25	46	4	5	2
sol2	RA0016501	100	99	35	20	20	11	11	3
sol2	PI5533001	100	131	18	3	60	12	5	3
sol2	AU0116001	100	155	33	27	31	3	4	4
sol2	PR5516000	100	134	30	31	18	10	8	3
sol2	CH0133001	90	130	24	23	23	14	14	2
sol2	MP00165_0008	120	145	28	19	18	20	14	1
sol2	LR5561001	100	108	10	25	17	26	20	2
sol2	BN5533001	80	84	30	25	20	14	6	5
sol2	PC0036002	90	119	25	40	30	2	1	2
sol2	CE5016001	100	133	30	12	8	23	25	3
sol2	LI0086501	80	121	10	10	10	37	30	3
sol2	BO0066501	100	124	15	22	30	11	20	2
sol2	BR0036500	100	145	18	22	18	16	24	1
sol2	AQ00860_0032	100	144	8	14	7	15	55	1
sol2	FC0016501	100	102	35	20	20	11	11	4
sol2	HN0066501	100	162	13	28	46	9	2	2
sol2	PL0563501	90	109	15	25	13	19	25	2
sol2	LO5016500	80	89	51	28	8	6	6	2
sol2	NO0066001	120	177	12	30	35	11	11	2
sol2	AL0066500	80	123	17	44	35	2	2	2
sol2	IF0066003	90	135	12	30	35	11	11	2
sol3	RA5016501	120	145	40	25	15	12	5	3
sol3	PI0066001	120	177	12	30	35	11	11	2
sol3	AU5016501	120	168	31	45	10	5	3	6
sol3	PR5066000	120	163	13	12	28	19	27	1
sol3	CH0036002	120	164	27	31	25	12	2	2
sol3	MP0036000	150	196	24	21	29	19	5	2
sol3	LR5116001	120	142	40	29	10	9	10	3
sol3	BN0066501	100	165	13	30	35	15	5	2
sol3	PC0036001	110	139	20	35	30	10	3	2
sol3	CE00365_0055	120	151	32	32	29	3	2	2
sol3	LI0066501	100	126	15	22	30	11	20	2
sol3	BO0036002	130	139	28	31	30	4	5	2
sol3	BR5016500	120	155	55	30	10	1	2	2
sol3	AQ0036001	120	165	22	32	14	18	11	3
sol3	FC5016001	120	162	35	25	15	17	5	3
sol3	HN0066001	120	177	12	30	35	11	11	2
sol3	PL0036000	110	139	20	35	30	10	3	2
sol3	LO0016501	100	117	59	24	8	4	2	4
sol3	NO0036001	140	238	22	27	41	4	4	2
sol3	AL0091000	50	63	8	8	6	14	65	1
sol3	IF0066001	120	177	12	30	35	11	11	2

III.3D. Presentation of water stress frequencies in France simulation over the last 20 years

- A. WSI all along the crop cycle for the five groups identified in the historical of water stress analysis. B, C, D, E, F are the spatial representation of the frequency for the different pattern.



Supplementary data III.4 : Statistical analysis, a specific model for a specific design.

Different statistical models were used to compute the adjusted means, depending on the experimental designs. To deal with the various experimental designs that were used, we decided to use the GxE two-stage approach describe by Frensham et al., (1997). All the models of the first stage analysis are presented in this table. A trial is defined by a combination of location, year and treatment. Column Ly et al. 2018 indicate the trial codification in Ly et al. 2018. The third and fourth columns describe the R-packages and the statistical models used for first stage trial analysis. Whatever the model, the genotypes were treated as fixed for calculation of the best linear unbiased estimator (BLUEs) or as random for the estimation of heritabilities.

<i>Trial</i>	<i>Ly et al. 2018</i>	<i>R-package</i>	<i>Statistical model</i>	<i>Nb Rep.</i>	<i>Design</i>
<i>Cle16RF</i>	-	ASREML-R	$Y_{ijk} = \mu + B_j + S_{k(check)} + G_i + \epsilon_{ijk}$	1	Augmented
<i>Cle16RO</i>	-	ASREML-R	$Y_{ijk} = \mu + B_j + S_{k(check)} + G_i + \epsilon_{ijk}$	1	Augmented
<i>Olm16RO</i>	-	ASREML-R	$Y_{ik} = \mu + S_{k(check)} + G_i + \epsilon_{ik}$	1	Augmented
<i>Cha12RF</i>	Châlons2012N+	ASREML-R	$Y_{jk} = \mu + S_{k(check)} + G_i + \epsilon_{jk}$	1	Augmented
<i>Rea13RF</i>	Réalville2013IRR	ASREML-R	$Y_{ijk} = \mu + R_j + S_{k(check)} + G_i + \epsilon_{ijk}$	4	Augmented
<i>Coi12IR</i>	-	ASREML-R	$Y_{ijk} = \mu + S_{k(check)} + G_i + \epsilon_{ijk}$	1.3	P-rep
<i>All14RF</i>	-	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Cap12RF</i>	Cappelle2012N+	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Cha14IR</i>	Champigny2014IRR	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Cha14RF</i>	Champigny2014DRY	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Cle13RF</i>	Clermont2013N+	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Lou12RF</i>	Louville2012N+	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Mau12RF</i>	Maule2012N+	SpATS	$Y_{ijk} = \mu + S_{jk} + G_i + \epsilon_{ik}$	2	Augmented
<i>Vra13RF</i>	Vraux2013N+	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Mon12RF</i>	Mons2012N+	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>And13RF</i>	Andelu2013N+	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Coi12RF</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	3	Augmented
<i>Gre12IR</i>	Gréoux2012IRR	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Gre12RF</i>	Gréoux2012DRY	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Gre13RF</i>	Gréoux2012DRY	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Mon13RF</i>	Mons2013N+	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Rec12RF</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Ver13RF</i>	Verneuil2013N+	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Wal14IR</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	RCBD
<i>Wal14RF</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	RCBD
<i>Wal16IR</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	RCBD
<i>Wal16RF</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	RCBD
<i>Wal17IR</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	RCBD
<i>Wal17RF</i>	-	SpATS	$Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$	2	RCBD
<i>Cas16IR</i>	-	SpATS	$Y_{ij} = \mu + R_j + G_i + \epsilon_{ij}$	2	RCBD
<i>Cas16RF</i>	-	SpATS	$Y_{ij} = \mu + R_j + G_i + \epsilon_{ij}$	2	RCBD
<i>Wal15IR</i>	-	SpATS	$Y_{ij} = \mu + R_j + G_i + \epsilon_{ij}$	2	RCBD
<i>Wal15RF</i>	-	SpATS	$Y_{ij} = \mu + R_j + G_i + \epsilon_{ij}$	2	RCBD
<i>Gre14IR</i>	Gréoux2014IRR	SpATS	$Y_{ijk} = \mu + R_j + RU + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented
<i>Gre14RF</i>	Gréoux2014DRY	SpATS	$Y_{ijk} = \mu + R_j + RU + S_{jk} + G_i + \epsilon_{ijk}$	2	Augmented

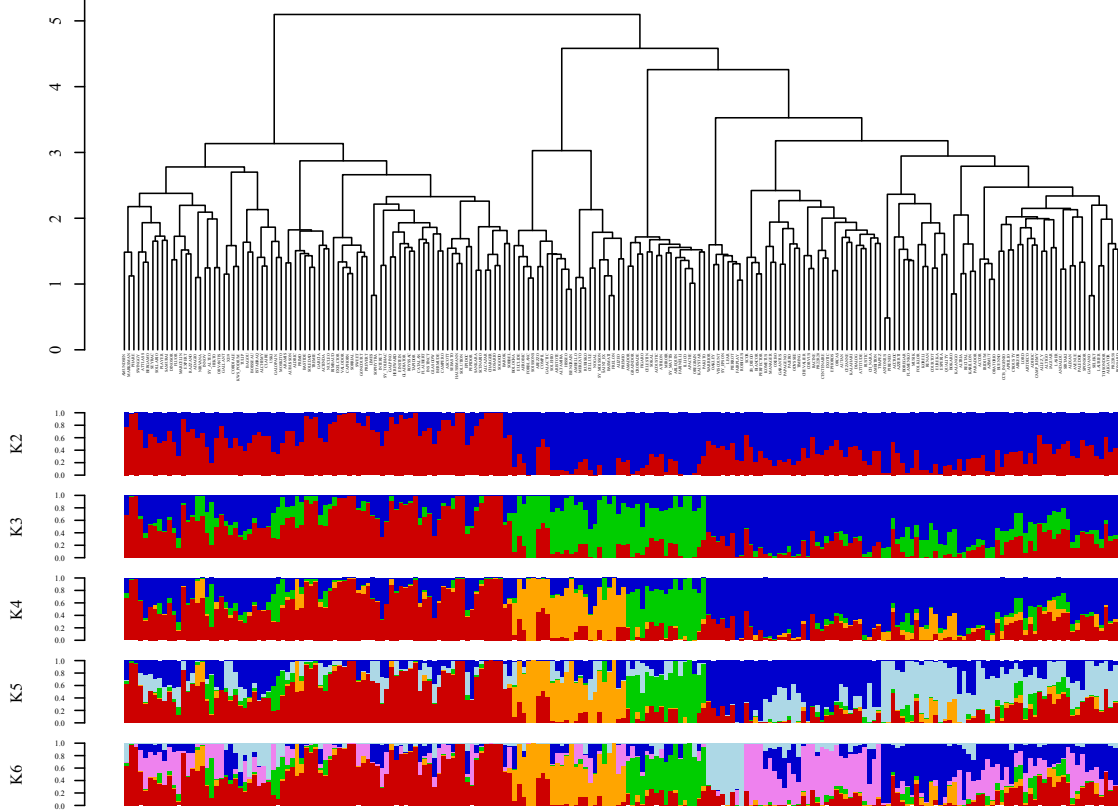
In each environment with two replicates and continuous design (28 trials), adjusted means and heritabilities were computed using 2D P-spline mixed models as implemented in the R package SpATS (Rodríguez-Álvarez et al. 2016). Three kinds of designs were used: the classical Randomized Complete Block Design (RCBD) with incomplete sub-block with check, a p-rep design where 30% of the varieties were randomly replicated two times, and an augmented design which was a RCBD with two blocks divided in six groups of earliness where four checks were present in each of the 12 sub-block. For 16 of the 35 trials the model

was $Y_{ijk} = \mu + R_j + S_{jk} + G_i + \epsilon_{ijk}$ with Y_{ijk} the observation of the replicate j of variety i , μ the environment effect, R_j the random effect of replicate j , S_{jk} the random effect of sub-block k in the replicate j (6 sub-blocks without replicate but with the four same checks repeated in each block). G_i the effect of variety i , ϵ_{ijk} are the residuals. For 6 trials, genotypes were not replicated in a continuous block, so replicate effect was discarded. As well as for the 4 trials without sub-block, where the S_{jk} was discarded. A quantitative covariate representing the available water capacity for each plot (RU) was added for Gre14IR and Gre14RF. For the remaining environments ASREML-R (Gilmour et al. 2006) was used. For the rain-out shelter experiment (Cle16IR and Cle16RF), adjusted means were estimated with the following statistical model $Y_{ijk} = \mu + B_j + S_{k(check)} + G_i + \epsilon_{ijk}$, with the same notation than before and with B_j the random effect of shelter.

Following Comier et al. 2014 for exploring genetic gain, the second order polynomial regression was used to correct adjusted mean for genotype heading dates by adding $b1_i HD_i + b2_i HD_i^2$, with $b1_i$ is the linear trend component for genotype i , $b2_i$ is the quadratic trend component for genotype i and HD the Heading Date of genotype i .

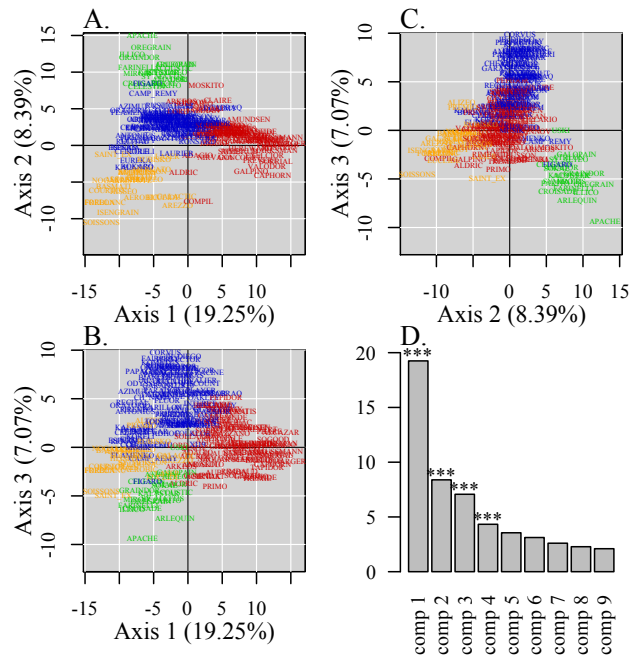
Supplementary data III.5 : Genetic relationship and structure of an European winter wheat panel

III.5A. Dendrogram derived from the kinship matrix and assignation of varieties to groups estimated by STRUCTURE software for K varying from 2 to 6



III.5B. Principal Coordinate Analysis based on the kinship matrix of the 210 Wheat.

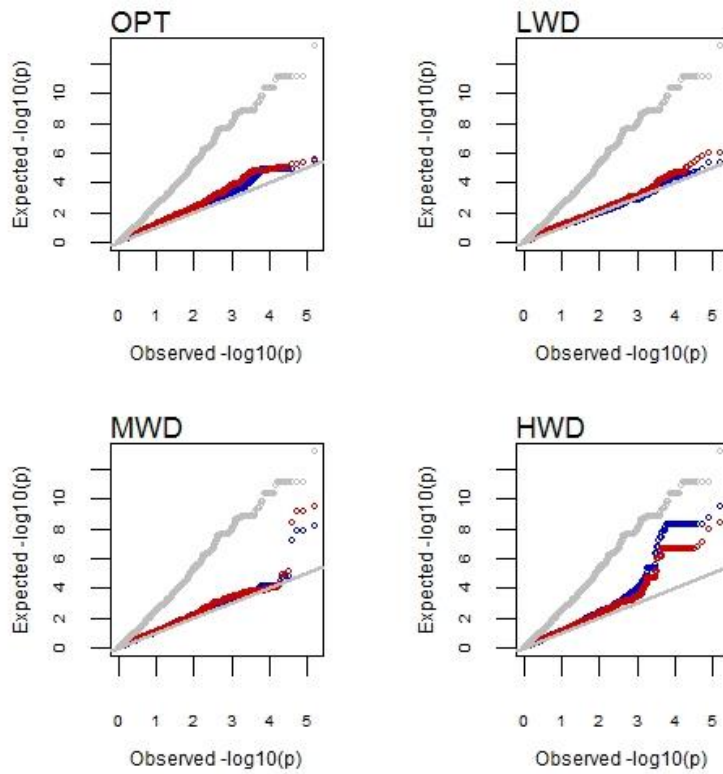
The subdivision of the varieties, assuming $K=4$ (based on Evanno¹ and the number of significant axis detected with the Tracy Widom test as Patterson et al. (2006)²



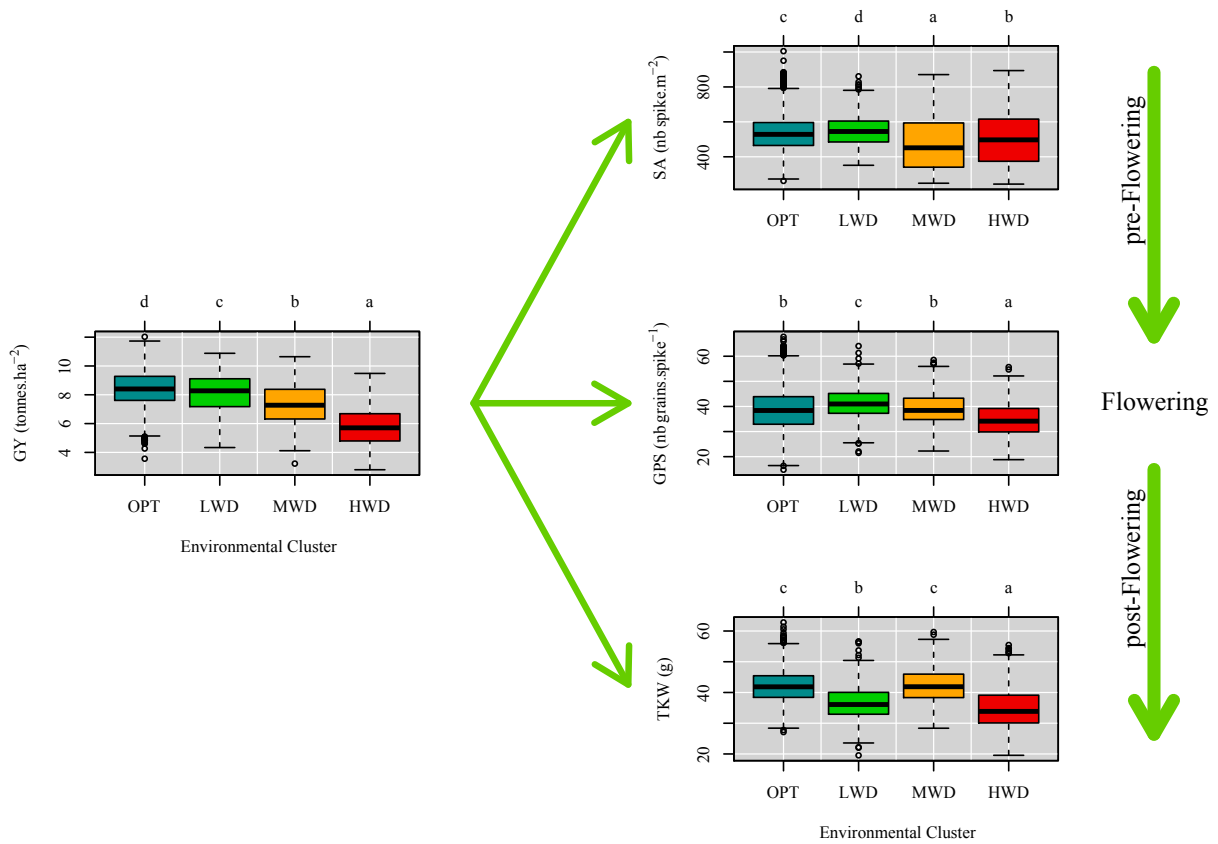
¹ Evanno, Regnaut, et Goudet, « Detecting the Number of Clusters of Individuals Using the Software Structure ».

² Patterson, Price, et Reich, « Population Structure and Eigenanalysis ».

III.5C. QQ-plots of the grain yield association in the four environments clusters (optimal (OPT), low water deficit (LWD), mild water deficit (MWD) and high water deficit (HWD)) for naive model (grey points), model K (red points) and K + PCA₁₋₄ (blue points)



•



Supplementary data III.6: Boxplot of grain yield and yield components in the four environmental clusters.

SA: Spike Per Area; GPS: Grain Per Spike; TKW: Thousand Kernel Weight; GY: Grain Yield. Horizontal and slanting arrows represent the grain yield break down and vertical arrows represent the development timing of component regarding the flowering times. The four different Environmental clusters (EC) related to drought stress were presented with different colors, OPT: Optimal condition (blue); LWD: Late Water Deficit (green), MWD: Medium Water Deficit (orange), HWD: High Water Deficit (red). Quartiles and median are used to construct the box. The whisker extends to 1.5 times the interquartile range from the box. Boxplots with the same letter above them are not significantly different ($\alpha = 0.001$), based on Tukey HSD P-value.

Supplementary data III.7: Correlation table between trait BLUEs over the four environmental clusters.

Lower triangle shows the Spearman's coefficient and upper triangle shows the Pearson's coefficient.

	OPT						LWD						MWD						HWD						
	HD	PH	SA	GPS	TKW	GY	HD	PH	SA	GPS	TKW	GY	HD	PH	SA	GPS	TKW	GY	HD	PH	SA	GPS	TKW	GY	
OPT	HD	1	0.4***	-0.12	0.41***	-0.18	0.53***	0.97***	0.14	0.11	0.2	-0.6***	-0.7***	0.95***	-0.15	0.09	0.18	-0.54***	-0.36***	0.92***	-0.24***	-0.27***	-0.06	-0.39***	-0.47***
	PH	0.42***	1	-0.21**	0.01	0.18**	0.1	0.37***	0.6***	-0.11	-0.06	-0.09	-0.37***	0.38***	0.65***	-0.19	-0.03	0	-0.42***	0.37***	0.6***	-0.33***	0.01	0.03	-0.13
	SA	-0.13	-0.24	1	-0.32***	-0.28**	0.13	-0.12	-0.08	0.56***	-0.27**	-0.18	-0.04	-0.1	-0.21	0.31***	-0.26*	-0.16	0.03	-0.09	-0.17*	0.51***	-0.18**	-0.25***	0.1
	GPS	0.41***	0.03	-0.32***	1	-0.44***	0.48***	0.42***	0	-0.12	0.56***	-0.5***	-0.2	0.33***	-0.11	-0.05	0.64***	-0.42***	0.07	0.36***	-0.18**	-0.3***	0.41***	-0.35***	-0.14*
	TKW	-0.17*	0.22	-0.28***	-0.45***	1	-0.1	-0.16	0.05	-0.22	-0.41***	-0.71***	0.19	-0.15	0.27*	-0.17	-0.42***	0.7***	0.07	-0.14*	0.3***	-0.09	-0.3***	0.69***	0.21**
	GY	0.5***	0.04	0.1	0.55***	-0.07	1	0.53***	-0.02	0.21	0.16	-0.39***	-0.2	0.48***	-0.14	0.15	0.23	-0.3**	0.16	0.48***	-0.24***	-0.01	0.05	-0.28***	-0.09
LWD	HD	0.96***	0.39***	-0.11	0.42***	-0.17*	0.51***	1	0.11	0.1	0.21	-0.59***	-0.67***	0.93***	-0.17	0.08	0.19	-0.53***	-0.35***	0.9***	-0.26***	-0.26***	-0.04	-0.38***	-0.43***
	PH	0.19**	0.64***	-0.11	-0.01	0.11	-0.01	0.18**	1	0.06	-0.14*	-0.05	-0.18*	0.16	0.46***	-0.08	-0.06	0.02	-0.2	0.13	0.48***	-0.12	-0.03	0.01	-0.08
	SA	0.11	-0.16*	0.58***	-0.13	-0.24***	0.2**	0.12	0	1	-0.47***	-0.25	-0.05	0.13	-0.18	0.42***	-0.2	-0.24	0.09	0.11	-0.18*	0.49***	-0.23***	-0.23***	0.03
	GPS	0.21**	-0.05	-0.28***	0.6***	-0.42***	0.23***	0.24***	-0.08	-0.49***	1	-0.5***	0.01	0.15	-0.21	-0.24	0.59***	-0.36***	-0.01	0.14*	-0.28***	-0.37***	0.51***	-0.39***	-0.12
	TKW	-0.61***	-0.11	-0.18**	-0.49***	0.69***	-0.36***	-0.62***	-0.1	-0.26***	-0.49***	1	0.52***	-0.55***	0.28**	-0.11	-0.39***	0.78***	0.24	-0.53***	0.36***	0.08	-0.22**	0.76***	0.35***
	GY	-0.68***	-0.44***	-0.01	-0.12	0.18*	-0.07	-0.66***	-0.17	0	0.05	0.51***	1	-0.67***	0.02	-0.05	-0.03	0.46***	0.5***	-0.68***	0.07	0.22**	0.08	0.38***	0.49***
MWD	HD	0.93***	0.4***	-0.09	0.3***	-0.11	0.47***	0.9***	0.19**	0.14*	0.13	-0.53***	-0.63***	1	-0.17	0.2	0.11	-0.54***	-0.32***	0.94***	-0.25***	-0.15*	-0.11	-0.38***	-0.42***
	PH	-0.06	0.75***	-0.23***	-0.09	0.3***	-0.18*	-0.1	0.51***	-0.23***	-0.18**	0.24***	-0.1	-0.09	1	-0.25***	-0.08	0.34***	-0.1	-0.17*	0.87***	-0.23***	0.02	0.32***	0.12
	SA	0.04	-0.23***	0.28***	-0.11	-0.12	0.13	0.04	-0.15*	0.39***	-0.29***	-0.05	-0.03	0.2**	-0.28**	1	-0.18	-0.4***	0.33***	0.12	-0.17*	0.6***	-0.28***	-0.29***	-0.21**
	GPS	0.17*	-0.08	-0.23***	0.66***	-0.43***	0.3***	0.19**	-0.07	-0.2**	0.62***	-0.39***	0.03	0.05	-0.14	-0.23***	1	-0.39***	0.24***	0.14	-0.1	-0.27***	0.46***	-0.36***	-0.18*
	TKW	-0.54***	-0.01	-0.16*	-0.42***	0.72***	-0.26***	-0.53***	0.05	-0.23***	-0.35***	0.76***	0.46***	-0.49***	0.3**	-0.35***	-0.41***	1	0.13	-0.48***	0.37***	-0.08	-0.11	0.8***	0.55***
	GY	-0.35***	-0.47***	0.06	0.1	0.06	0.26***	-0.32***	-0.22**	0.11	0	0.22**	0.53***	-0.26***	-0.23	0.4**	0.26*	0.12	1	-0.36***	-0.07	0.35***	0.03	0.08	0.2**
HWD	HD	0.91***	0.38***	-0.07	0.34***	-0.11	0.5***	0.88***	0.15	0.14	0.15	-0.53***	-0.62***	0.94***	-0.1	0.13	0.08	-0.44***	-0.29***	1	-0.26*	-0.17	-0.07	-0.37***	-0.38***
	PH	-0.18	0.66***	-0.18	-0.14	0.28**	-0.25*	-0.22	0.48***	-0.2	-0.24	0.31***	-0.02	-0.19	0.86***	-0.21	-0.14	0.32***	-0.22	-0.19**	1	-0.15*	0.04	0.36***	0.17*
	SA	-0.23	-0.32***	0.45***	-0.26*	-0.08	0.02	-0.23	-0.17	0.47***	-0.36***	0.08	0.19	-0.01	-0.25*	0.67***	-0.27*	-0.04	0.37***	-0.01	-0.15	1	-0.41***	-0.14	0.19**
	GPS	-0.03	-0.03	-0.18	0.43***	-0.3**	0.13	-0.01	-0.07	-0.23	0.5***	-0.26*	0.1	-0.01	-0.02	-0.3**	0.45***	-0.1	-0.01	-0.04	0.01	-0.36***	1	-0.21	0.34***
	TKW	-0.41***	-0.01	-0.26*	-0.36***	0.68***	-0.25*	-0.42***	0	-0.25*	-0.36***	0.76***	0.37***	-0.39***	0.25*	-0.27*	-0.35***	0.8***	0.06	-0.38***	0.31***	-0.12	-0.2**	1	0.44***
	GY	-0.47***	-0.19	0.11	-0.08	0.16	-0.01	-0.43***	-0.09	0.05	-0.09	0.33***	0.5***	-0.34***	0.04	-0.11	-0.16	0.53***	0.21	-0.28***	0.13	0.33***	0.35***	0.4***	1

Supplementary data III.8 : Years of registration effects on grain yield and its components in the environmental clusters.

III.8A. Anova table for grain yield from eq. 6

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
ETs	3	624980535	208326845	722.5963	< 2.2e-16***
ETs:Years	4	10359686	2589922	8.9833	4.218e-07***
Residuals	812	234102220	288303		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

III.8B. Summary table of the ANOVA on grain yield

	Estimate	Std. Error	t value	Pr(> t)
Intercept	5663.799	142.939	39.624	< 2e-16***
ETs: OPT	2136.713	202.147	10.57	< 2e-16***
ETs: LWD	2256.441	202.147	11.162	< 2e-16***
ETs: MWD	1922.361	202.147	9.51	< 2e-16***
OPT: Years	38.072	8.903	4.276	2.13E-05***
LWD: Years	8.606	8.903	0.967	0.334023
MWD: Years	19.625	8.903	2.204	0.027783*
HWD: Years	30.653	8.903	3.443	6.05E-04***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

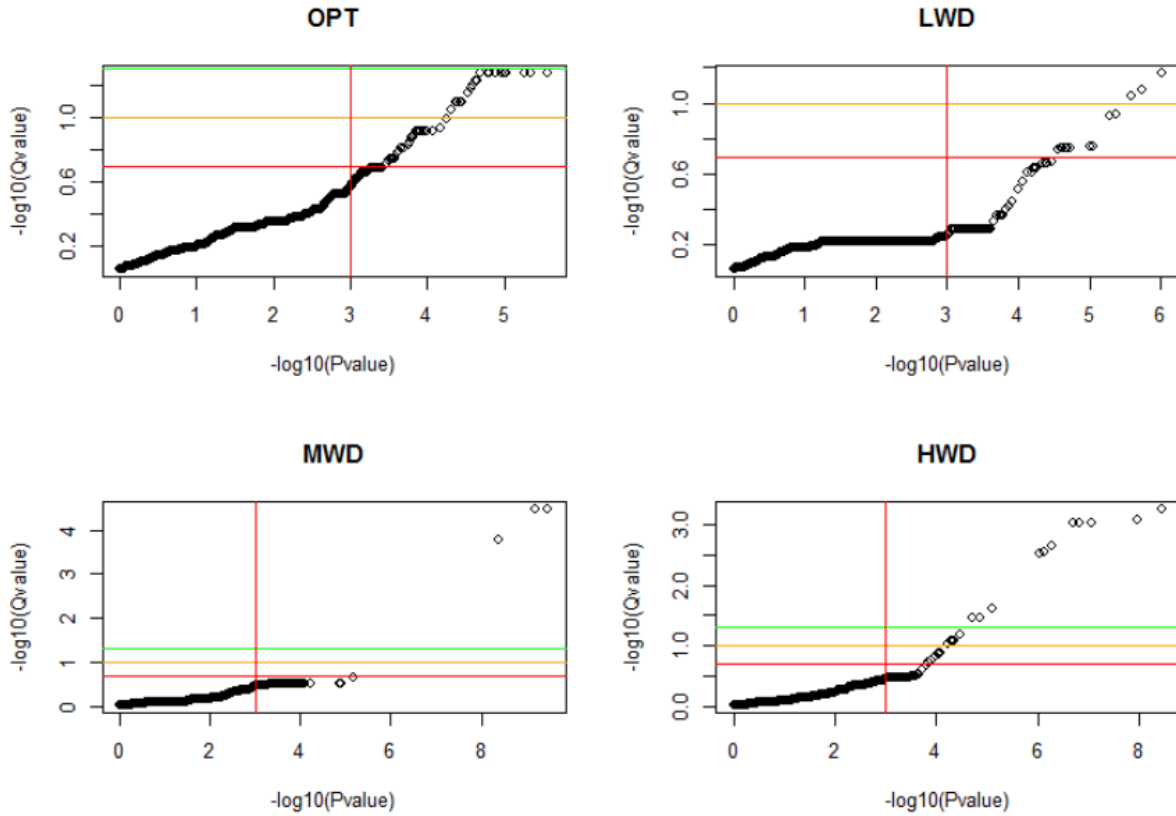
III.8C. Coefficient from eq. 6 for other traits than grain yield

Effect	PH	SA	GPS	TKW
Intercept	33.432***	538.362***	33.653***	33.432***
ETs: OPT	8.754***	9.251	2.441.	8.754***
ETs: LWD	2.689*	13.937	8.865***	2.689*
ETs: MWD	8.972***	2.434	3.485*	8.972***
OPT: Years	0.005	0.694	0.127*	0.005
LWD: Years	-0.011	1.497	-0.059	-0.011
MWD: Years	-0.009	1.075	0.079	-0.009
HWD: Years	-0.027	1.339	0.059	-0.027

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

III.9B. Relationships between SNP P-value and Q-value (from FDR) for grain yield in the four environments types

The three horizontal lines represent the FDR threshold at 0.05, 0.1 and 0.2 in green, orange and red, respectively. OPT: optimal condition; LWD: late water deficit; MWD: medium water deficit; HWD: high water deficit.



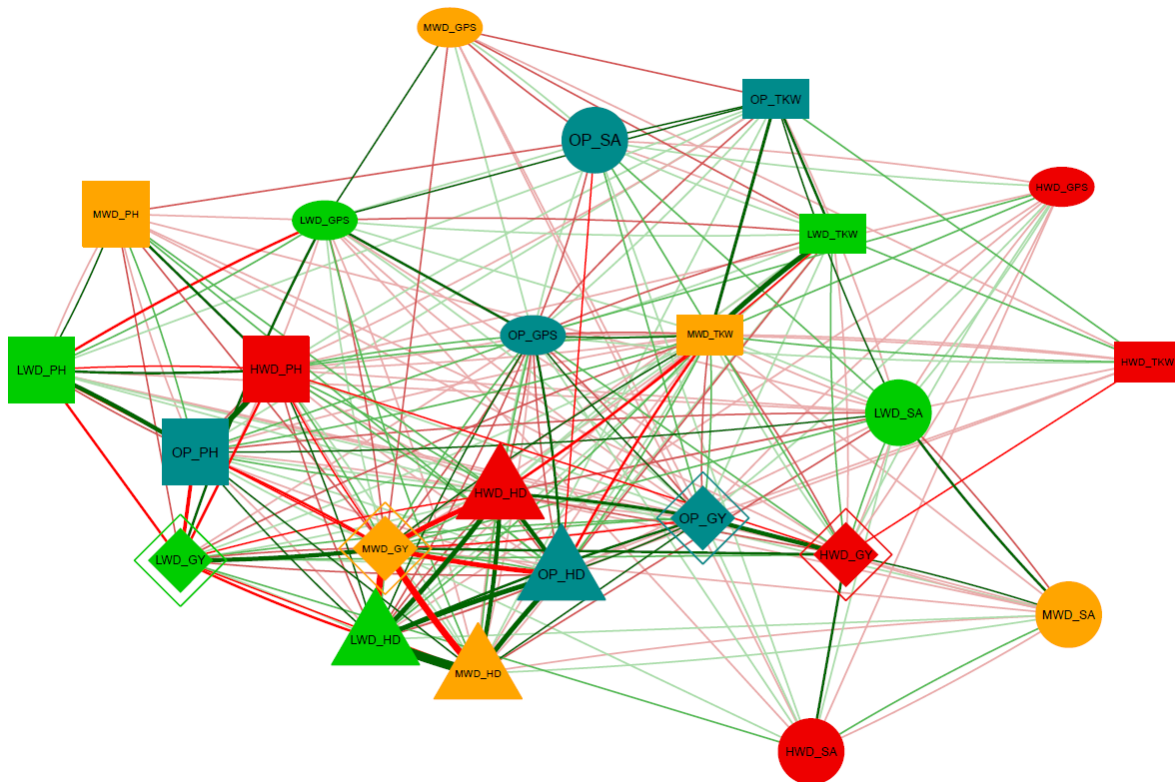
III.9C. Distribution of percentage of SNP with $-\log_{10}(P\text{-value})$ lower than 0.001 at different level of FDR threshold

This table gives the percentage of SNPs with a P-value lower than 0.001 for GWAS model and Qval from FDR test lower than different threshold (*e.g.* for heading date in OPT on the 213 SNP with $P\text{-val}<0.001$ only 11.74% have a $Q\text{-val}<0.2$). ETs: Environment types with (OPT) optimal condition, (LWD) late water deficit, (MWD) moderate water deficit and (HWD) high water deficit. Pheno: GY: grain yield, SA: spike per area, GPS: grain per spike, TKW: thousand kernel weights, HD: heading date, PH: plant height.

ETs	Pheno	nbSNP	Qval<0.05	Qval<0.1	Qval<0.2	Qval<0.3	Qval<0.4	Qval<0.5	Qval<0.6	Qval<0.7	Qval<0.8	Qval<0.9
OPT	HD	213	3.76	8.92	11.74	13.15	100	100	100	100	100	100
OPT	PH	521	6.91	7.49	65.83	100	100	100	100	100	100	100
OPT	SA	173	0	0	0	0	0	0	100	100	100	100
OPT	GY	518	0	14.86	47.68	100	100	100	100	100	100	100
OPT	TKW	1380	51.88	100	100	100	100	100	100	100	100	100
OPT	GPA	1439	80.96	99.72	100	100	100	100	100	100	100	100
OPT	GPS	501	0	0	0	100	100	100	100	100	100	100
LWD	HD	136	2.94	2.94	2.94	12.50	34.56	36.03	38.97	47.06	50.00	100
LWD	PH	335	17.01	22.69	24.48	48.96	77.61	100	100	100	100	100
LWD	SA	237	0	0	0	0	100	100	100	100	100	100
LWD	GY	246	0	1.63	8.54	17.89	20.73	26.83	100	100	100	100
LWD	TKW	1122	68.89	94.12	100	100	100	100	100	100	100	100
LWD	GPA	807	16.48	41.26	100	100	100	100	100	100	100	100
LWD	GPS	502	0	0	0	100	100	100	100	100	100	100
MWD	HD	174	2.30	4.02	5.17	5.17	9.20	100	100	100	100	100
MWD	PH	254	0	0	0	1.97	46.06	77.56	99.61	100	100	100
MWD	SA	168	0	0	0	0	0	100	100	100	100	100
MWD	GY	449	0.89	0.89	0.89	1.56	100	100	100	100	100	100
MWD	TKW	489	0.82	0.82	0.82	100	100	100	100	100	100	100
MWD	GPA	1206	0.33	51.41	100	100	100	100	100	100	100	100
MWD	GPS	183	0	0	0	0	0	100	100	100	100	100
HWD	HD	484	13.22	14.67	15.91	100	100	100	100	100	100	100
HWD	PH	370	10.27	11.08	20	54.59	95.95	100	100	100	100	100
HWD	SA	66	0	0	0	0	0	0	0	0	0	100
HWD	GY	439	18.68	22.32	26.65	27.79	100	100	100	100	100	100
HWD	TKW	732	0	0	90.44	100	100	100	100	100	100	100
HWD	GPA	1050	6.95	67.81	100	100	100	100	100	100	100	100
HWD	GPS	216	6.94	12.50	30.09	62.04	64.35	100	100	100	100	100

- **Supplementary data 10: QTL colocalization network for 6 traits measured on the 210 lines among the four environment types.**

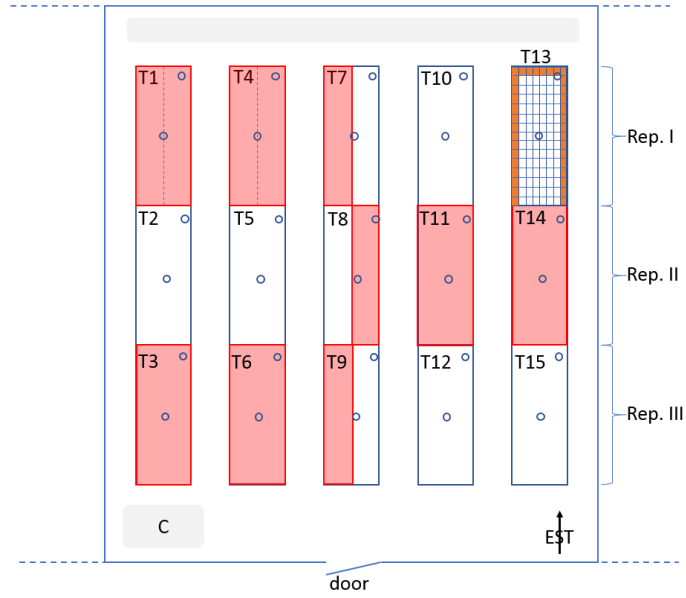
This network is based on the percentage of common QTL between traits within environment types (ETs). The shape refers to the trait as: plant height in square, heading date in triangle, grain yield in diamond, spike per area in circle, grain per spike in ellipse, and thousand kernel weights in rectangle. Similarly, color refers to the different ETs as: (OPT) optimal water condition in blue, (LWD) late water deficit in green, (MWD) medium water deficit in orange, and (HWD) high water deficit in red. Lines between QTL was corrected using a hypergeometric law to determine significant colocalization ($P < 0.001$). Lines thickness is a function of the percentage of common QTL, from 5% for thinnest to 100% for the thickest (cf table 6). Common QTL with common effect sign are in green, while opposite effects are in red.



Annexes IV

• **Supplementary data IV.1 : Wheat varieties tested, year of release (YR) and country where first registered.**

Variety	YR	Origin	Variety	YR	Origin	Variety	YR	Origin
ACCOR	2007	France	DINOSOR	2005	France	OREGRAIN	2011	France
ACCROC	2010	France	EPHOROS	2004	France	ORVANTIS	2000	France
ACIENDA	2004	France	EPIDOC	2006	France	OXEBO	2010	France
ACOUSTIC	2010	France	ESKET	2007	Germany	PAJERO	1995	Belgium
ADAGIO	2009	France	EUCLIDE	2007	France	PAKITO	2010	France
ADHOC	2010	France	EUREKA	1992	France	PALEDOR	2005	France
AEROBIC	2009	France	EXELCIOR	2007	France	PAPAGENO	2007	Austria
ALCAZAR	2004	France	EXOTIC	2005	France	PARADOR	2000	France
ALCHEMY	2005	UK	EXPERT	2007	France	PEPIDOR	2007	France
ALDRIC	2007	France	FAIRPLAY	2011	France	PERFECTOR	2004	France
ALIGATOR	2010	France	FARINELLI	2010	France	PHARE	2007	France
ALIXAN	2005	France	FIGARO	2011	France	PIERROT	2009	Danemark
ALIZEO	2010	France	FLAMENKO	2010	France	PIRENEO	2005	Austria
ALLEZ_Y	2010	France	FLAUBERT	2010	France	PLAYER	2010	France
ALTAMIRA	2009	France	FLUOR	2010	France	PR22R20	2002	France
ALTIGO	2007	France	FOLKLOR	2010	France	PREMIO	2007	France
ALTRIA	1996	France	FORBLANC	2009	France	PREVERT	2010	France
AMADOR	2010	France	FRELON	2000	France	PRIMO	2009	Italy
AMBELLO	2010	France	GALACTIC	2007	France	RACINE	2010	France
AMBITION	2004	Danemark	GALLANT	2007	UK	RAZZANO	2010	France
AMUNDSEN	2007	France	GALOPAIN	2009	France	RENAN	1989	France
ANDALOU	2001	France	GALPINO	2010	Spain	RIMBAUD	2010	France
ANTONIUS	2006	France	GALVANO	2008	Danemark	ROBIGUS	2002	UK
APACHE	1998	France	GARANTUS	2007	Poland	RONCARD	2011	France
APRILIO	2010	France	GARCIA	2006	France	ROSARIO	2004	France
ARAMIS	2009	France	GLADIATOR	2002	UK	ROYSSAC	2002	France
AREZZO	2007	France	GLASGOW	2004	UK	RUBISKO	2011	France
ARISTOTE	2010	France	GONCOURT	2009	France	RUSTIC	2005	Belgium
ARKEOS	2010	France	GRAINDOR	2006	France	SAINT_EX	2010	France
ARLEQUIN	2007	France	HAUSSMANN	2006	France	SAMURAI	2005	France
ARTDECO	2011	France	HEREWARD	1992	France	SANKARA	2004	France
ASCOTT	2011	France	HISSEO	2010	Italy	SCENARIO	2010	France
ATHLON	2010	France	ILICO	2010	France	SCOR	2009	France
ATTITUDE	2009	France	INNOV	2005	France	SEBASTO	2007	France
ATTLASS	2005	France	INSTINCT	2006	France	SELEKT	2007	France
AUBUSSON	2001	France	INTERET	2007	France	SEYRAC	2006	France
AUTAN	2000	France	IRIDIUM	2007	France	SIRTAKI	2008	Italy
AZZERTI	2010	France	ISENGRAIN	1997	France	SOBBEL	2009	France
BAGOU	2007	France	JAGUAR	2009	Italy	SOGOOD	2006	France
BAROK	2009	France	JB_DIEGO	2006	UK	SOISSONS	1988	France
BASTIDE	2002	France	KALAHARI	2010	Germany	SOKAL	2010	France
BERGAMO	2011	France	KALANGO	2002	Italy	SOLEDAD	2008	Italy
BERMUDE	2007	France	KALYSTAR	2010	France	SOLEHIO	2009	France
BIANCOR	2010	France	KARILLON	2010	France	SOLLARIO	2007	France
BOISSEAU	2007	France	KOMETUS	2011	Germany	SOLUTION	2007	France
BOLOGNA	2002	Spain	KORELI	2007	France	SOPHYTRA	2007	France
BOREGAR	2007	France	KWS_PODIUM	2010	UK	SORRIAL	2009	France
BUENNO	2007	France	LASER	2010	Italy	SPONSOR	1995	France
CAMP_REMY	1980	France	LAURIER	2011	France	SWEET	2010	France
CAPHORN	2000	France	LEAR	2007	UK	SWINGGY	2009	France
CATALAN	2002	France	LIMES	2002	France	SY_ALTEO	2010	France
CCB_INGENIO	2006	Spain	LORD	2009	France	SY_EPSON	2011	France
CELESTIN	2010	France	MANAGER	2006	France	SY_MATTIS	2010	France
CELLULE	2011	France	MARCELIN	2009	France	SY_MOISSON	2011	France
CENTENAIRE	2002	Belgium	MARKSMAN	2006	UK	SY_TOLBIAC	2011	France
CEZANNE	1998	France	MERCATO	2005	France	TAPIDOR	2001	France
CH_NARA	2009	Switzerland	MH_09.17	NA	NA	TIMING	2010	France
CHARGER	1997	France	MIROIR	2010	France	TOISONDOR	2004	France
CHEVALIER	2006	Austria	MOSKITO	2011	France	TRAPEZ	2009	France
CLAIRE	1997	UK	MUSIK	2010	France	TREMIÉ	1992	France
COMPIL	2010	France	NIRVANA	2001	France	TULIP	2010	France
CORDIALE	2005	France	NUCLEO	2010	France	USKI	2009	France
CORVUS	1998	Belgium	OAKLEY	2006	UK	VALODOR	2007	France
COURTOT	1974	France	ODYSSEE	2011	France	VISCOUNT	2007	UK
CROISADE	2010	France	ORATORIO	1996	France	WARRIOR	2009	UK
CROUSTY	1995	France	ORCAS	2010	Germany	XII9	2001	UK
DIALOG	2007	France						



- Supplementary data IV.2 : Representation of the experimental design in the greenhouse.**

Split-plot design, with sub-plots nested into 5 sub-blocks nested in turn into 2 whole-plots arranged in 3 replicates on 15 tables. Whole-plot (or treatment) arrangements are represented in light red and white. Sub-plot (or genotype) arrangements inside the table are represented on table 13, each table containing 120 pots. The NB1 genotype was positioned around the table to limit border effects (orange pots). Each table was split into two incomplete sub-blocks (e.g. tables 1 and 4) in which three checks were positioned at random. Thermo-buttons screening the temperature evolution are represented in open blue circles. The grey areas on the top and bottom left-hand side represent the air extractors.

• **Supplementary data IV.3 : Variance components for each trait in the 2016 and 2017 experiments**

Mean is the mean of fitted values form Mod [3], Sd is the standard deviation for fitted values, h^2 is heritability, SI is the stress index, ssT is the sum of squares explained by the treatment, varG is the variance explained by the genotype, varGxT is the variance explained by the GxT interaction and varRes the residual variance.

		Units	Mean	H ²	SI	ssT	varG	varGT	varRes			
2016	Flowering Date	D.Z65	Days	174.12	0.95	0.00	3.86E-04	26.88***	7.71E-07	7.69		
		D	$\mu\text{g.cm}^{-3}$	4.20	0.23	5.54	1.44	0.46.	0.7**	7.35		
	Senescence Parameters	K	$\mu\text{g.cm}^{-3}$	45.56	0.75	-1.31	21.67	8.94***	1.94E-06		17.85	
		a		2.32E-04	1.11E-02	-192.44	5.95E-06	6.91E-09	2.26E-13		3.71E-06	
		r		2.66E-02	0.30	-5.31	3.97E-05	6.12e-06**	3.76e-06*		7.26E-05	
		XPI	$^{\circ}\text{C.days}$	613.81	0.79	6.89	19288.11*	3174.94***	135.47		4672.52	
		YPI	$\mu\text{g.cm}^{-3}$	26.98	0.75	-0.21	8.23E-03	2.87***	0.35**		4.58	
		VIT	per $^{\circ}\text{C.days}$	-0.30	0.50	-7.38	6.69E-03	1.51e-03***	4.88e-04*		0.01	
	Senescence traits	START	$^{\circ}\text{C.day}$	516.95	0.72	7.54	10990.11	3162.87***	583.09**		5594.55	
		fSTART	$\mu\text{g.cm}^{-3}$	45.58	0.84	-0.62	1.19	8.13***	0.17		8.88	
		END	$^{\circ}\text{C.day}$	710.67	0.77	6.44	54812.36**	3297.82***	2.69E-04		5921.87	
		fEND	$\mu\text{g.cm}^{-3}$	8.38	0.29	2.14	0.77	0.55*	0.65**		6.16	
		AUC		24901.76	0.82	5.82	22360141.43	10314251.87***	6.98		13856378.54	
		AUCK		30852.30	0.85	6.41	29393172.45	15566752.51***	9.04		15901778.66	
	2017	Number of Spikes	SPP		4.54	0.78	3.14	0.83	0.94***	0.05	1.45	
			NbGrains		42.18	0.82	5.13	150.86	47.03***	1.47	58.79	
		Main Shoot	GW	g	1.71	0.73	29.31	2.91***	0.07***	0.01***		0.11
			TKW	g	40.92	0.66	25.73	1721.39***	15.02***	4.35***		33.97
			Grain.Area		14.89	0.74	15.28	75.15***	0.97***	0.15***		1.54
			Grain.Width		3.46	0.69	11.05	1.92***	0.02***	4.21e-03***		0.03
Grain.Length				5.90	0.85	4.30	1.16***	0.05***	2.84e-03**		0.04	
			NbGrains		139.71	0.78	8.26	6179.54.	1176.72***	51.9		1795.69
Tillers		GW	g	5.52	0.71	26.51	47.17***	1.54***	0.21**		3.09	
		TKW	g	39.79	0.69	20.40	730.52***	13.13***	2.98***		26.30	
		Grain.Area	mm ²	14.25	0.77	12.88	35.23***	0.93***	0.14***		1.28	
		Grain.Width	mm ²	3.40	0.73	8.44	0.91***	0.02***	2.81e-03***		0.03	
		Grain.Length	mm ²	5.77	0.84	4.65	0.74***	0.05***	5.28e-03***		0.04	
			NbGrain		136.33	0.70	13.92	17145.46**	808.25***	71.88		1883.59
2017	Flowering Date	D.Z65	Days	177.22	0.95	-0.12	3.41	32.49***	1.08E-06	10.49		
		D	$\mu\text{g.cm}^{-3}$	6.00	0.13	-1.01	0.38	0.13.	2.15E-07	4.97		
	Senescence Parameter	K	$\mu\text{g.cm}^{-3}$	46.72	0.79	0.78	19.02	10.55***	4.26E-06		16.58	
		a		1.00E-03	4.24E-02	-8.66	9.09E-07	2.66E-07	2.74E-07		3.55E-05	
		r		1.89E-02	0.45	3.50	9.27E-06	8.77e-06***	7.67E-07		6.20E-05	
		XPI	$^{\circ}\text{C.days}$	700.63	0.77	0.98	1388.29	5385.5***	401.45		8465.80	
		YPI	$\mu\text{g.cm}^{-3}$	29.36	0.80	0.35	1.34	2.58***	4.03E-07		3.97	
		VIT	per $^{\circ}\text{C.days}$	-0.22	0.54		1.71E-03	1.65e-03***	1.24E-04		0.01	
	Stay-green Trait	START	$^{\circ}\text{C.day}$	547.95	0.75	3.38	14542.59	3890.84***	2.31E-03		7943.47	
		fSTART	$\mu\text{g.cm}^{-3}$	48.43	0.85	0.50	9.96	8.54***	9.33E-07		9.20	
		END	$^{\circ}\text{C.day}$	853.31	0.70	-0.59	630.34	8468.4***	788		19708.98	
		fEND	$\mu\text{g.cm}^{-3}$	10.29	0.20	-0.23	0.06	0.17*	2.41E-06		4.17	
		AUC		28782.44	0.81	2.00	17372593.86	13775145.15***	86987		19515502.45	
		AUCK		37491.58	0.82	1.54	8239642.45	21997032.13***	1136065.62		25439236.26	
2017	Number of Spikes	SPP		5.04	0.71	4.85	7.89*	0.81***	0.02	1.90		
		NbGrain		36.74	0.77	9.25	144.31	43.12***	8.03***		54.69	
	Main Shoot	GW	g	1.38	0.72	27.94	1.8***	0.07***	0.02***		0.10	
		TKW	g	37.40	0.71	21.88	773.45***	18.16***	5.35***		29.48	
		Grain.Area		14.20	0.74	13.48	34.57***	1.14***	0.33***		1.46	
		Grain.Width		3.39	0.70	8.93	0.88***	0.02***	7.25e-03***		0.03	
		Grain.Length		5.78	0.83	4.78	0.8***	0.05***	7.43e-03***		0.04	
			NbGrain		136.33	0.70	13.92	17145.46**	808.25***	71.88		1883.59
	Tillers	GW	g	5.21	0.73	25.87	80.68***	1.47***	0.14		2.92	
		TKW	g	38.01	0.78	14.49	255.4***	15.39***	4.05***		14.08	
		Grain.Area	mm ²	13.66	0.81	8.35	12.08***	1.06***	0.21***		0.87	
		Grain.Width	mm ²	3.35	0.76	4.81	0.24***	0.02***	4.37e-03***		0.02	
		Grain.Length	mm ²	5.65	0.88	3.66	0.48***	0.06***	5.33e-03***		0.03	

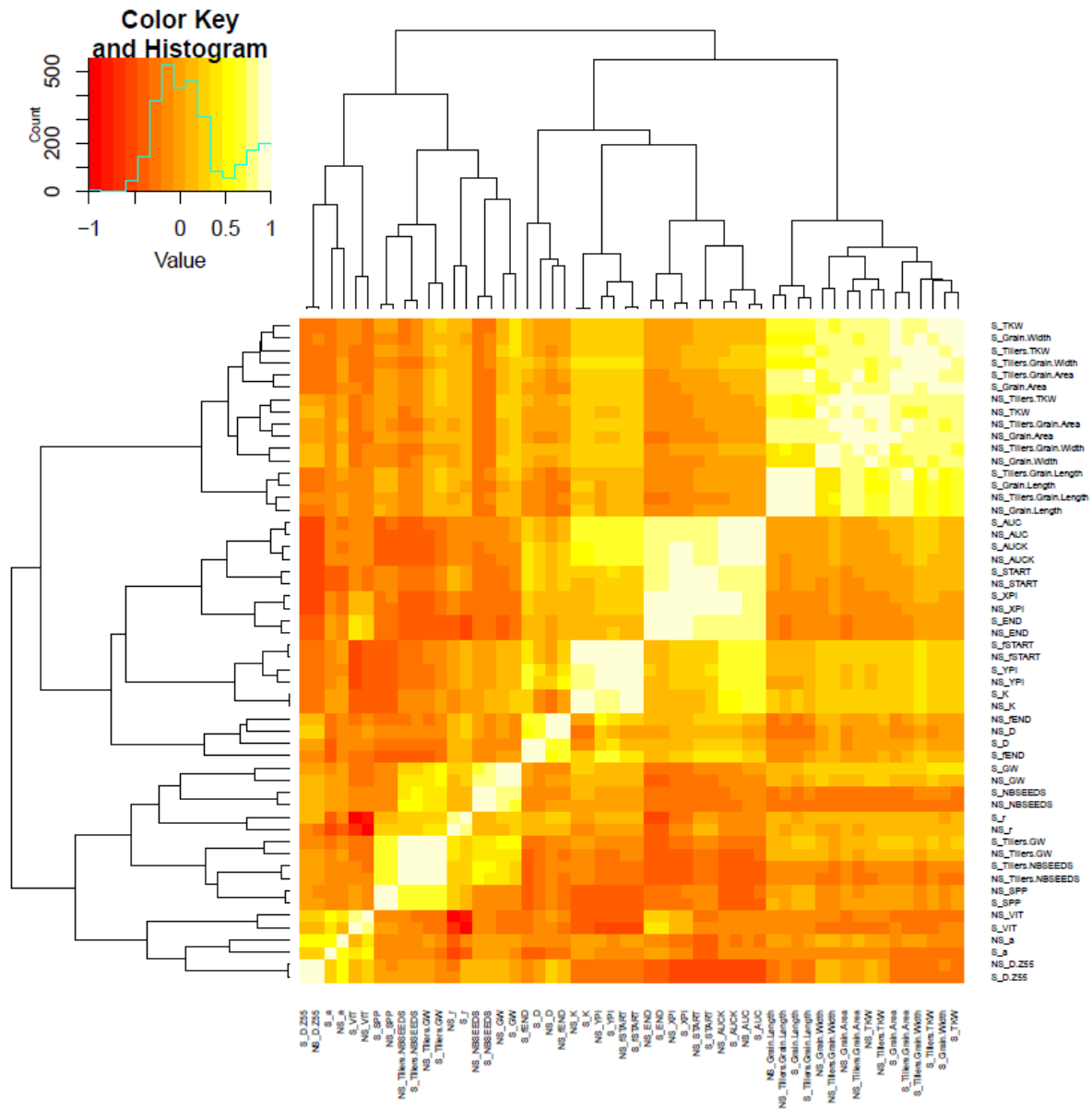
Wald and LTR test: ***, P-values <0.001; **, P-values <0.01; *, P-values <0.05 and '.', P-values <0.1

• **Supplementary data IV.4 : Genotypes performances for TKW in no stress (NS) and stress (S) treatments (with standard deviation), plus the coefficient of Genotype by treatment interaction using Model 9**

GENOTYPE	NS		S		Genotype by treatment
	gBLUP	gBLUP_sd	gBLUP	gBLUP_sd	
ACCOR	33.28	1.95	27.40	1.95	1.25
ACCROC	40.23	1.98	33.12	2.09	1.44
ACIENDA	35.91	1.91	28.93	1.91	0.97
ACOUSTIC	40.62	1.89	29.20	1.89	-0.94
ADAGIO	44.76	1.87	32.01	1.86	-1.16
ADHOC	42.95	1.97	34.03	1.97	0.77
AEROBIC	44.91	1.94	36.92	1.94	1.54
ALCAZAR	46.71	1.92	32.81	2.00	-1.56
ALCHEMY	46.29	2.02	36.76	1.92	0.85
ALDRIC	40.65	1.90	31.67	1.90	0.44
ALIGATOR	40.18	1.97	29.80	1.97	-0.41
ALIXAN	37.05	1.95	28.04	1.95	-0.03
ALIZEO	45.69	1.92	31.99	1.92	-1.58
ALLEZ_Y	46.14	1.95	34.29	1.95	-0.48
ALTAMIRA	45.32	1.91	38.16	1.92	2.06
ALTIGO	48.92	1.97	39.71	1.97	1.37
ALTRIA	44.99	1.96	34.59	1.96	0.19
AMADOR	34.94	1.92	28.94	1.92	1.39
AMBELLO	50.16	1.86	37.06	1.86	-0.67
AMBITION	44.82	2.08	33.38	1.97	-0.42
AMUNDSEN	42.30	1.87	28.40	1.87	-2.12
ANDALOU	45.08	1.93	34.24	1.93	-0.04
ANTONIUS	41.78	1.87	32.17	1.87	0.23
APACHE	47.64	1.78	38.74	1.72	1.37
APRILIO	44.31	1.97	35.27	2.08	0.87
ARAMIS	45.94	1.91	35.97	1.91	0.56
AREZZO	47.16	1.90	36.88	1.90	0.54
ARISTOTE	41.45	1.94	32.57	1.94	0.60
ARKEOS	31.88	1.94	26.39	1.94	1.29
ARLEQUIN	38.05	1.89	26.69	1.89	-1.24
ARTDECO	40.77	1.99	37.07	1.99	3.43
ASCOTT	42.77	1.90	29.98	1.90	-1.44
ATHLON	46.34	1.93	33.69	1.92	-0.91
ATTITUDE	42.33	1.96	29.50	2.06	-1.52
ATTLASS	46.69	1.93	30.70	1.93	-2.74
AUBUSSON	42.35	1.89	32.62	1.89	0.23
AUTAN	52.82	1.97	41.59	1.97	0.72
AZZERTI	41.58	1.98	32.51	1.98	0.51
BAGOU	39.17	2.03	28.86	1.93	-0.50
BAROK	42.63	1.97	30.66	2.07	-1.00
BASTIDE	42.60	2.43	31.90	2.43	-0.28
BERGAMO	44.41	1.93	36.13	1.93	1.31
BERMUDE	41.43	1.07	32.38	1.08	0.50
BIANCOR	42.36	1.94	32.73	2.03	0.29
BOISSEAU	46.55	1.96	35.67	2.06	0.12
BOLOGNA	37.74	1.97	30.54	1.97	1.07
BOREGAR	45.80	1.08	35.91	1.08	0.59
BUENNO	46.76	1.97	41.06	1.97	3.07
CAMP_REMY	44.08	1.98	33.65	1.98	0.06
CAPHORN	43.16	1.74	34.07	1.75	0.70
CATALAN	44.86	1.97	33.94	1.98	-0.11
CCB_INGENIO	50.09	2.00	39.10	2.11	0.51
CELESTIN	47.95	1.94	37.97	2.04	0.81
CELLULE	46.35	1.97	37.09	1.97	1.01
CENTENAIRE	47.04	1.90	35.91	1.90	0.05
CEZANNE	57.15	1.97	46.04	1.97	1.34
CH_NARA	42.99	2.02	32.09	2.02	-0.34
CHARGER	43.67	1.74	31.38	1.75	-1.04
CHEVALIER	47.29	1.93	32.35	2.04	-2.07
CLAIRE	42.17	1.83	29.88	1.83	-1.23
COMPIL	40.50	1.88	31.27	1.87	0.28
CORDIALE	43.35	1.92	30.89	1.92	-1.18
CORVUS	40.65	1.91	27.09	2.00	-2.15
COURTOT	41.82	2.14	36.78	2.02	2.81
CROISADE	44.03	1.93	36.37	1.93	1.62
CROUSTY	46.61	1.97	36.64	1.97	0.64
DIALOG	45.80	1.97	31.98	2.08	-1.63
DINOSOR	46.25	1.91	34.78	1.91	-0.25
EPHOROS	45.37	1.92	34.41	1.92	-0.08

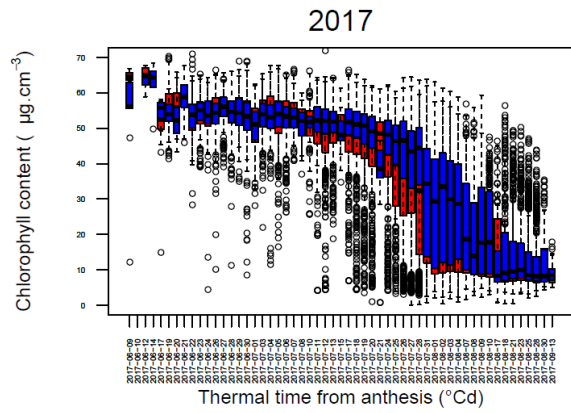
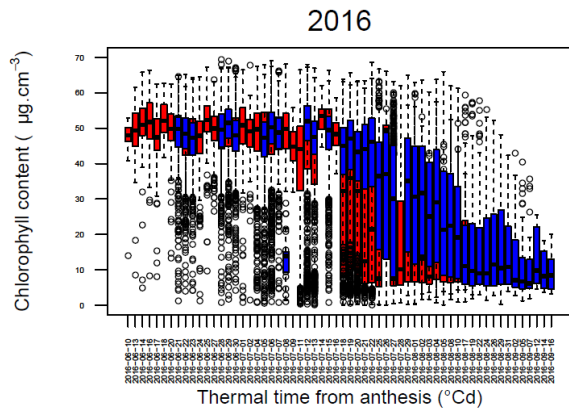
EPIDOC	43.87	2.08	30.98	1.97	-1.35
ESKET	40.41	1.88	30.88	1.97	0.10
EUCLIDE	45.29	1.92	36.62	2.01	1.21
EUREKA	52.00	1.99	43.03	2.00	1.89
EXELCIOR	45.87	2.48	38.32	2.60	1.91
EXOTIC	46.71	1.92	38.31	1.92	1.54
EXPERT	56.57	1.91	40.43	1.91	-1.57
FAIRPLAY	42.56	1.93	30.29	1.93	-1.17
FARINELLI	49.15	1.91	40.14	1.91	1.51
FIGARO	50.68	1.98	44.44	1.98	3.27
FLAMENKO	42.20	1.96	33.77	1.96	0.95
FLAUBERT	38.33	1.94	27.35	1.94	-0.98
FLUOR	39.22	1.94	30.29	1.94	0.29
FOLKLOR	46.02	1.99	33.85	2.10	-0.67
FORBLANC	49.01	1.97	40.41	1.89	1.72
FRELON	42.23	1.82	33.03	1.82	0.52
GALACTIC	44.97	2.12	32.87	1.93	-0.77
GALLANT	38.15	1.91	28.21	2.01	-0.42
GALOPAIN	40.74	1.87	32.11	1.87	0.65
GALPINO	42.59	1.87	34.26	1.87	1.05
GALVANO	41.77	1.91	29.92	1.91	-1.04
GARANTUS	38.34	1.87	28.11	1.88	-0.56
GARCIA	45.04	1.93	34.04	1.93	-0.14
GLADIATOR	42.96	1.92	29.43	1.92	-1.83
GLASGOW	37.96	1.97	27.97	1.97	-0.47
GONCOURT	42.27	2.00	32.42	2.00	0.16
GRAINDOR	49.38	1.88	41.29	1.88	2.06
HAUSSMANN	42.33	1.81	32.06	1.81	-0.07
HEREWARD	44.13	1.91	30.32	1.91	-1.84
HISSEO	44.92	1.91	36.88	2.00	1.52
ILLICO	48.13	1.86	41.49	1.86	2.71
INNOV	49.06	1.91	39.80	2.00	1.35
INSTINCT	39.39	1.96	25.35	2.20	-2.57
INTERET	47.19	2.05	36.52	1.95	0.32
IRIDIUM	48.40	1.92	36.02	1.92	-0.49
ISENGRAIN	45.12	1.78	36.89	1.77	1.43
JAGUAR	36.32	1.96	30.22	2.06	1.51
JB_DIEGO	43.19	1.90	29.80	1.90	-1.72
KALAHARI	46.39	1.95	36.72	1.95	0.78
KALANGO	41.61	1.99	32.99	1.99	0.76
KALYSTAR	43.78	1.91	31.09	1.91	-1.25
KARILLON	44.47	1.95	32.52	1.95	-0.75
KOMETUS	45.05	2.08	28.69	1.97	-3.16
KORELI	48.66	2.01	36.94	2.13	-0.09
KWS_PODIUM	45.32	1.96	34.92	1.96	0.23
LASER	40.39	1.98	30.38	1.98	-0.17
LAURIER	55.06	1.97	40.62	1.96	-0.80
LEAR	43.19	1.92	33.27	1.92	0.23
LIMES	53.15	1.82	39.19	1.82	-0.77
LORD	43.96	1.96	37.69	1.96	2.39
MANAGER	44.08	1.95	32.16	1.95	-0.78
MARCELIN	47.76	1.88	36.28	1.89	-0.06
MARKSMAN	40.34	1.82	28.25	1.82	-1.35
MERCATO	48.16	1.84	36.71	1.84	0.00
MH_09.17	41.65	1.93	30.66	1.93	-0.57
MIROIR	46.40	1.93	38.26	2.02	1.65
MOSKITO	46.58	1.90	34.93	1.90	-0.31
MUSIK	41.36	1.99	33.74	1.99	1.30
NIRVANA	45.97	1.83	34.36	1.83	-0.36
NUCLEO	43.62	1.90	32.36	1.90	-0.47
OAKLEY	41.42	2.02	26.36	1.92	-2.89
ODYSSEE	44.72	1.99	34.81	1.98	0.43
ORATORIO	39.47	2.00	28.11	1.91	-1.05
ORCAS	46.69	1.92	36.04	1.93	0.27
OREGRAIN	44.59	1.85	35.33	1.93	0.78
ORVANTIS	50.27	1.85	37.50	1.85	-0.47
OXEBO	48.51	1.97	35.74	1.97	-0.70
PAJERO	41.23	1.97	32.80	1.97	0.83
PAKITO	49.02	1.89	38.91	1.89	0.87
PALEDOR	38.57	2.06	29.46	1.96	0.10
PAPAGENO	41.54	1.99	30.26	1.99	-0.74
PARADOR	46.08	2.04	33.03	1.94	-1.17
PEPIDOR	43.39	1.91	30.95	1.91	-1.16
PERFECTOR	46.57	2.55	35.32	2.43	-0.09
PHARE	38.69	1.84	26.82	1.92	-1.44
PIERROT	42.48	1.92	31.04	2.02	-0.71
PIRENEO	42.53	1.87	32.75	1.87	0.23

PLAYER	43.41	2.08	29.22	2.08	-2.15
PR22R20	41.77	2.07	29.64	1.97	-1.19
PREMIO	49.37	1.09	37.65	1.07	0.00
PREVERT	45.72	1.85	36.02	1.84	0.68
PRIMO	34.11	1.92	28.16	2.02	1.32
RACINE	40.18	1.90	27.29	1.90	-1.82
RAZZANO	51.16	1.91	36.69	1.90	-1.31
RENAN	48.62	2.01	36.14	2.01	-0.52
RIMBAUD	43.95	1.90	36.32	1.90	1.62
ROBIGUS	39.14	1.83	27.06	1.83	-1.50
RONCARD	45.62	1.99	30.75	1.99	-2.25
ROSARIO	44.24	1.80	33.28	1.80	-0.22
ROYSSAC	45.55	1.89	32.99	1.89	-0.95
RUBISKO	52.48	1.87	39.66	1.88	-0.22
RUSTIC	46.24	1.94	33.28	1.94	-1.09
SAINT_EX	39.07	1.99	34.45	1.90	2.70
SAMURAI	50.17	1.91	38.21	1.91	-0.02
SANKARA	40.28	2.01	30.13	1.91	-0.27
SCENARIO	40.76	1.97	28.81	1.88	-1.22
SCOR	46.99	1.89	35.46	1.88	-0.19
SEBASTO	46.07	1.89	32.22	1.89	-1.61
SELEKT	41.47	1.82	29.06	1.82	-1.39
SEYRAC	43.84	1.93	34.82	1.93	0.82
SIRTAKI	44.66	1.91	32.33	1.91	-0.94
SOBBEL	55.67	1.93	46.77	1.93	2.40
SOGOOD	49.13	1.81	34.71	1.81	-1.54
SOISSONS	42.81	1.86	35.48	1.86	1.64
SOKAL	43.30	1.95	28.86	1.95	-2.30
SOLEDAD	45.01	1.87	36.30	1.87	1.15
SOLEHIO	50.04	1.95	42.54	1.95	2.47
SOLLARIO	42.87	1.96	32.86	1.96	0.14
SOLUTION	38.35	1.87	26.58	1.87	-1.42
SOPHYTRA	52.66	1.82	39.49	1.82	-0.39
SORRIAL	46.45	1.88	37.48	1.81	1.19
SPONSOR	35.51	2.04	24.27	1.94	-1.49
SWEET	54.57	1.90	41.11	1.90	-0.31
SWINGGY	47.75	1.92	32.54	1.92	-2.17
SY_ALTEO	55.25	1.87	43.47	1.87	0.73
SY_EPSON	42.69	1.91	32.10	1.91	-0.21
SY_MATTIS	43.73	1.92	34.75	1.92	0.84
SY_MOISSON	43.25	1.93	33.78	2.02	0.50
SY_TOLBIAC	47.20	1.90	36.97	1.90	0.57
TAPIDOR	37.55	1.89	27.51	1.90	-0.55
TIMING	44.79	2.00	37.11	2.11	1.70
TOISONDOR	41.41	1.97	31.91	1.97	0.25
TRAPEZ	53.16	1.94	37.83	1.94	-1.55
TREMIE	40.39	1.78	30.98	1.78	0.16
TULIP	44.26	1.97	31.14	1.97	-1.43
USKI	47.98	1.95	37.58	1.95	0.57
VALODOR	48.14	1.81	37.90	1.81	0.68
VISCOUNT	43.18	1.92	33.94	1.92	0.61
WARRIOR	36.43	1.94	25.39	2.04	-1.26
XI19	37.24	1.99	29.36	2.11	0.62



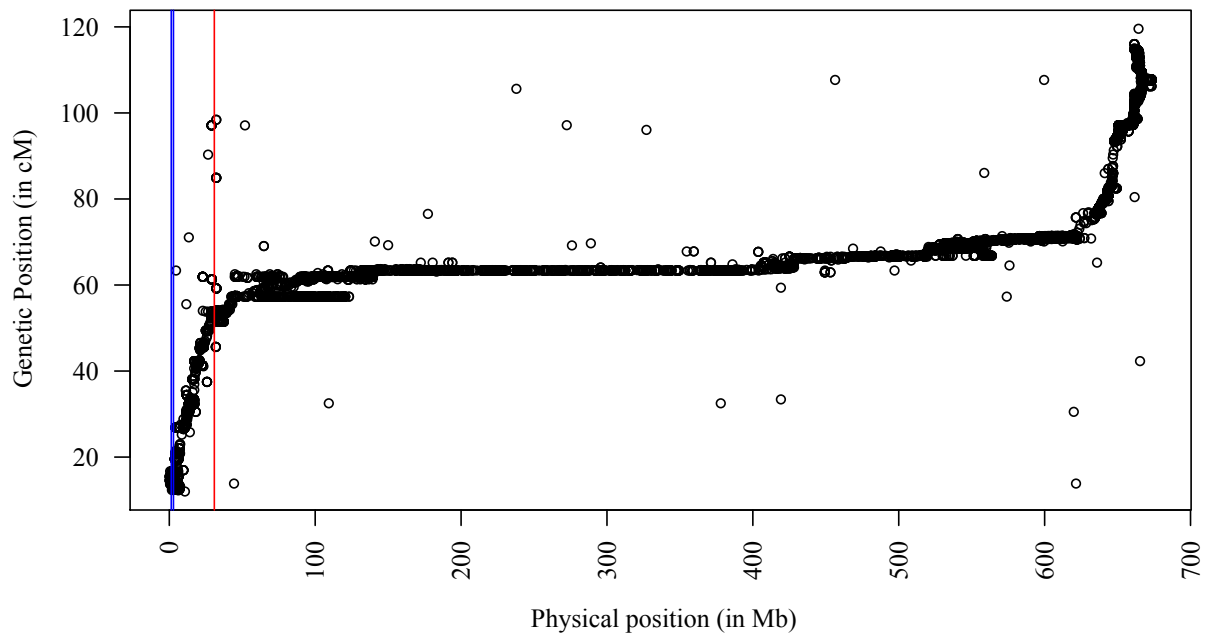
- Supplementary data IV. 5 : Correlation matrix for gBLUP of all the traits scored in the two treatments.**

Colors are a function of Pearson's pairwise correlations. Treatment and traits are separated by an underscore. Rows and columns are ranked as a function of their Euclidian distance.



- **Supplementary data IV.6 : Chlorophyll content as a function of Julian days for 2016 and 2017 experiments.**

Raw data were used to build the boxplot. The boxplot in blue shows no-stress treatment and the boxplot in red shows stress treatment



- Supplementary data IV.6 : The Marey map of chromosome 4B and the physical position of Rht-B1 (red) and the QTL link with TKW heat stress tolerance (blue).**
 Physical positions are the position on Chinese Spring refSeq1.0 (IWGSC et al 2018) and genetic position come from a consensus map between the Renan/Recital genetic map (Rimbert et al. 2018) and the Limagrain genetic map (unpublished data).

Annexes V

• **Annexe V.1 : Table descriptive des 212 covariables environnementales (EC) calculées sur les 17 environnements**

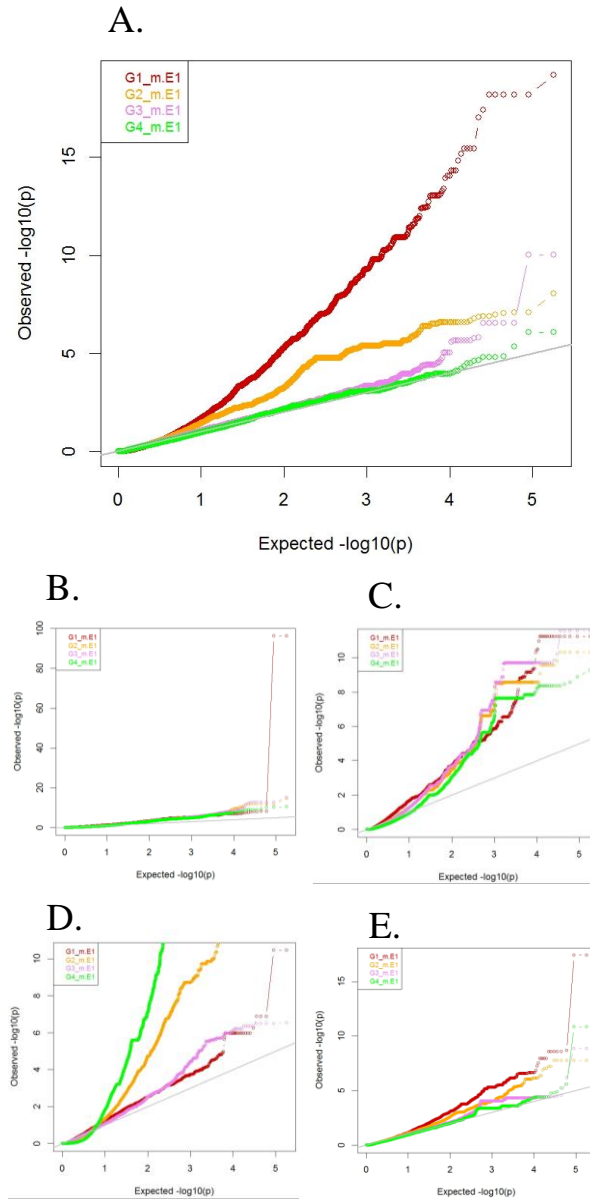
Les EC ont été regroupées en 17 clusters et l'EC la plus représentative de chaque cluster est notée en gras.

EC	Type	Début	Fin	Variable	Critère Limite	Unitée	Composante	SD	Clustering
EC024	Deficit hydrique	Z00	Z33	Dh	s	mm	SA	0.07	1
EC026	Offre climatique	Z00	Z33	Rg	s	calorie	SA	5512.24	1
EC027	Offre climatique	Z00	Z33	Rg	m	calorie	SA	42.70	1
EC028	Offre climatique	Z00	Z33	RgsurT	s	calorie/°C	SA	1010.47	1
EC029	Offre climatique	Z00	Z33	RgsurT	m	calorie/°C	SA	7.40	1
EC030	Exces eau	Z00	Z33	Rui.	s	mm	SA	24.24	1
EC049	Offre climatique	Z11	Z21	Rg	s	calorie	SA	2625.42	1
EC050	Offre climatique	Z11	Z21	Rg	m	calorie	SA	37.67	1
EC052	Offre climatique	Z11	Z21	RgsurT	s	calorie/°C	SA	729.87	1
EC053	Offre climatique	Z11	Z21	RgsurT	m	calorie/°C	SA	10.81	1 T
EC073	Offre climatique	Z21	Z33	Rg	m	calorie	SA	59.60	1
EC087	Deficit hydrique	Z21	Z33	Dh	s	mm	SA	0.07	1
EC103	Offre climatique	Z33	Z55	RgsurT	m	calorie/°C	SA & GPS	3.69	1
EC105	Gel	Z33	Z55	TMini	n	<0 °C	SA & GPS	1.76	1
EC107	Gel	Z33	Z55	TMini	s	<0 °C	SA & GPS	1.80	1
EC178	Gel	Z55 -350°j	Z55 +350°j	TMini	n	<0 °C	SA & GPS	0.34	1
EC179	Gel	Z55 -350°j	Z55 +350°j	TMini	s	<0 °C	SA & GPS	0.22	1
EC211	Maladie	All	All	Dégat gèle	m	note 1-9		0.001	1
EC005	Offre climatique	Z00	Z11	Rg	m	calorie	SA	35.83	2
EC034	Offre climatique	Z00	Z33	TMoy	s	°C	SA	199.76	2
EC035	Offre climatique	Z00	Z33	TMoy	m	°C	SA	1.62	2
EC057	Offre climatique	Z11	Z21	TMoy	s	°C	SA	183.99	2
EC058	Offre climatique	Z11	Z21	TMoy	m	°C	SA	3.15	2 T
EC077	Exces eau	Z21	Z33	Rui.	s	mm	SA	8.23	2
EC110	Deficit hydrique	Z33	Z55	Dh	s	mm	SA & GPS	18.45	2
EC032	Gel	Z00	Z33	TMini	n	<0 °C	SA	18.09	3
EC033	Gel	Z00	Z33	TMini	s	<0 °C	SA	145.50	3
EC071	Deficit hydrique	Z21	Z33	P-etp	s	<0 mm	SA	21.57	3
EC122	Fortes temperatures	M	Z55	TMaxi	s	>25 °C	SA & GPS	7.17	3
EC128	Fortes temperatures	M -5j	M +5j	TMaxi	s	>25 °C	SA & GPS	8.08	3
EC168	Deficit hydrique	Z55	Z92	Dh	s	mm	TKW	59.69	3
EC201	Deficit hydrique	Z75	Z92	Dh	s	mm	TKW	34.76	3 T
EC016	Deficit hydrique	Z00	Z11	Dh	s	mm	SA	0.00	4
EC096	Stress Hydrique	Z33	Z55	Fa.H.LAI	m		SA & GPS	0.10	4
EC109	Offre climatique	Z33	Z55	TMoy	m	°C	SA & GPS	0.94	4
EC112	Stress Hydrique	Z33	Z55	Fa.H.Bio	m		SA & GPS	0.10	4
EC123	Fortes temperatures	M	Z55	TMaxi	n	>25 °C	SA & GPS	1.69	4
EC129	Fortes temperatures	M -5j	M +5j	TMaxi	n	>25 °C	SA & GPS	2.03	4
EC181	Stress Azote	Z75	Z92	Fa.A.Bio	n	<0.333	TKW	1.21	4 T
EC194	Stress Hydrique	Z75	Z92	Fa.H.LAI	n	<0.666	TKW	1.21	4
EC195	Stress Hydrique	Z75	Z92	Fa.H.LAI	n	<0.333	TKW	1.21	4
EC197	Stress Hydrique	Z75	Z92	Fa.H.Bio	n	<0.666	TKW	1.21	4
EC198	Stress Hydrique	Z75	Z92	Fa.H.Bio	n	<0.333	TKW	1.21	4
EC200	Stress Azote	Z75	Z92	Fa.A.LAI	n	<0.666	TKW	1.21	4
EC202	Stress Azote	Z75	Z92	Fa.A.LAI	n	<0.333	TKW	1.21	4
EC204	Stress Azote	Z75	Z92	Fa.A.Bio	n	<0.666	TKW	1.21	4
EC002	Deficit hydrique	Z00	Z11	P-etp	s	<0 mm	SA	6.65	5
EC008	Exces eau	Z00	Z11	Rui.	s	mm	SA	18.95	5
EC025	Deficit hydrique	Z00	Z33	P-etp	s	<0 mm	SA	30.10	5
EC063	Deficit hydrique	Z11	Z21	Dh	s	mm	SA	0.00	5
EC097	Deficit hydrique	Z33	Z55	P-etp	s	<0 mm	SA & GPS	27.20	5
EC098	Offre climatique	Z33	Z55	Rg	s	calorie	SA & GPS	2647.14	5
EC099	Offre climatique	Z33	Z55	Rg	m	calorie	SA & GPS	51.61	5
EC131	Deficit hydrique	Z55	Z75	P-etp	s	<0 mm	TKW	21.88	5
EC133	Offre climatique	Z55	Z75	Rg	m	calorie	TKW	65.78	5
EC137	Fortes temperatures	Z55	Z75	TMaxi	s	>25 °C	TKW	33.15	5 T
EC138	Fortes temperatures	Z55	Z75	TMaxi	n	>25 °C	TKW	5.40	5
EC141	Offre climatique	Z55	Z75	TMoy	s	°C	TKW	47.98	5
EC142	Offre climatique	Z55	Z75	TMoy	m	°C	TKW	1.45	5
EC155	Deficit hydrique	Z55	Z92	P-etp	s	<0 mm	TKW	37.29	5
EC156	Offre climatique	Z55	Z92	Rg	s	calorie	TKW	3165.33	5
EC157	Offre climatique	Z55	Z92	Rg	m	calorie	TKW	56.71	5

EC165	Offre climatique	Z55	Z92	TMoy	s		°C	TKW	84.17	5	
EC166	Offre climatique	Z55	Z92	TMoy	m		°C	TKW	1.73	5	
EC182	Deficit hydrique	Z75	Z92	P-etp	s	<0	mm	TKW	20.05	5	
EC183	Offre climatique	Z75	Z92	Rg	s		calorie	TKW	1299.89	5	
EC184	Offre climatique	Z75	Z92	Rg	m		calorie	TKW	53.42	5	
EC188	Fortes temperatures	Z75	Z92	TMaxi	s	>25	°C	TKW	50.91	5	
EC189	Fortes temperatures	Z75	Z92	TMaxi	n	>25	°C	TKW	3.66	5	
EC148	Stress Azote	Z55	Z75	Fa.A.LAI	m			TKW	0.14	6	
EC149	Deficit hydrique	Z55	Z75	Dh	s		mm	TKW	38.22	6	
EC152	Stress Azote	Z55	Z75	Fa.A.Bio	m			TKW	0.14	6	
EC172	Stress Azote	Z55	Z92	Fa.A.LAI	m			TKW	0.14	6	
EC175	Stress Azote	Z55	Z92	Fa.A.Bio	m			TKW	0.14	6	
EC180	Azote	Z65 -1j	Z65	INN	v			TKW	0.29	6	
EC199	Stress Azote	Z75	Z92	Fa.A.LAI	m			TKW	0.15	6	
EC203	Stress Azote	Z75	Z92	Fa.A.Bio	m			TKW	0.15	6	T
EC205	Azote	Z92 -1j	Z92	INN	v			TKW	0.29	6	
EC037	Stress Hydrique	Z00	Z33	Fa.H.Bio	m			SA	0.00	7	
EC045	Stress Hydrique	Z00	Z33	Fa.H.LAI	m			SA	0.00	7	
EC074	Stress Hydrique	Z21	Z33	Fa.H.LAI	m			SA	0.00	7	
EC084	Stress Hydrique	Z21	Z33	Fa.H.Bio	m			SA	0.00	7	T
EC154	Stress Hydrique	Z55	Z92	Fa.H.LAI	m			TKW	0.24	8	
EC169	Stress Hydrique	Z55	Z92	Fa.H.Bio	m			TKW	0.24	8	
EC187	Stress Hydrique	Z75	Z92	Fa.H.LAI	m			TKW	0.34	8	
EC196	Stress Hydrique	Z75	Z92	Fa.H.Bio	m			TKW	0.34	8	T
EC013	Offre climatique	Z00	Z11	TMoy	m		°C	SA	3.76	9	
EC070	Stress Azote	Z21	Z33	Fa.A.Bio	n	<0.333		SA	21.58	9	
EC072	Offre climatique	Z21	Z33	Rg	s		calorie	SA	4735.37	9	
EC075	Offre climatique	Z21	Z33	RgsurT	s		calorie/°C	SA	817.30	9	
EC078	Gel	Z21	Z33	TMini	n	<0	°C	SA	16.17	9	
EC079	Gel	Z21	Z33	TMini	s	<0	°C	SA	77.20	9	
EC082	Stress Hydrique	Z21	Z33	Fa.H.LAI	n	<0.666		SA	21.58	9	
EC083	Stress Hydrique	Z21	Z33	Fa.H.LAI	n	<0.333		SA	21.58	9	
EC085	Stress Hydrique	Z21	Z33	Fa.H.Bio	n	<0.666		SA	21.58	9	
EC086	Stress Hydrique	Z21	Z33	Fa.H.Bio	n	<0.333		SA	21.58	9	
EC089	Stress Azote	Z21	Z33	Fa.A.LAI	n	<0.666		SA	21.58	9	
EC090	Stress Azote	Z21	Z33	Fa.A.LAI	n	<0.333		SA	21.58	9	
EC092	Stress Azote	Z21	Z33	Fa.A.Bio	n	<0.666		SA	21.58	9	T
EC134	Stress Hydrique	Z55	Z75	Fa.H.LAI	m			TKW	0.27	9	
EC145	Stress Hydrique	Z55	Z75	Fa.H.Bio	m			TKW	0.27	9	
EC185	Offre climatique	Z75	Z92	RgsurT	s		calorie/°C	TKW	69.75	9	
EC031	Stress Hydrique	Z00	Z33	Fa.H.LAI	n	<0.666		SA	11.34	10	
EC036	Stress Hydrique	Z00	Z33	Fa.H.LAI	n	<0.333		SA	11.34	10	
EC038	Stress Hydrique	Z00	Z33	Fa.H.Bio	n	<0.666		SA	11.34	10	
EC039	Stress Hydrique	Z00	Z33	Fa.H.Bio	n	<0.333		SA	11.34	10	
EC041	Stress Azote	Z00	Z33	Fa.A.LAI	n	<0.666		SA	11.34	10	
EC042	Stress Azote	Z00	Z33	Fa.A.LAI	n	<0.333		SA	11.34	10	
EC044	Stress Azote	Z00	Z33	Fa.A.Bio	n	<0.666		SA	11.34	10	
EC046	Stress Azote	Z00	Z33	Fa.A.Bio	n	<0.333		SA	11.34	10	T
EC102	Offre climatique	Z33	Z55	RgsurT	s		calorie/°C	SA & GPS	224.72	11	
EC106	Stress Hydrique	Z33	Z55	Fa.H.LAI	n	<0.666		SA & GPS	6.13	11	
EC108	Offre climatique	Z33	Z55	TMoy	s		°C	SA & GPS	75.24	11	
EC111	Stress Hydrique	Z33	Z55	Fa.H.LAI	n	<0.333		SA & GPS	6.13	11	
EC113	Stress Hydrique	Z33	Z55	Fa.H.Bio	n	<0.666		SA & GPS	6.13	11	
EC114	Stress Hydrique	Z33	Z55	Fa.H.Bio	n	<0.333		SA & GPS	6.13	11	
EC116	Stress Azote	Z33	Z55	Fa.A.LAI	n	<0.666		SA & GPS	6.13	11	
EC117	Stress Azote	Z33	Z55	Fa.A.LAI	n	<0.333		SA & GPS	6.13	11	
EC119	Stress Azote	Z33	Z55	Fa.A.Bio	n	<0.666		SA & GPS	6.13	11	T
EC120	Stress Azote	Z33	Z55	Fa.A.Bio	n	<0.333		SA & GPS	6.13	11	
EC130	Stress Azote	Z55	Z75	Fa.A.Bio	n	<0.333		TKW	2.22	12	T
EC143	Stress Hydrique	Z55	Z75	Fa.H.LAI	n	<0.666		TKW	2.22	12	
EC144	Stress Hydrique	Z55	Z75	Fa.H.LAI	n	<0.333		TKW	2.22	12	
EC146	Stress Hydrique	Z55	Z75	Fa.H.Bio	n	<0.666		TKW	2.22	12	
EC147	Stress Hydrique	Z55	Z75	Fa.H.Bio	n	<0.333		TKW	2.22	12	
EC150	Stress Azote	Z55	Z75	Fa.A.LAI	n	<0.666		TKW	2.22	12	
EC151	Stress Azote	Z55	Z75	Fa.A.LAI	n	<0.333		TKW	2.22	12	
EC153	Stress Azote	Z55	Z75	Fa.A.Bio	n	<0.666		TKW	2.22	12	
EC164	Stress Hydrique	Z55	Z92	Fa.H.LAI	n	<0.666		TKW	2.82	12	
EC167	Stress Hydrique	Z55	Z92	Fa.H.LAI	n	<0.333		TKW	2.82	12	
EC170	Stress Hydrique	Z55	Z92	Fa.H.Bio	n	<0.666		TKW	2.82	12	
EC171	Stress Hydrique	Z55	Z92	Fa.H.Bio	n	<0.333		TKW	2.82	12	

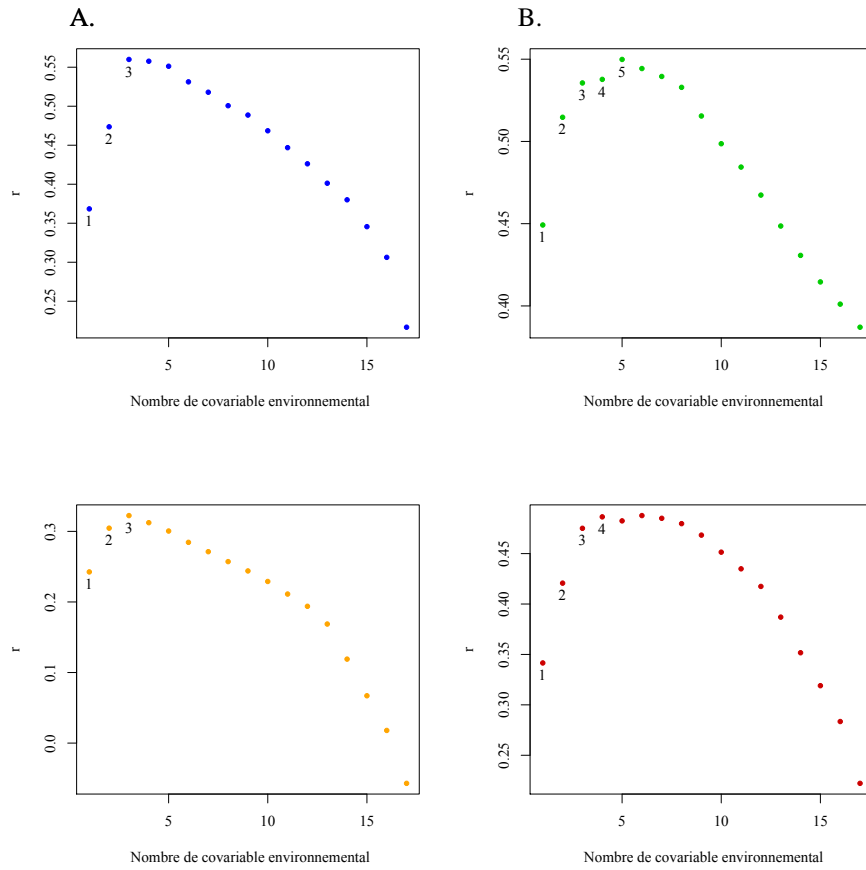
EC173	Stress Azote	Z55	Z92	Fa.A.LAI	n	<0.666		TKW	2.82	12	
EC174	Stress Azote	Z55	Z92	Fa.A.LAI	n	<0.333		TKW	2.82	12	
EC176	Stress Azote	Z55	Z92	Fa.A.Bio	n	<0.666		TKW	2.82	12	
EC177	Stress Azote	Z55	Z92	Fa.A.Bio	n	<0.333		TKW	2.82	12	
EC001	Stress Azote	Z00	Z11	Fa.A.Bio	n	<0.333		SA	10.35	13	T
EC004	Offre climatique	Z00	Z11	Rg	s		calorie	SA	923.72	13	
EC006	Offre climatique	Z00	Z11	RgsurT	s		calorie/°C	SA	332.55	13	
EC009	Gel	Z00	Z11	TMini	n	<0	°C	SA	10.01	13	
EC010	Gel	Z00	Z11	TMini	s	<0	°C	SA	73.98	13	
EC012	Stress Hydrique	Z00	Z11	Fa.H.LAI	n	<0.666		SA	10.35	13	
EC014	Stress Hydrique	Z00	Z11	Fa.H.LAI	n	<0.333		SA	10.35	13	
EC017	Stress Hydrique	Z00	Z11	Fa.H.Bio	n	<0.666		SA	10.35	13	
EC018	Stress Hydrique	Z00	Z11	Fa.H.Bio	n	<0.333		SA	10.35	13	
EC020	Stress Azote	Z00	Z11	Fa.A.LAI	n	<0.666		SA	10.35	13	
EC021	Stress Azote	Z00	Z11	Fa.A.LAI	n	<0.333		SA	10.35	13	
EC023	Stress Azote	Z00	Z11	Fa.A.Bio	n	<0.666		SA	10.35	13	
EC192	Offre climatique	Z75	Z92	TMoy	s		°C	TKW	48.83	13	
EC193	Offre climatique	Z75	Z92	TMoy	m		°C	TKW	2.21	13	
EC047	Stress Azote	Z11	Z21	Fa.A.Bio	n	<0.333		SA	13.69	14	
EC048	Deficit hydrique	Z11	Z21	P-etp	s	<0	mm	SA	21.35	14	
EC055	Gel	Z11	Z21	TMini	n	<0	°C	SA	15.38	14	
EC056	Gel	Z11	Z21	TMini	s	<0	°C	SA	153.65	14	
EC059	Stress Hydrique	Z11	Z21	Fa.H.LAI	n	<0.666		SA	13.69	14	
EC060	Stress Hydrique	Z11	Z21	Fa.H.LAI	n	<0.333		SA	13.69	14	
EC062	Stress Hydrique	Z11	Z21	Fa.H.Bio	n	<0.666		SA	13.69	14	
EC064	Stress Hydrique	Z11	Z21	Fa.H.Bio	n	<0.333		SA	13.69	14	T
EC066	Stress Azote	Z11	Z21	Fa.A.LAI	n	<0.666		SA	13.69	14	
EC067	Stress Azote	Z11	Z21	Fa.A.LAI	n	<0.333		SA	13.69	14	
EC069	Stress Azote	Z11	Z21	Fa.A.Bio	n	<0.666		SA	13.69	14	
EC081	Offre climatique	Z21	Z33	TMoy	m		°C	SA	2.22	14	
EC007	Offre climatique	Z00	Z11	RgsurT	m		calorie/°C	SA	10.56	15	
EC040	Stress Azote	Z00	Z33	Fa.A.LAI	m			SA	0.05	15	
EC043	Stress Azote	Z00	Z33	Fa.A.Bio	m			SA	0.05	15	
EC065	Stress Azote	Z11	Z21	Fa.A.LAI	m			SA	0.05	15	
EC068	Stress Azote	Z11	Z21	Fa.A.Bio	m			SA	0.05	15	
EC080	Offre climatique	Z21	Z33	TMoy	s		°C	SA	86.77	15	
EC088	Stress Azote	Z21	Z33	Fa.A.LAI	m			SA	0.11	15	T
EC091	Stress Azote	Z21	Z33	Fa.A.Bio	m			SA	0.11	15	
EC093	Azote	Z32-lj	Z32	INN	v			SA	0.33	15	
EC115	Stress Azote	Z33	Z55	Fa.A.LAI	m			SA & GPS	0.14	15	
EC118	Stress Azote	Z33	Z55	Fa.A.Bio	m			SA & GPS	0.14	15	
EC054	Exces eau	Z11	Z21	Rui.	s		mm	SA	12.52	16	
EC076	Offre climatique	Z21	Z33	RgsurT	m		calorie/°C	SA	8.64	16	
EC104	Exces eau	Z33	Z55	Rui.	s		mm	SA & GPS	4.29	16	
EC132	Offre climatique	Z55	Z75	Rg	s		calorie	TKW	2271.24	16	
EC135	Offre climatique	Z55	Z75	RgsurT	s		calorie/°C	TKW	140.29	16	
EC136	Offre climatique	Z55	Z75	RgsurT	m		calorie/°C	TKW	3.51	16	
EC160	Offre climatique	Z55	Z92	RgsurT	s		calorie/°C	TKW	178.86	16	
EC161	Offre climatique	Z55	Z92	RgsurT	m		calorie/°C	TKW	2.81	16	T
EC186	Offre climatique	Z75	Z92	RgsurT	m		calorie/°C	TKW	3.06	16	
EC011	Offre climatique	Z00	Z11	TMoy	s		°C	SA	42.16	17	
EC094	Deficit rayonnement	Z33	M	Rg	s	<250	calorie	SA & GPS	376.94	17	
EC095	Deficit rayonnement	Z33	M	Rg	n	<250	calorie	SA & GPS	3.70	17	
EC100	Deficit rayonnement	Z33	Z55	Rg	s	<250	calorie	SA & GPS	429.50	17	
EC101	Deficit rayonnement	Z33	Z55	Rg	n	<250	calorie	SA & GPS	4.37	17	
EC121	Deficit rayonnement	M	Z55	Rg	s		calorie	SA & GPS	72.88	17	
EC124	Deficit rayonnement	M -100°j	Z55	Rg	s	<250	calorie	SA & GPS	134.33	17	
EC125	Deficit rayonnement	M -100°j	Z55	Rg	n	<250	calorie	SA & GPS	1.68	17	T
EC126	Deficit rayonnement	M -5j	M +5j	Rg	s	<250	calorie	SA & GPS	68.36	17	
EC127	Deficit rayonnement	M -5j	M +5j	Rg	n	<250	calorie	SA & GPS	0.87	17	
EC158	Deficit rayonnement	Z55	Z92	Rg	s	<250	calorie	TKW	145.03	17	
EC159	Deficit rayonnement	Z55	Z92	Rg	n	<250	calorie	TKW	2.23	17	
EC003	Stress Hydrique	Z00	Z11	Fa.H.LAI	m			SA	0	NA	
EC015	Stress Hydrique	Z00	Z11	Fa.H.Bio	m			SA	0	NA	
EC019	Stress Azote	Z00	Z11	Fa.A.LAI	m			SA	0	NA	
EC022	Stress Azote	Z00	Z11	Fa.A.Bio	m			SA	0	NA	
EC051	Stress Hydrique	Z11	Z21	Fa.H.LAI	m			SA	0	NA	
EC061	Stress Hydrique	Z11	Z21	Fa.H.Bio	m			SA	0	NA	
EC139	Gel	Z55	Z75	TMini	n	<0	°C	TKW	0	NA	
EC140	Gel	Z55	Z75	TMini	s	<0	°C	TKW	0	NA	

EC162	Gel	Z55	Z92	TMini	n	<0	°C	TKW	0	NA
EC163	Gel	Z55	Z92	TMini	s	<0	°C	TKW	0	NA
EC190	Gel	Z75	Z92	TMini	n	<0	°C	TKW	0	NA
EC191	Gel	Z75	Z92	TMini	s	<0	°C	TKW	0	NA
EC206	Maladie	All	All	Rouille Brune	m		note 1-9		0	NA
EC207	Maladie	All	All	Fusariose de l'épis	m		note 1-9	TKW & GPS	0	NA
EC208	Maladie	All	All	Verse	m		note 1-9		0	NA
EC209	Maladie	All	All	Oidium	m		note 1-9		0	NA
EC210	Maladie	All	All	Septoriose	m		note 1-9		0	NA
EC212	Maladie	All	All	Rouille jaune	m		note 1-9		0	NA



• **Annexe V.2 : QQ-plot issus des modèles GWAS multi-environnement pour le rendement et les composantes**

Les couleurs rouge, orange, violet et vert représentent la distribution des P-values calculées sur l'ensemble des SNP avec les modèles FA1, FA2, FA3 et FA4, respectivement. La figure A, B, C, D et E sont les QQ-plot pour le rendement en grains, la date de floraison, le nombre d'épis par m², le nombre de grains par épis et le poids de mille grains, respectivement.



• **Annexe V.3 : Corrélation entre la matrice de VarCov Σ_G pour différents traits et la matrice de distance Σ_{EG} entre environnements en fonction du nombre de covariables environnementales utilisées pour construire la matrice Σ_{EG}**

A. Date de floraison ; B. Nombre d'épis par m² ; C. nombre de grain par épis ; D. Poids de mille grains.

- **Annexe V.4 : Description des SNP sélectionnés dans le modèle multi-environnement multi-locus (Mod [4]) et relation avec les covariables environnementales.**

Les positions des QTL ont été obtenues en ajoutant ou soustrayant l'étendue du LD dans la zone ; le LOG et les effets sont issus du Modèle [4], le R² pour chaque SNP et le R² de la régression entre la matrice d'effet du SNP et la matrice des valeurs $Y_{ij}-\mu_j$; Pour chaque covariable sélectionnée, le coefficient de régression de Spearman est représenté avec son niveau de significativité, ce coefficient est issu de la régression des effets des SNP dans chaque environnement et de la valeur des covariables dans chaque environnement.

- **A. Date d'épiaison (D.Z55)**

<i>SNP</i>	Chr	Pos	Min	Max	MAF	LOG	Min	Max	Moy.	R²	Nenv	EC058	EC084	EC092
<i>AX-89774410</i>	2B	53,152,857	51,907,756	54,397,958	0.09	8.77	-6.06	-1.92	-3.88	0.07	17	0.66**	-0.64**	0.01
<i>AX-89417224</i>	2B	793,034,008	792,262,877	793,805,139	0.06	3.27	-1.86	1.13	0.02	0.06	1	0	0.21	0.25
<i>PPD</i>	2D	33,955,686	31,523,326	36,388,046	0.47	< 2.2e-16	-10.29	-2.74	-6.36	0.40	17	0.66**	-0.61**	0
<i>AX-89411693</i>	2D	507,831,658	507,497,670	508,165,646	0.05	5.83	-4.11	-1.05	-2.12	0.11	14	0.5*	-0.29	0.21
<i>AX-89458321</i>	5D	545,979,738	544,427,324	547,532,152	0.06	3.21	-3.07	-1.16	-1.84	0.04	14	0.5*	-0.36	0.24
<i>AX-86167911</i>	6B	688,302,220	687,835,327	688,769,113	0.05	9.51	-4.22	1.07	-1.84	0.13	11	0.65**	-0.64**	0.07
<i>AX-89729836</i>	7A	10,860,633	10,313,195	11,408,071	0.24	3.56	-0.47	0.84	0.23	0.02	0	-0.24	0.11	0.12
<i>AX-89463792</i>	7A	575,270,076	569,620,744	580,919,408	0.10	11.22	-1.10	1.73	0.34	0.01	2	-0.11	0.53*	0.02
<i>AX-89379451</i>	7B	547,280,837	532,582,403	561,979,271	0.11	7.34	0.65	2.88	1.42	0.05	11	-0.58*	0.25	-0.18

- **B. Nombre d'épis par m²**

<i>SNP</i>	Chr	Pos	Min	Max	MAF	LOG	Min	Max	Moy.	R²	Nenv	EC064	EC084	EC203	EC181	EC130
<i>AX-89448426</i>	1B	453877190	447740210	460014170	0.20	4.97	-37.37	5.74	-15.07	0.02	13	0.29	0.44.	0.08	0.14	0.32
<i>AX-89313298</i>	2A	507316373	490440038	524192708	0.38	2.87	-45.08	-1.60	-21.36	0.00	15	-0.65**	0.18	-0.04	0.25	0.51*
<i>AX-89475760</i>	2B	658209385	653601192	662817578	0.21	4.02	-12.40	47.11	21.82	0.11	14	0.06	-0.66**	-0.42.	-0.7**	-0.06
<i>PPD</i>	2D	33955686	31523326	36388046	0.47	10.76	-38.64	45.22	-2.00	0.01	13	0.01	0.16	-0.06	0.44.	-0.61*
<i>AX-89562015</i>	3B	391644588	347310734	435978442	0.31	4.54	-38.67	20.22	-0.67	0.07	12	0.55*	-0.29	-0.38	-0.45.	-0.4
<i>AX-89518325</i>	4A	29158632	27965184	30352080	0.20	6.62	-4.52	40.00	20.40	0.02	15	0.38	0.16	-0.07	-0.03	-0.2
<i>AX-89401469</i>	4A	567209135	551836163	582582107	0.16	3.47	-34.69	25.19	-2.23	0.15	14	0.54*	-0.16	-0.1	-0.54*	-0.17
<i>AX-89650790</i>	4A	661997409	659864088	664130730	0.42	3.49	-33.22	21.65	-7.61	0.03	14	0.12	-0.16	-0.23	-0.37	0.23
<i>AX-89730025</i>	4B	643018792	642397398	643640186	0.21	2.97	-17.44	40.65	17.84	0.07	14	0.37	-0.02	-0.3	-0.37	0.02
<i>AX-89686631</i>	4D	8596114	8596114	8596114	0.50	4.11	-14.04	16.24	3.13	0.00	10	-0.15	0.36	-0.27	0.33	0.32
<i>AX-89406945</i>	4D	460831218	451533964	470128472	0.23	7.50	-11.08	46.31	13.74	0.00	14	-0.44.	-0.01	0.21	0.2	0.1
<i>AX-89446584</i>	5B	580852037	578955714	582748360	0.18	4.15	-28.30	28.96	-3.62	0.05	9	0.09	0.03	-0.16	0.13	-0.33
<i>AX-89349453</i>	6B	643431025	637525064	649336986	0.07	3.00	-60.68	5.29	-23.41	0.02	13	-0.17	0.23	0.19	0	0.76***
<i>AX-89564204</i>	6B	713500860	713103787	713897933	0.22	4.98	-48.77	18.32	-8.95	0.01	13	0.06	0.48.	-0.18	0.18	0.32
<i>AX-89681284</i>	7B	444476128	434473489	454478767	0.18	3.68	-46.12	10.64	-9.79	0.00	10	-0.07	-0.62**	0.26	-0.29	-0.4
<i>AX-89751245</i>	7D	12616216	11069072	14163360	0.13	5.20	-30.23	35.46	5.28	0.07	16	0.55*	-0.33	-0.22	-0.4	-0.73**

• C. Nombre de grains par épis

SNP	Chr	Pos	Min	Max	MAF	LOG	Min	Max	Moy.	R ²	Nenv	EC001	EC181	EC201
AX-89614433	1A	534545508	533387013	535704003	0.16	3.77	-4.00	0.21	-1.48	0.03	6	0.55*	0.38	0.33
AX-89629904	1B	658410658	657817008	659004308	0.35	4.17	-2.42	1.65	0.03	0.01	2	-0.14	0.38	-0.34
AX-89608609	2A	76519199	73030596	80007802	0.41	4.82	-1.11	2.69	0.61	0.01	4	0.13	-0.06	-0.08
AX-89675223	2A	711969402	711383409	712555395	0.06	5.52	-5.36	0.82	-2.29	0.24	8	-0.16	-0.22	-0.15
PPD	2D	33955686	31523326	36388046	0.47	<15	-3.35	2.04	-0.69	0.02	8	-0.16	-0.52*	-0.37
AX-89320644	2D	102663338	98615904	106710772	0.34	3.82	-1.31	2.31	0.82	0.01	4	-0.24	-0.74***	-0.08
AX-89716856	3B	729167792	726807044	731528540	0.07	3.02	-0.90	4.73	1.22	0.04	5	0.06	0.29	0.02
AX-89691021	4B	4529364	3301953	5756775	0.40	5.07	-0.69	2.33	0.55	0.00	4	0.56*	0.5*	-0.17
AX-89469167	5A	451718534	450499779	452937289	0.13	3.62	-0.99	5.30	1.05	0.21	2	0.13	0.17	-0.18
AX-89663549	5A	680886582	680068203	681704961	0.16	12.31	-3.46	3.11	-1.27	0.01	10	0.01	-0.03	-0.37
AX-89427621	5B	64640833	52984970	76296696	0.17	6.96	-0.85	2.70	0.54	0.01	4	0.09	0.43	0.35
AX-89336004	6B	12120300	11691057	12549543	0.08	5.10	-2.19	5.47	0.80	0.14	8	-0.42	0.18	-0.26
AX-89364136	6B	483502666	479308515	487696817	0.06	4.64	-5.15	2.84	-0.95	0.18	7	0.21	0.19	0.23
AX-89706745	6D	6443318	5805840	7080796	0.29	7.75	-1.70	13.42	5.32	0.01	15	-0.41	0.17	0.06
AX-89772931	6D	6553418	5917449	7189387	0.29	6.23	-33.19	9.81	-9.41	0.01	15	0.11	-0.01	-0.37
AX-89389438	6D	6567768	5931993	7203543	0.30	5.02	-7.05	28.55	3.60	0.02	9	-0.05	-0.24	0.41
AX-86170398	7A	120574833	116988231	124161435	0.50	7.30	-1.64	2.53	-0.07	0.00	2	0.32	-0.34	-0.22
AX-89376373	7B	358752754	344149799	373355709	0.05	9.99	-2.23	5.99	0.41	0.41	5	0.04	0.17	-0.16
AX-89739920	7B	625904472	623597694	628211250	0.31	6.01	-1.57	3.45	0.38	0.07	3	-0.46	0.27	0

• **C. Poids de mille grains**

<i>SNP</i>	Chr	Pos	Min	Max	MAF	LOG	Min	Max	Moy.	R²	Nenv	EC001	EC181	EC201	EC137
<i>AX-89626732</i>	1B	4965255	4794627	5135883	0.05	4.01	-3.19	1.60	-0.73	0.02	6	-0.23	0.1	0.14	0.41
<i>AX-89588588</i>	1B	477405267	467625838	487184696	0.17	3.24	-2.85	-0.16	-1.54	0.01	10	0.18	0.05	-0.07	0.68**
<i>AX-89507056</i>	1D	488311472	482951513	493671431	0.48	5.19	-1.03	0.91	0.05	0.00	0	0.24	-0.28	-0.25	-0.31
<i>AX-89635032</i>	2A	712351340	711770570	712932110	0.41	8.65	-2.06	-0.28	-1.21	0.05	11	0.54*	0.04	-0.18	0.13
<i>PPD</i>	2D	33955686	31523326	36388046	0.47	< 2.2e-16	-0.27	3.73	2.18	0.14	13	0.21	-0.15	-0.16	0.48.
<i>AX-89481481</i>	2D	617409406	600889976	633928836	0.34	2.98	-1.05	0.72	0.13	0.00	0	0.18	0.14	0	0.38
<i>AX-89466001</i>	4B	594851314	590757259	598945369	0.46	5.51	-1.03	1.10	0.13	0.00	1	-0.32	0.21	-0.24	-0.07
<i>AX-89766169</i>	4D	17255339	17255339	17255339	0.31	7.31	-1.75	1.32	0.41	0.00	4	-0.56*	0.43.	-0.01	0.06
<i>AX-89474360</i>	5A	438471661	436780227	440163095	0.12	4.63	-2.01	0.65	-0.82	0.01	3	-0.41	-0.32	0.34	-0.39
<i>AX-89364136</i>	6B	483502666	479308515	487696817	0.06	9.19	-1.95	2.55	-0.36	0.00	3	0.25	0.26	0.08	0.14
<i>AX-89351462</i>	7A	675080484	673784162	676376806	0.31	3.48	-0.55	1.32	0.28	0.00	1	-0.25	0.05	0.11	-0.63**
<i>AX-89535064</i>	7A	696821354	696211545	697431163	0.06	5.53	0.88	4.95	3.57	0.06	16	0.28	-0.49*	-0.44.	-0.48.

ANALYSE GENETIQUE ET ECOPHYSIOLOGIQUE DE LA TOLERANCE A LA SECHERESSE ET AU STRESS THERMIQUE CHEZ LE BLE TENDRE (*T. AESTIVUM* L.)

RESUME :

Dans un contexte de changement climatique, la caractérisation des variétés de blé tendre en réponse à des événements de sécheresse et de stress thermique est un des défis de l'agriculture. Cette thèse, issue d'un partenariat public-privé entre Arvalis-Institut du Végétal, Biogemma et l'INRA (Institut National de la Recherche Agronomique), avait pour but de développer des connaissances et des outils nécessaires à l'identification de variétés tolérantes à la sécheresse et au stress thermique et à la création de variétés répondant à cette exigence. Pour ce faire, nous avons analysé un panel de 220 variétés commerciales, génotypées avec 280K SNP et testées dans 35 environnements variés (combinaison d'année, lieu et régime hydrique), plus une expérimentation en conditions contrôlées où un stress thermique a été appliqué pendant le remplissage du grain. La complexité de l'étude de la tolérance à la sécheresse nous a conduit à présenter cette thèse en séparant, dans un premier temps, l'étude des stress hydriques et thermiques, puis de prospecter une méthode d'analyse multi-stress. Nous avons montré que même si la sélection a amélioré la performance des variétés en condition hydrique optimale, le progrès génétique doit être accéléré et mieux réparti en fonction des différents types de stress. Nous proposons pour cela plusieurs déterminants génétiques qui pourraient permettre un gain dans des environnements stressants. Nos résultats et méthodes sont discutés au regard des besoins en préconisation et amélioration variétale. Des pistes de recherche complémentaires et des améliorations ont aussi été suggérées.

Mots-clés : Analyse multi-stress, Génétiques d'association, Progrès génétique, Stress hydrique et thermique

GENETIC AND ECOPHYSIOLOGICAL ANALYSES OF TOLERANCE TO DROUGHT AND HIGH TEMPERATURE IN BREAD WHEAT (*TRITICUM AESTIVUM* L.)

SUMMARY:

In a context of climate change, the characterization of wheat varieties in response to drought and heat stress events is one of the major challenges of agriculture. This PhD thesis, resulting from a private-public partnership between Arvalis 'Institut du Végétal', Biogemma and INRA ("Institut National de la Recherche Agronomique"), aimed at providing necessary knowledge and tools to identify drought or heat-tolerant varieties and breed for varieties that meet these requirements. Analyses were conducted using a panel of 220 commercial varieties, genotyped with 280K SNP and tested in 35 environments (combination of year, location and water regime) and an experiment under controlled conditions where heat stress was applied during grain filling. The complexity of the study of drought and heat tolerance led us to present this thesis by first separating hydric and thermal stresses, and then to explore a multi-stress analysis method. Even if breeding has improved the performance of varieties under optimal water conditions, we showed that genetic progress must be accelerated and better distributed according to different stress scenarios. We propose several genetic determinants that could allow genetic gain in stressful environments. Our results and methods are discussed in view of the needs for varietal recommendation and improvement. Additional research strategies and methods improvements were also suggested.

Keywords: Multi-stress analysis, Genome Wide Association Studies, Genetic progress, Drought stress.