REPUBLIQUE DU CAMEROUN
Paix – Travail – Patrie
****
UNIVERSITE DE YAOUNDE I
FACULTE DES SCIENCES
DEPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VEGETALES
****

REPUBLIC OF CAMEROUN
Peace – Work – Fatherland
****
UNIVERSITY OF YAOUNDE I
FACULTY OF SCIENCE
DEPARTMENT OF PLANT BIOLOGY
****

Genetic diversity, QTL mapping for essential yield components and aluminum tolerance in hexaploid wheat (Triticum aestivum L.) by Next-Generation Sequencing

THESIS submitted for the award of the Doctorate/PhD degree in Plant Biology
THESE Présentée en vue de l'obtention du diplôme de Doctorat/PhD en Biologie Végétale

Par : **TEKEU Honoré**
Master in Plant Biotechnology

Sous la direction de
DJOCGOUE Pierre François
Associate Professor

Année Académique : 2018

UNIVERSITE DE YAOUNDE I
UNIVERSITY OF YAOUNDE I

FACULTE DES SCIENCES
FACULTY OF SCIENCE

DEPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VEGETALES
DEPARTMENT OF PLANT BIOLOGY

ATTESTATION DE CORRECTION

****

Nous soussignés, membres du Jury de soutenance du mémoire de Doctorat/Ph.D en Biologie des Organismes Végétaux option Biotechnologies Végétales de l'étudiant **TEKEU Honoré**, Matricule **07S240**, soutenu publiquement le 07 Juin 2018 sur le sujet «**Genetic diversity, QTL mapping for essential yield components and aluminum tolerance in hexaploid wheat (*Triticum aestivum* L.) by Next-Generation Sequencing**» attestons que les corrections conformément aux remarques et recommandations du jury lors de la soutenance de la dite thèse de Doctorat/Ph.D ont été effectuées par le candidat.

En foi de quoi, la présente attestation lui est délivrée pour servir et valoir ce que de droit./-

Rapporteur

DJOCGOUE Pierre François
Maître de Conférences

Membres

BELL Joseph Martin
Professeur

MBOUOBDA Hermann Désiré
Maître de Conférences

NIEMENAK Nicolas
Maître de Conférences

NGONKEU MANGAPTCHE Eddy Léonard
Maître de Conférences

Président

YOUMBI Emmanuel
Professeur

UNIVERSITE DE YAOUNDE I FACULTE DES SCIENCES Division de la Programmation et du Suivi des Activités Académiques		The University of Yaoundé I Faculty of Science Division of Programming and follow-up of Academic Affaires

LISTE DES ENSEIGNANTS PERMANENTS	LIST OF PERMENENT TEACHING STAFF

ANNEE ACADEMIQUE 2017/2018
(Par Département et par Grade)

DATE D'ACTUALISATION : 31 Janvier 2018

ADMINISTRATION

DOYEN : AWONO ONANA Charles, Professeur
VICE-DOYEN / DPSAA : DONGO Etienne, Professeur
VICE-DOYEN / DSSE : OBEN Julius ENYONG, Professeur
VICE-DOYEN / DRC : MBAZE MEVA'A Luc Léonard, Maître de Conférences
Chef Division Affaires Académiques, Scolarité et Recherche : ABOSSOLO Monique, Chargée de Cours
Chef Division Administrative et Financière : NDOYE FOE Marie C.F., Maître de Conférences

1-Département de Biochimie (BC) (40)			
N°	Noms et Prénoms	Grade	Observations
1	BENG née NINTCHOM PENLAP V.	Professeur	En poste
2	FEKAM BOYOM Fabrice	Professeur	En poste
3	MBACHAM Wilfried	Professeur	Chef Dpt/FMSB
4	MOUNDIPA FEWOU Paul	Professeur	Chef de Département
5	OBEN Julius ENYONG	Professeur	Vice-Doyen DSSE/Coord. CRFD-STG
ICE	BIGOGA DIAGA Jude	Maître de Conférences	En poste
7	BOUDJEKO Thaddée	Maître de Conférences	En poste
8	FOKOU Elie	Maître de Conférences	En poste
9	KANSCI Germain	Maître de Conférences	En poste
10	MINKA Samuel	Maître de Conférences	En poste

11	NGONDI Judith Laure	Maître de Conférences	En poste
12	NGUEFACK Julienne	Maître de Conférences	En poste
13	NJAYOU Frédéric Nico	Maître de Conférences	En poste
14	WAKAM née NANA Louise	Maître de Conférences	En poste
15	BELINGA née NDOYE FOE Marie C.	Maître de Conférences	Chef DAF / FS
16	ACHU Merci BIH	Chargée de Cours	En poste
17	ATOGHO Barbara Mma	Chargée de Cours	En poste
18	DEMMANO Gustave	Chargé de Cours	En poste
19	DJOKAM TAMO Rosine	Chargé de Cours	En poste
20	DJUIDJE NGOUNOUE Marcelline	Chargé de Cours	En poste
21	DJUIKWO NKONGA Ruth Viviane	Chargée de Cours	En poste
22	EFFA ONOMO Pierre	Chargé de Cours	En poste
23	EVEHE BEBANDOUE Marie – Solange	Chargée de Cours	En disponibilité
24	EWANE Cécile Anne	Chargée de Cours	En poste
25	KOTUE TAPTUE Charles	Chargé de Cours	En poste
26	MBONG ANGIE MOUGANDE Mary	Chargée de Cours	En poste
27	MOFOR née TEUGWA Clautilde	Chargée de Cours	IA4/MINESUP
28	Palmer MASUMBE NETONGO	Chargé de Cours	En poste
29	TCHANA KOUATCHOUA Angèle	Chargée de Cours	En poste
30	AKINDEH MBUH NJI	Assistant	En poste
31	BEBEE FADIMATOU	Assistante	En poste
32	BEBOY EDJENGUELE Sara Nathalie	Assistante	En poste
33	DAKOLE DABOY Charles	Assistant	En poste
34	DONGMO LEKAGNE Joseph Blaise	Assistant	En poste
35	FONKOUA Martin	Assistant	En poste
36	LUNGA Paul KAILAH	Assistant	En poste
37	MANANGA Marlyse Joséphine	Assistante	En poste

38	MBOUCHE FANMOE Marcelline Joëlle	Assistante	En poste
39	PECHANGOU NSANGOU Sylvain	Assistant	En poste
40	TIENTCHEU DJOKAM Léopold	Assistant	En poste
Département de BIOLOGIE ET PHYSIOLOGIE ANIMALES (B.P.A.) (44)			
1	BILONG BILONG Charles Félix	Professeur	Chef de Département
2	DIMO Théophile	Professeur	En Poste
3	FOMENA Abraham	Professeur	En Poste
4	KAMTCHOUING Pierre	Professeur	en poste
5	MIMPFOUNDI REMY	Professeur	En poste
6	NJAMEN Dieudonné	Professeur	En poste
7	NJIOKOU Flobert	Professeur	En Poste
8	NOLA Moïse	Professeur	En poste
9	TAN Paul Vernyuy	Professeur	En poste
10	TCHUEM TCHUENTE Louis	Professeur	Coord. Progr. MINSANTE
11	AJEAGAH Gidéon AGHAINDOUM	Maître de Conférences	C.S. D.P.E.R/FS
12	DJIETO Lordon Champlain	Maître de Conférences	En poste
13	DZEUFIET DJOMENI Paul Désiré	Maître de Conférences	En poste
14	ESSOMBA née NTSAMA MBALLA	Maître de Conférences	VDRC/Chef Dépt FMSB
15	FOTO MENBOHAN Samuel	Maître de Conférences	En poste
16	KAMGANG René	Maître de Conférences	C.E. MINRESI
17	KEKEUNOU Sévilor	Maître de Conférences	En poste
18	MEGNEKOU Rosette	Maître de Conférences	En poste
19	MONY NTONE Ruth	Maître de Conférences	En poste
20	ZEBAZE TOGOUET Serge Hubert	Maître de Conférences	En poste
21	ALENE Désirée Chantal	Chargée de Cours	En poste
22	ATSAMO Albert Donatien	Chargé de Cours	En poste
23	BELLET EDIMO Oscar Roger	Chargé de Cours	En poste
24	BILANDA Danielle Claude	Chargée de Cours	En poste
25	DJIOGUE Séfirin	Chargé de Cours	En poste

26	GOUNOUE KAMKUMO Raceline	Chargée de Cours	En poste
27	JATSA MEGAPTCHE Hermine	Chargée de Cours	En poste
28	MAHOB Raymond Joseph	Chargé de Cours	En poste
29	MBENOUN MASSE Paul Serge	Chargé de Cours	En poste
30	MOUGANG NGAMENI Luciane	Chargée de Cours	En poste
31	LEKEUFACK FOLEFACK Guy Benoît	Chargé de Cours	En poste
32	NGOUATEU KENFACK Omer BEBE	Chargé de Cours	En poste
33	NGUEGUIM TSOFACK Florence	Chargée de Cours	En poste
34	NGUEMBOCK	Chargé de Cours	En poste
35	NJUA Clarisse YAFI	Chargé de Cours	En poste
36	TADU Zéphirin	Chargé de Cours	En poste
37	TOMBI Jeannette	Chargée de Cours	En poste
38	YEDE	Chargé de Cours	En poste
39	ETEME ENAMA Serge	Assistant	En poste
40	KANDEDA KAVAYE Antoine	Assistant	En poste
41	KOGA MANG'Dobara	Assistant	En poste
42	METCHI DONGFACK Mireille Flore	Assistante	En poste
43	NOAH E. O. Vivien	Assistant	En poste
44	MVEYO NDANKEU Yves Patrick	Assistant	En poste
3-DEPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VEGETALES (B. P. V.) (27)			
1	YOUMBI Emmanuel	Professeur	Chef de Département
2	AMBANG Zachée	Professeur	CD A/ASR-UYII
3	MOSSEBO Dominique Claude	Professeur	En poste
4	BELL Joseph Martin	Professeur	En poste
5	DJOCGOUE Pierre François	Maître de Conférences	En poste
6	KENGNE NOUMSI Ives Magloire	Maître de Conférences	En poste
7	MBOLO Marie.	Maître de Conférences	En poste
8	NDONGO BEKOLO	Maître de Conférences	En poste

9	ZAPFACK Louis	Maître de Conférences	En poste
10	ANGONI Hyacinthe	Chargé de Cours	En poste
11	BIYE Elvire Hortense	Maître de Conférences	En poste
12	MAHBOU SOMO TOUKAM Gabriel	Chargé de Cours	En poste
13	MALLA Armand William	Maître de Conférences	En poste
14	MBARGA BINDZI Marie Alain.	Chargé de Cours	Inspecteur académ. N°1 MINESUP
15	NGALLE Hermine BILLE	Chargé de Cours	Coordo. Programme MINFOF
16	NGODO MELINGUI Jean Baptiste	Chargé de Cours	En poste
17	NGONKEU MAGAPTCHE Eddy Léonard	Maître de Conférences	En poste
18	NGOUO Lucas Vincent	Chargé de Cours	En poste
19	NSOM ZAMO Annie Claude ép. Pial	Chargée de Cours	Expert national. /UNESCO
20	ONANA Jean Michel	Chargé de Cours	En poste
21	TONFACK Libert Brice	Chargé de Cours	En poste
22	TSOATA Esaïe	Chargé de Cours	En poste
23	DJEUANI Astride Carole	Assistante	En poste
24	GONMADGE Christelle	Assistante	En poste
25	MAFFO MAFFO Nicole Liliane	Assistante	En poste
26	NNANGA MEBENGA Ruth Laure	Assistante	En poste
27	NOUKEU KOUAKAM Armelle	Assistante	En poste
4-DEPARTEMENT DE CHIMIE INORGANIQUE (C.I.) (35)			
1	KETCHA MBADCAM Joseph	Professeur	Chef de Département
2	NDIFON Peter TEKE	Professeur	CT MINRESI
3	NGAMENI Emmanuel	Professeur	Doyen/ UDS
4	GHOGOMU Paul MINGO	Professeur	Directeur Cabinet PM
5	LAMINSI Samuel	Professeur	En poste
6	MELO née CHINJE Uphie F.	Professeur	Recteur Université de N'Gaoundéré
7	AGWARA ONDOH Moïse	Professeur	Vice-Recteur
8	NANSEU Charles Péguy	Professeur	En poste

9	NENWA Justin	Professeur	En poste
10	BABALE DJAM DOUDOU	Maître de Conférences	En poste
11	DJOUFAC WOUMFO Emmanuel	Maître de Conférences	En poste
12	ELIMBI Antoine	Maître de Conférences	En poste
13	KONG SAKEO	Maître de Conférences	En poste
14	NDIKONTAR Maurice KOR	Maître de Conférences	Vice-Doyen/UBa
15	NGOMO Horace MANGA	Maître de Conférences	Vice-recteur
16	YOUNANG Elie	Maître de Conférences	En poste
17	ACAYANKA Elie	Chargé de Cours	En poste
18	CHEUMANI YONA Arnaud	Chargé de Cours	En poste
19	EMADACK Alphonse	Chargé de Cours	En poste
20	GWET Simon – Pierre	Chargé de Cours	En poste
21	KAMGANG YOUBI Georges	Chargé de Cours	En poste
22	KEUMEGNE MBOUGUEM Jean Claude	Chargé de Cours	En poste
23	KENNE DEDZO Gustave	Chargé de Cours	. En poste
24	MBEY Jean Aimé	Chargé de Cours	En poste
25	NDI Julius NSAMI	Chargé de Cours	En poste
26	NDOSIRI Bridget NDOYE	Chargée de Cours	En poste
27	NJIOMOU Chantale épse DJANGANG	Chargé de Cours	En poste
28	NJOYA Dayirou	Chargé de Cours	En poste
29	NYAMEN Linda Dyorisse	Chargée de Cours	En poste
30	PABOUDAM GBAMBIE Awaou	Chargé de Cours	En poste
31	TCHAKOUTE KOUAMO Hervé	Chargé de Cours	En poste
32	BELIBI BELIBI Placide Désiré	Assistant	En poste
33	KOUOTOU DAOUDA	Assistant	En poste
34	MAKON Thomas Beauregar	Assistant	En poste
35	NCHIMI NONO Katia	Assistante	En poste
5-DEPARTEMENT DE CHIMIE ORGANIQUE (C.O.) (33)			
1	DONGO Etienne	Professeur	Vice-Doyen/ DPSAA

2	GHOGOMU TIH ROBERT RALPH	Professeur	Directeur I.B.A. Foumban
3	MBAFOR Joseph	Professeur	En poste
4	NGADJUI TCHALEU B.	Professeur	Chef de dépt FMBS
5	NGOUELA Silvère Augustin	Professeur	Chef de dépt. UDs
6	NKENGFACK Augustin Ephraïm	Professeur	Chef de Département
7	NYASSE Barthélemy	Professeur	Directeur
8	PEGNYEMB Dieudonné Emmanuel	Professeur	Directeur au MINESUP
9	WANDJI Jean	Professeur	En poste
10	Alex de Théodore ATCHADE	Maître de Conférences	CS Rectorat/CD D.E.P.E.
11	FOLEFOC Gabriel NGOSONG	Maître de Conférences	En poste
12	KEUMEDJIO Félix	Maître de Conférences	En poste
13	KOUAM Jacques	Maître de Conférences	En poste
14	MBAZOA née DJAMA Céline	Maître de Conférences	En poste
15	NOUNGOUE TCHAMO Diderot	Maître de Conférences	En poste
16	TCHOUANKEU Jean-Claude	Maître de Conférences	C.T. UYII, Vice-recteur
17	YANKEP Emmanuel	Maître de Conférences	En poste
18	TCHUENDEM KENMOGNE Marguerite	Maître de Conférences	En poste
19	TIH née NGO BILONG E. Anastasie	Maître de Conférences	En poste
20	AMBASSA Pantaleon	Chargé de Cours	En poste
21	EYONG Kenneth OBEN	Chargé de Cours	En poste
22	FOTSO WABO Ghislain	Chargé de Cours	En poste
23	KAMTO Eutrophe Ledoux	Chargé de Cours	En poste
24	MKOUNGA Pierre	Chargé de Cours	En poste
25	NGO MBING Joséphine	Chargée de Cours	En poste
26	NGONO BIKOBO Dominique Serge	Chargé de Cours	En poste
27	NOTE LOUGBOT Olivier	Chargé de Cours	En poste
28	OUAHOUO WACHE Blandine Marlyse	Chargée de Cours	En poste
29	TABOPDA KUATE Turibio	Chargé de Cours	En poste

30	TAGATSING FOTSING Maurice	Chargé de Cours	En poste
31	ZONDEGOUMBA Ernestine	Chargée de Cours	En poste
32	NGINTEDO Dominique	Assistant	En poste
33	NGOMO Orléans	Assistant	En poste
6-DEPARTEMENT D'INFORMATIQUE (IN) (28)			
1	TCHUENTE Maurice	Professeur	PCA UB
2	FOTSO Pauline Laure	Professeur	Député à l'Assemblée Nationale
3	FOUDA NDJODO Marcel	Professeur	IGA-MINESUP/Chef Dpt ENS
4	ATSA ETOUNDI Roger	Professeur	Chef de Département/CD MINFOPRA
5	NDOUNDAM René	Maître de Conférences	En poste
6	CHEDOM FOTSO Donatien	Chargé de Cours	En poste
7	DJAM Xaviera Youth KIMBI	Chargé de Cours	En poste
8	KOUOKAM KOUOKAM Etienne Appolin	Chargé de Cours	En poste
9	MELATAGIA YONTA Paulin	Chargé de Cours	En poste
10	MOTO MPONG Serge Alain	Chargé de Cours	En poste
11	TINDO Gilbert	Chargé de Cours	En poste
12	TSOPZE Norbert	Chargé de Cours	En poste
13	WAKU KOUAMOU Jules	Chargé de Cours	En poste
14	ABESSOLO ALO'O Gislain	Assistant	En poste
15	AMINOU Halilou	Assistant	En poste
16	BAYEM Jacques Narcisse	Assistant	En poste
17	DOMGA KOMGUEM Rodrigue	Assistant	En poste
18	EBELE Serge	Assistant	En poste
19	HAMZA Adamou	Assistant	En poste
20	KAMDEM KENGNE Christiane	Assistant	En poste
21	KAMGUEU Patrick Olivier	Assistant	En poste
22	JIOMEKONG AZANZI Fidel	Assistant	En poste
23	MAKEMBE S. Fritz Oswald	Assistant	En poste

24	MEYEMDOU Nadège Sylvianne	Assistante	En poste
25	MONTHE DJIADEU Valery Martial	Assistant	En poste
26	NKONDOCK MI BAHANACK Nicolas	Assistant	En poste
27	OLE OLE David Claude	Assistant	En poste
28	TAPAMO KENFACK Hyppolite	Assistant	En poste
7-DEPARTEMENT DE MATHEMATIQUES (MA) (31)			
1	BEKOLLE David	Professeur	Vice-Recteur UN
2	BITJONG NDOMBOL	Professeur	En poste
3	DOSSA COSSY Marcel	Professeur	En poste
4	AYISSI Raoul Domingo	Maître de Conférences	Chef de Département
5	EMVUDU WONO Yves S.	Maître de Conférences	Chef Division MINESUP
6	NKUIMI JUGNIA Célestin	Maître de Conférences	En poste
7	NOUNDJEU Pierre	Maître de Conférences	En poste
8	AGHOUKENG JIOFACK Jean Gérard	Chargé de Cours	En poste
9	BOGSO Antoine M	Chargé de Cours	En poste
10	CHENDJOU Gilbert	Chargé de Cours	En poste
11	DOUANLA YONTA Hermann	Chargé de Cours	En poste
12	FOMEKONG Christophe	Chargé de Cours	En poste
13	KIANPI Maurice	Chargé de Cours	En poste
14	KIKI Maxime Armand	Chargé de Cours	En poste
15	MBAKOP Guy Merlin	Chargé de Cours	En poste
16	MBANG Joseph	Chargé de Cours	En poste
17	MBEHOU Mohamed	Chargé de Cours	En poste
18	MBELE BEDIMA Martin	Chargé de Cours	En poste
19	MENGUE MENGUE David Joe	Chargé de Cours	En poste
20	NGUEFACK Bertrand	Chargé de Cours	En poste
21	NGUIMTSA Charles	Chargé de Cours	En poste
22	POLA DOUNDOU Emmanuel	Chargé de Cours	En poste
23	TAKAM SOH Patrice	Chargé de Cours	En poste

24	TCHANGANG Roger Duclos	Chargé de Cours	En poste
25	TCHOUNDJA Edgar Landry	Chargé de Cours	En poste
26	TIAYA TSAGUE N. Anne- Marie	Chargée de Cours	En poste
27	DJIADEU NGAHA Michel	Assistant	En poste
28	MBIAKOP Hilaire George	Assistant	En poste
29	NIMPA PEFOUKEU Romain	Assistant	En poste
30	TANG AHANDA Barnabé	Assistant	Chef Serv. MINPLAMAT
31	TETSADJIO TCHILEPECK Mesmin Erick	Assistant	En poste
8-DEPARTEMENT DE MICROBIOLOGIE (MB) (13)			
1	ETOA François-Xavier	Professeur	Recteur UD/Chef de Département
2	BOYOMO ONANA	Maître de Conférences	En poste
3	ESSIA NGANG Jean Justin	Maître de Conférences	Chef de Division recherche IMPM
4	NYEGUE Maximilienne Ascension	Maître de Conférences	En poste
5	NWAGA Dieudonné M.	Maître de Conférences	En poste
6	RIWOM Sara Honorine	Maître de Conférences	En poste
7	SADO KAMDEM Sylvain Leroy	Maître de Conférences	En poste
8	BODA Maurice	Chargé de Cours	En poste
9	ENO Anna Arey	Chargée de Cours	En poste
10	ESSONO OBOUGOU Germain Gabriel	Chargé de Cours	En poste
11	BOUGNOM Blaise Pascal	Chargé de Cours	En poste
12	NJIKI BIKOÏ Jacky	Assistante	En poste
13	TCHIKOUA Roger	Assistant	En poste
9-DEPARTEMENT DE PHYSIQUE (PH) (41)			
1	ESSIMBI ZOBO Bernard	Professeur	En poste
2	KOFANE Timoléon Crépin	Professeur	En poste
3	NJOMO Donatien	Professeur	En poste
4	TABOD Charles TABOD	Professeur	Doyen/UBa

5	WOAFO Paul	Professeur	En poste
6	NDJAKA Jean Marie Bienvenu	Professeur	Chef de Département
7	PEMHA Elkana	Professeur	En poste
8	TCHAWOUA Clément	Professeur	En poste
9	BIYA MOTTO Frédéric	Maître de Conférences	Dir. Gén. B. MEKIM
10	BEN- BOLIE Germain Hubert	Maître de Conférences	En poste
11	DJUIDJE KENMOE Gemaine épse ALOYEM KAZE	Maître de Conférences	En poste
12	EKOBENA FOUDA Henri Paul	Maître de Conférences	Chef Dépt UN
13	EYEBE FOUDA Jean Sire	Maître de Conférences	En poste
14	FEWO Serge Ibraïd	Maître de Conférences	En poste
15	MBANE BIOUELE	Maître de Conférences	En poste
16	NANA NBENDJO Blaise	Maître de Conférences	En poste
17	NJANDJOCK NOUCK Philippe	Maître de Conférences	Chef Serv. MINRESI
18	NOUAYOU Robert	Maître de Conférences	En poste
19	OUMAROU BOUBA	Maître de Conférences	En poste
20	SAIDOU	Maître de Conférences	En poste
21	SIEWE SIEWE Martin	Maître de Conférences	En poste
22	ZEKENG Serge Sylvain	Maître de Conférences	En poste
23	BODO Bernard	Chargé de Cours	En poste
24	ENYEGUE A NYAM Françoise épouse BELINGA	Chargé de Cours	En poste
25	EDONGUE HERVAIS	Chargé de Cours	En poste
26	FOUEDJIO David	Chargé de Cours	En Poste
27	HONA Jacques	Chargé de Cours	En poste
28	MBINACK Clément	Chargé de Cours	En Poste
29	MBONO SAMBA Yves Christian U.	Chargé de Cours	En poste
30	NDOP Joseph	Chargé de Cours	En poste
31	OBOUNOU Marcel	Chargé de Cours	Dir.acad/Univ.Int.Etat Cam-Congo
32	SIMO Elie	Chargé de Cours	En poste

33	TABI Conrad Bertrand	Chargé de Cours	En poste
34	TCHOFFO Fidèle	Chargé de Cours	En poste
35	VONDOU DERBETINI Appolinaire	Chargé de Cours	En Poste
36	WAKATA née BEYA Annie	Chargée de Cours	Chef Serv. MINESUP
37	WOULACHE Rosalie Laure	Chargée de Cours	En poste
38	ABDOURAHIMI	Assistant	En Poste
39	CHAMANI Roméo	Assistant	En Poste
40	MELI'I Jorelle Larissa	Assistante	En Poste
41	MVOGO Alain	Assistant	En Poste
10- DEPARTEMENT DE SCIENCES DE LA TERRE (S.T.) (43)			
1	NZENTI Jean-Paul	Professeur	En poste
2	BITOM Dieudonné Lucien	Professeur	Doyen / UN
3	NDJIGUI Paul-Désiré	Professeur	Chef de Département
4	FOUATEU Rose épouse YONGUE	Maître de Conférences	En poste
5	KAMGANG Pierre	Maître de Conférences	En poste
6	MEDJO EKO Robert	Maître de Conférences	Dir/IUT Bois Mbalmayo
7	MOUNDI Amidou	Maître de Conférences	CT/MINMIDT
8	NDAM NGOUPAYOU Jules-Remy	Maître de Conférences	En poste
9	NGOS III Simon	Maître de Conférences	D.A.A.C./UM
10	NJILAH Isaac KONFOR	Maître de Conférences	En poste
11	NKOUMBOU Charles	Maître de Conférences	En poste
12	TEMDJIM Robert	Maître de Conférences	En poste
13	YENE ATANGANA Joseph Q.	Maître de Conférences	Chef Cellule MINTP
14	ABOSSOLO née ANGUE Monique	Chargée de Cours	Chef DAASR
15	BEKOA Etienne	Chargé de Cours	Barage Memve'ele
16	BISSO Dieudonné	Chargé de Cours	Directeur Projet Barage Memve'ele
17	EKOMANE Emile	Chargé de Cours	Délégué Régional/MINMIDT
18	ESSONO Jean	Chargé de Cours	En poste
19	EYONG John TAKEM	Chargé de Cours	En poste

20	FUH Calistus Gentry	Chargé de Cours	En poste
21	GANNO Sylvestre	Chargé de Cours	En poste
22	GHOGOMU Richard TANWI	Chargé de Cours	Chef de Dépt IMIP/Maroua
23	LAMILEN BILLA Daniel	Chargé de Cours	En poste
24	MBIDA YEM	Chargé de Cours	CS/LABOGENIE
25	MINYEM Dieudonné	Chargé de Cours	Chef de Dépt IMIP/Maroua
26	MOUAFO Lucas	Chargé de Cours	En poste
27	NGO BELNOUN Rose Noël	Chargée de Cours	En poste
28	NGO BIDJECK Louise Marie	Chargée de Cours	En poste
29	NGUEUTCHOUA Gabriel	Chargé de Cours	CEA MINRESI
30	NJOM Bernard de Lattre	Chargé de Cours	En poste
31	NYECK Bruno	Chargé de Cours	En poste
32	ONANA Vincent	Chargé de Cours	En poste
33	TCHAKOUNTE Jacqueline ép. NUMBEM	Chargée de cours	CEA MINRESI
34	TCHOUANKOUE Jean-Pierre	Chargé de Cours	En poste
35	ZO'O ZAME Philémon	Chargé de Cours	S.G. MINTP
36	TEMGA Jean Pierre	Chargé de Cours	En poste
37	ANABA ONANA Achille Basile	Assistant	En Poste
38	BINELI BETSI Thierry Olivier	Assistant	En Poste
39	METANG Victor	Assistant	En poste
40	NOMO NEGUE Emmanuel	Assistant	En poste
41	SABABA Elisé	Assistante	En Poste
42	TCHAPTCHET TCHATO De Pesquidoux I	Assistant	En poste
43	TEHNA Nathanaël	Assistant	CS/ MINMIDT

Répartition chiffrée des enseignants permanents par Département

Département	Nombre d'enseignants				
	Pr	MC	CC	ASS	Total
BC	5 (1)	10 (4)	14 (9)	11 (5)	40 (19)
BPA	10 (0)	10 (3)	18 (7)	6 (1)	44 (11)
BPV	3 (0)	6 (1)	13 (3)	5 (5)	27 (9)
C.I.	9 (1)	7 (0)	15 (3)	4 (2)	35 (6)
C.O.	9 (0)	10 (3)	12 (3)	2 (0)	33 (6)
IN	4 (1)	1 (0)	8 (0)	17 (4)	30 (5)
MA	3 (0)	4 (0)	19 (1)	7(0)	33 (1)
MB	1 (0)	6 (2)	4 (1)	2 (0)	13 (3)
PH	8 (0)	14 (1)	15 (3)	4 (1)	41 (5)
ST	3 (0)	10 (1)	23 (4)	7 (0)	43 (5)
Total	55 (3)	78 (15)	141 (34)	65 (18)	339 (70)

Soit un total de 340 (70) dont :

- Professeurs 55 (3)
- Maîtres de Conférences 78 (15)
- Chargés de Cours 141 (34)
- Assistants 65 (18)

( ) = Nombre de femmes

Le Doyen de la Faculté des Sciences

DEDICATION

- To all my family in general,
- Particularly to the DJIELA Pierre family,
  - my Mother FOKOU Géneviève,
  - my late Mother MEBONG Marie,
  - my Uncle NGOUNOU Augustin,

ACKNOWLEDGEMENTS

The work presented in this thesis was supported by “the Department of Plant Biology (University of Yaounde I)”, “Institute of Integrative Biology and Systems (Laval University, Canada)”; “Agence Universitaire de la Francophonie (Ecole Doctoral BIOVEG-AGRO, AUF)”, “European Commission through PAFROID project”, “MIC-CERES project (Microbial ecocompatible strategies for improving wheat quality traits and rhizospheric soil sustainability)”, “Institute of Agricultural Research for Development (IRAD)”, and by “Plant Breeding Laboratory, Department of Genetics (University of Stellenbosch, South Africa)”.

- I am grateful to the Department of Plant Biology Staff, Faculty of Sciences, University of Yaoundé I, for their dedication in my training;

- Pr. YOUMBI Emmanuel, Professor, Head of Department of Plant Biology, for his availability and dedication in our training in the Department;

- Pr. DJOCGOUE Pierre François, Associate Professor, for having accepted to directed this thesis. Also, for the suggestions and orientations that he brought to these works. Thanks for your constant support/encouragements and concern towards the completion of this work. I benefited from your great experience as a researcher from my first steps in scientific research up to now;

- Pr. NGONKEU M. Eddy Leonard, Associate Professor, for having accepted to guide my first steps in scientific research. This work has benefited from his scientific rigor and critical thinking. He received the project to enhance wheat productivity in Cameroon and knew how to build team spirit in our research team during the phenotyping trials in fields;

- I am grateful to the Staff of the Department of Phytology, Institute of Integrative Biology and Systems (IBIS), Laval University (Canada) for their constant support/encouragements and concern towards the completion of this work;

- Pr. BELZILE François, Professor, Laval University, you welcomed me into your rich and dynamic lab and I have learned a lot from you in molecular genetics analysis; Thanks for the high-level scientific research, as well as socially and professionally you taught me, these experiences were very important for my career and allowed me to consolidate my knowledges and develop new ones;

- Dr. MARTINE Jean, for her availability at the lab at any times, she supported me and taught me a lot about the use of NGS tools through genetic, bioinformatic and genomic analysis techniques.

- I express my deepest gratitude to all the members of François Belzile Laboratory, for the notions of genetic analysis in Next-Generation Sequencing technologies and for the long discussions time that have enriched my stay at the IBIS. I think of: Amina Abed, Sébastien Belanger, Sidiki Malle, Vincent-thomas Boucher St-Amour, Marc-André Lemay, Boudhrioua Chiheb, Manel Fallah,

Aurélie Tardivel, Davoud Torkamaneh, Martin Lacroix, Suzanne Marchand, Patricio Esteves, Magdalena Mastropierro, Waldiodio Seck ;

- I specially thank Brian Boyle and all the team of genomic analysis and sequencing platform for their support in sequencing and libraries preparing.

- My sincere thanks to Jérôme Laroche, Stéphane Larose and all the team of bioinformatics and genomic analysis who brought me a lot of expertise in computer programming.

- I express my sincere thanks to Dr. Ryan Peter and Dr. Leon Kochian for their suggestions and orientations on the chapter of wheat selection for tolerance to aluminum toxicity.

- I am grateful to the "Agence Universitaire de la Francophonie (College Doctoral BIOVEG-AGRO, AUF)" for many regional groupings of the PhD students which sharpened my curiosity and strengthened my abilities in scientific research;

- I especially thank all the teachers of the BIOVEG-AGRO Doctoral College, who were able to transmit us and train us in high-level research. I particularly thank Prof. Serge HAMON, Prof. DOUADI Khelifi, Prof. Sy Mame-Oureye;

- I would like to thank all my friends of the College for their frank collaboration.

- I am grateful to the Department of Genetics, University of Stellenbosch (South Africa) for their dedication in my training;

- Dr. BOTES Willem, Stellenbosch University, for the time you spent with me during my doctoral internship mobility in your lab; I would like to tell you how much I have been faithful to your teachings and I express my deepest gratitude for having guided and supported me.

- Mrs Aletta Ellis, Mrs Lezaan Springfield and Corrine, Lab managers, for their technical assistance in microsatellites markers analysis and PCR amplifications. They also taught me a lot about conducting a field experiment and I benefited from their multiple advice and suggestions;

- Pr. Aletta Van der Merwe, and Dr. Simo Maduna, Stellenbosch University, I have learned a lot from you in methods for populations genetics analyses;

- My deep gratitude goes to the members of the Plant Breeding Laboratory, for all the basic agronomic and technical knowledges in varietal selection of wheat that we learned. I cited: Cebekhulu Nondumido, Wandile Ngcamphalala, Jabolanie Mthembu, Swmeintjie, Irma Van Zyl, Andre;

- At the Department of Plant biology, University of Yaoundé, I express my gratitude to:

- Pr. BELL Joseph Martin, Professor, for having sharpened my curiosity on the basics of genetics; I would also like to thank him for the help he has given to this work;

- Pr. NIEMENAK Nicolas, Associate Professor, for the experiences in scientific communication that he transmitted to us during the BIOSCIENCES workshops and the 2nd grouping of the Doctoral College BIOVEG-AGRO in Yaoundé;

-Pr. MALLA Armand William, Associate Professor, for his technical orientations and discussions to improve the present document;

- Dr. TONFACK Brice Libert, for his scientific advice and discussions that have strengthened my abilities, since we were together in South Africa up to now in the Department;

- Pr. BOYOMO Onana, Associate Professor, for his assistance during my first steps in Master degree and for scientific discussions about molecular work done in this document;

- I expressed my immense gratitude to people at the Institute of Agricultural Research for Development (IRAD), for their technical assistance during this work:

- Dr. VENASIUS Lendzemo, head of IRAD North West, for his great assistance during socio-economic works and collection of local wheat cultivars form farmers in Cameroon;

- Mr. Tume George Tardzenyuy and Aga Valentine Enwe for their technical assistance for examining the perceptions of farmers in the Regions of North-west and North about local wheat varieties and production constraint;

- The late head of IRAD_Adamawa, for his help during the PRA with wheat farmers in Wassande, Adamawa;

- Dr. TANDZI Liliane, head of cereal IRAD_Nkolbisson, for his assistance to conserve seeds;

- I thank all of the members of our research team in Cameroon, especially TSIMI Patrick, NGO Adrienne, Mam Charly, TCHIECHOUA Yves, Thomas MBALLA, PETMI Carine, MONGOUE Aminatou, Ariane DAMDJO, Arlette FOKO, MANET Linda, TSETCHOUA Rosine, for their multiform support;

- I express my immense gratitude to all my brothers and sisters, nephews and nieces, for their availability, moral and warm support from my first days in school up to now; particularly to Dr. NGOMSI Augustin, DONKENG Marie, SIELINOU David, LATTA Joseph, DJIOKO Jean Paul, DOMMO Oscarine and NGUIEKO Linda.

- I express a deepest gratitude to Michèle TCHUINGUEM, for her constant love, frank friendship and tireless moral support;

- I thank my friends, especially MANGA Jude, MELACHIO Tresor, GUY LANDRY, KAMGA Raoul, NONGNI Gael, NKWAMENI Achille, for their multiform support;

- My deep acknowledgements go to the members of the jury, who accepted to examine this work. Thank you for your suggestions and orientations to improve the quality of this work.

TABLE OF CONTENTS

LISTE DES ENSEIGNANTS PERMANENTS i
LIST OF PERMENENT TEACHING STAFF i
DEDICATION xv
ACKNOWLEDGEMENTS xvi
TABLE OF CONTENTS xix
LIST OF FIGURES xxii
LIST OF TABLES xxiv
LIST OF ABBREVIATIONS xxv
ABSTRACT xxvi
RÉSUMÉ xxviii
CHAPTER I: GENERALITIES 1
I.1. Introduction 2
Context and justification 2
I.2. Literature review 7
I.2.1. Wheat 7
I.2.1.1. Domestication of wheat 7
I.2.1.2. Origin and evolution of wheat 7
I.2.1.3. Genome complexity in wheat 8
I.2.1.4. Evolution of genome sequencing in wheat 10
I.2.2. Importance and production of wheat 11
I.2.3. Wheat production in Cameroon 12
I.2.4. Genetic diversity in hexaploid wheat 16
I.2.5. Technologies of genomic sequencing 17
I.2.6. Bioinformatics Algorithms and Pipelines for DNA Sequence Analysis 19
I.2.7. Genetic mapping 19
I.2.7.1. Bi-parental (two-parent) mapping 20
I.2.7.2. Association Mapping 22
I.2.8. Genetic mapping for grain size 28
I.2.9. Impact of acid soils on wheat 30
I.2.9.1. Characteristics of acid soils 30
I.2.9.2. Impact of aluminum toxicity 31
I.2.9.3. Identification of functional genes for aluminum tolerance 32
I.2.9.4. Development of molecular markers for assisted selection 33
I.2.9.5. QTL mapping and inheritance of Al tolerance 34

I.2.9.6. Association Mapping Analysis for Al tolerance 35
CHAPTER II: MATERIALS AND METHODS 36
II.1. Genetic diversity of Cameroonian bread wheat (*Triticum aestivum* L.) cultivars revealed by microsatellite markers 37
II.1.1. Plant material 37
II.1.2. Genomic DNA extraction 37
II.1.3. Microsatellite markers and PCR amplification 38
II.1.4. Data analysis 39
II.2. Population structure in a global accession of hexaploid wheat (*Triticum aestivum* L.) breeding lines using Genotyping by Sequencing approach 39
II.2.1. Plant material 39
II.2.2. Agro-morphological characterization 46
II.2.2.1. Description of the experimental site 46
II.2.2.2. Experimental design 47
II.2.2.3. Statistical analyses 48
II.2.3. Genotyping-by-Sequencing 48
II.2.3.1. DNA extraction 48
II.2.3.2. Library preparation and sequencing 49
II.2.3.3. Genotyping-by-Sequencing analysis pipelines 50
II.2.3.4. Genotyping-by-Sequencing accuracy 53
II.3. Identification of QTL for essential yield components in wheat 53
II.3.1. Plant material 53
II.3.2. Phenotyping 53
II.3.3. Genotyping 54
II.3.4. Statistical and population structure analyses 54
II.3.5. Genome wide association analysis (GWAS) 55
II.3.6. Identification of candidate genes 55
II.3.7. Examination of haplotype in and around the candidate gene 55
II.4. Ultra-Dense SNP Genotyping to discover new genomic sources for tolerance to acid soil with aluminum toxicity in wheat 56
II.4.1. Plant material 56
II.4.2. Greenhouse sand screening 56
II.4.3. Data collection and statistical analysis 58
II.4. 4. Citrate efflux analysis 58
II.4.5. Genotyping and systematics analysis of haplotypes around the MATE1B gene 59
CHAPTER III: RESULTS AND DISCUSSION 60

III.1. RESULTS 61
III.1.1. Genetic diversity of Cameroonian bread wheat (*Triticum aestivum* L.) cultivars revealed by microsatellite markers 61
III.1.1.2. Characteristics of markers and genetic diversity 61
III.1.1.3. Genetic relationship and diversity among different geographical regions 62
III.1.2. Genetic population in a global accession of hexaploid wheat breeding lines 65
III.1.2.1. Agro-morphological variation 65
III.1.2.3. Variant discovery in the wheat genome 68
III.1.2.4. Accuracy validation of SNP calling GBS 69
III.1.2.5. Distribution of polymorphic SNP markers on wheat genomes 69
III.1.2.6. Relationship between global germplasm collections 71
III.1.3. Detection of QTLs for important yield components 75
III.1.3.1. Phenotypic variation 75
III.1.3.2. Molecular diversity and analysis of population structure 77
III.1.3.3. Marker–trait associations for grain size in bread wheat 77
III.1.3.4. Candidate gene detection for grain size 79
III.1.3.5. Characterization of haplotype around the D11 gene 80
III.1.3.6. Grouping of haplotypes using phenotypic traits 82
III.1.4. Ultra-Dense SNP Genotyping to discover new genomic sources for tolerance to acid soil with aluminum toxicity in wheat. 82
III.1.4.1. Performance of wheat lines for acid soil tolerance in greenhouse 82
III.1.4.2. Physiological response of genotypes for acid soil tolerance 85
III.1.4.3. Identification of haplotypes around the MATE1B gene 87
III.2. Discussion 91
III.2.1. Genetic diversity in Cameroonian bread wheat cultivars revealed by SSR markers 91
III.2.2. Characterization of a global wheat collection using Genotyping-by-Sequencing 92
III.2.3. Identification of QTLs for essential yield component in wheat 94
III.2.4. Ultra-Dense SNP Genotyping for genomic selection of wheat lines for acid soil under aluminum tolerance 97
CHAPTER IV: CONCLUSION AND PERSPECTIVES 100
IV.1. Conclusion 101
IV.2. Perspectives 102
REFERENCES 103
APPENDIX 117
PUBLICATION 138

LIST OF FIGURES

Fig.1. Wheat domestication events (Matsuoka, 2011) 8

Fig. 2. The three genomes (A, B and D) of bread wheat with each of the seven sets of chromosomes of each ancestor. 8

Fig. 3. Schematic diagram of the relationships between wheat genomes and the history of polyploidization and genealogy (IWGSC, 2014). 9

Fig. 4. Relative genome sizes of various species 10

Fig. 5. Evolution of hexaploid wheat genome sequencing from 2005 (A) to 2014 (B). 11

Fig. 6. Top 10 global (A) and African wheat producers (B) in 2016. 12

Fig. 7. Twelve Mega environments (MEs) or favorable zones for wheat cultivation in Africa, of which three are exploitable in the Nord-west (A), North (B) and Adamawa (C) regions of Cameroon. 13

Fig. 8. Evolutions of productions and importations of wheat in Cameroon (USDA, 2014) 14

Fig. 9. SODEBLE from 1976 to 1982 in Wassandé (Adamawa, Cameroon).. 15

Fig. 10. Steps for preparation of GBS' libraries (Elshire et al., 2011). 19

Fig. 11. Illustration of the two genetic mapping approaches (Soto-Cerda and Cloutier, 2012). 23

Fig. 12. Principles of linkage disequilibrium (LD) and association mapping. 24

Fig. 13. Genetic mapping by Association Mapping approach. 25

Fig. 14. Comparison of different mapping models by association analysis. 27

Fig. 15. Global distribution of acid soils (Von Uexkull et Mutert, 1995; http://websoilsurvey.nrcs.usda.gov/). 30

Fig.16. Relationship between element availability and soil pH (Goedert et al., 1997). 31

Fig.17. Methods illustrating Al exclusion, tolerance and detoxification mechanisms (Kochian et *al*., 2015). 33

Fig.18. Interview and collection of wheat cultivars from Cameroonian farmers in North-west (A), North (B) and Adamawa (C). 37

Fig. 19. Phenotyping of wheat breeding lines in Mbankol (Yaounde, Cameroon) 48

Fig. 20. Steps of libraries preparation and sequencing (The Genotyping-by-Sequencing libraries were prepared in 96-plex following co-digestion with two restriction enzymes *Pst*I (CTGCAG) and *Msp*I (CCGG), as per Elshire et al. (2011). 50

Fig. 21. Bioinformatics FastGBS' steps of filtering and imputation 51

Fig. 22. Schematic representation of the analytical steps in the FastGBS pipeline (Torkamaneh et al., 2016). 52

Fig. 23. Bioinformatics UNEAK analysis pipeline 53

Fig. 24. Screening of wheat breeding lines for acid soil tolerance in greenhouse 57

Fig. 25. Correlation between gene diversity and the number of alleles over 11 microsatellite loci in hexaploid wheat. 61

Fig. 26. Grouping according to the dissimilarities between 17 accessions of hexaploid wheat on the basis of the SSR profiles of 11 loci. 63

Fig. 27. Principal Component Analysis of 17 hexaploid wheat accessions from 6 villages in Cameroon. The grouping is based on Dice's similarity coefficients. 64

Fig.28. Dendrogram using 39 agro-morphological wheat descriptors. Colors indicated the memberships in group 65

Fig. 29. Venn diagrams of the number of SNPs identified in each dataset and with the respective bioinformatics pipeline. 69

Fig.30. Genomic coverage of polymorphic SNP markers over physical map of 21 chromosomes in hexaploid wheat lines. 70

Fig. 31. Phylogenetic relationship of the 278 wheat lines accessions based on the analysis of SNPs generated by GBS. 72

Fig.32. Principal coordinate plot (multidimensional scaling) of all 278 lines (across the A, B and D genomes of hexaploid wheat) against 96667 SNP-markers. 73

Fig 33. Bimodal distribution of grain width (A) and length (B) 76

Fig 34. Linear regression between grain size traits 76

Fig 35. Population structure of 170 hexaploid wheat cultivars, where each vertical line represents a cultivar and each color, a separate subpopulation. 77

Fig.36. Genome-wide association studies of grain traits in hexaploid wheat. 79

Fig. 37. Candidate gene (D11) in associated region on chromosome 2D 80

Fig. 39. Three haplotypes around the D11 gene. Individuals of haplotype A carry alleles (A, T); haplotype B with alleles (C, T) and the haplotype C with alleles (C, C). 81

Fig. 40. Histograms showing the impact of haplotypes on the grain yield and its components. 82

Fig.41. Relative roots length of best wheat lines for Al tolerance 84

Fig.42. Tolerants and sensitives standard checks of wheat lines grown on acid soil with Al toxicity 85

Fig.43. Citrate efflux from excised root apices of evaluated wheat lines for Aluminium tolerance. 86

Fig. 44. Population structure analysis of 45 bread wheat lines for aluminum tolerance 88

Fig. 45. Alleles of wheat cultivars haplotypes around the MATE1B gene 89

Fig. 46. Groups of wheat lines haplotypes around the MATE1B gene 89

Fig. 47. Impact of wheat haplotypes for the MATE1B gene on the phenotypic traits 90

LIST OF TABLES

Table I. Wheat cultivars used and their origins in Cameroon 38

Table II: List of cultivars and their origins 40

Table III: Physico-chemical composition of soils in the Mbankolo area, Cameroon 47

Table IV. The pH and aluminum (Al) concentration of sand plot soil samples taken at 0 to 10 cm depths in the Stellenbosch field used for acidity tolerance screening. 57

Table V. Description of SSR Markers 61

Table VI. Analysis of geographical regions 64

Table VII. Principal Component Analyzes of 39 wheat descriptors 67

Table VIII. Comparison of GBS SNP data between UNEAK and FastGBS pipelines 69

Table IX. Distribution of polymorphic SNP markers across the A, B and D genomes of hexaploid wheat. 70

Table X. SNP diversity summary assessed in cultivars collection 74

Table XI. Number of SNP markers shared between populations (above diagonal) and the estimates of pairwise FST (below diagonal) * 75

Table XII: Descriptive statistics, broad sense heritability (h^2) and F-value from analysis of variance for the grain size descriptors 75

Table XIII: Correlation between seed traits and yield 76

Table XIV. Details of loci associated with grain size traits identified via a genome-wide association study in hexaploid wheat 78

Table XV. Analysis of variances of the phenotypic traits for acid soil tolerance 83

Table XVI. Pearson's correlation coefficient between phenotypic traits 84

Table XVII. Mean performance of wheat lines evaluated under acid soil with aluminum toxicity conditions. 86

LIST OF ABBREVIATIONS

AFLP : Amplified Fragment Length Polymorphism

DUS : Distinction, Uniformity and Stability

GBS : Genotyping-by-Sequencing

LD : Linkage Disequilibrium

PCR : Polymerase Chain Reaction

QTL : Quantitative trait locus

RAPD : Random Amplified Polymorphic DNA

RFLP: Restriction Fragment Length Polymorphism

SNP : Single nucleotide polymorphism

SSR : Simple Sequence Repeat

ABSTRACT

Next Generation Sequencing (NGS) has revolutionized research in plants and animals by developing new high throughput genotyping methods to greatly accelerate the study of genome composition and function. Bread wheat is one of the world's three major cereal crops, with a complex genome whose study requires the use of NGS tools. In addition, It's has been reported that wheat production in Cameroon faces several constraints such as low grain yields and acidic soils with aluminum toxicity. This thesis project was investigated to characterize the genetic diversity and select high-yielding and tolerant wheat varieties to aluminum toxicity. By evaluating the genetic diversity in the 17 wheat cultivars collected in Cameroon, 11 microsatellite markers revealed 77 alleles allowing to discriminate these cultivars. The number of alleles per locus ranged from 2 to 13 with an average of 7. The results obtained in our study provided new information on the genetic relationships in Cameroonian wheat cultivars. The set of microsatellite markers used showed a high level of polymorphism and necessaries information to discriminate these cultivars of hexaploid wheat. Thus, the Cameroonian wheat varieties have been grouped into 5 main groups according to their geographical origin, showing a level of genetic diversity moderately high but insufficient to cope with the biotic and abiotic constraints. In order to increase the level of diversity in Cameroonian cultivars, several wheat varieties have been collected from the major world producers and characterized using the Genotyping-by-Sequencing (GBS) protocol, which is an approach based on the sequencing of short DNA segments (~100-150pb) scattered throughout the genome. The resulting DNA sequences (~ 2.4 million per line) were analyzed using the FastGBS pipeline. This allowed to identify 96,667 SNP markers distributed across all 21 wheat chromosomes. The analysis of genetic diversity revealed the presence of six distinct groups in this international collection, with strong genetic structuring. Very high degrees of molecular diversity have been observed among the populations studied, with a low level of material exchanges between Cameroon and other countries. After the introduction of exotic cultivars, the level of wheat genetic diversity in Cameroon was improved. Using a subset of 71 lines for which chip data (90K SNP Wheat Infinium) were available, we found a high level of agreement (> 95%) between the two genotyping technologies. In order to discover the genomic regions controlling seed size variation, a Genome-Wide Association Study (GWAS) was conducted using the GWAS-GBS approach. A total of 7 SNPs was associated with both traits (grain length and width), identifying 2 quantitative trait loci (QTLs) located on the chromosomes 1D and 2D. In the vicinity of the SNP peak on the chromosome 2D, we found a promising candidate gene (D11), whose homolog had previously been involved in the regulation of grain size in rice. The D11 gene has been

reported to encode a novel cytochrome P450 with homology to the enzymes involved in the biosynthesis of brassinosteroids. In addition, 102 high yielding wheat varieties were selected. In order to select varieties for aluminum tolerance, a total of 45 bread wheat lines were phenotyped in a greenhouse experiment and genotyped by GBS. Citrate efflux of excised root apices was measured on individual seedlings using as few as four apices and its concentration was estimated with coupled enzyme assays that detect the production or consumption of NADH. The root growth rate of all genotypes was reduced with the addition of Al to the pots and Al-susceptible and Al-tolerant wheat genotypes were clearly identified. Correlations between root length and plant height, fresh biomass of shoots / roots, shoot / root dry biomass, released citrate, or tolerance indices in non-acid soil versus acid soil conditions were very significant ($P < 0.01$). Thus, 17 wheat varieties were selected for acidic soil tolerance under aluminum toxicity conditions; Finally, we were able to identify seven (7) high yielding wheat varieties, which are also tolerant to aluminum toxicity (having both MATE1B and D11 functional genes). The present project has made it possible to highlight the level of genetic diversity available in Cameroonian wheat cultivars and in an international collection of wheat lines, selecting high-yielding wheat varieties tolerant to aluminum toxicity, which could be useful to boost wheat productivity in breeding programs.

Keywords: Bread wheat (*Triticum aestivum* L.), Genetic diversity, QTL mapping, Genotyping-by-Sequencing, aluminum tolerance

RÉSUMÉ

Le Séquençage de Nouvelle Génération (NGS) a révolutionné la recherche chez les plantes et les animaux en développant de nouvelles méthodes de génotypage à haut débit pour accélérer considérablement l'étude de la composition et de la fonction du génome. Le blé tendre est l'une des trois principales cultures céréalières au monde, possédant un génome complexe dont son étude nécessite l'utilisation des outils de NGS. Par ailleurs, il a été rapporté que la production de blé au Cameroun est confrontée à plusieurs contraintes telles que les faibles rendements en grains et des sols acides à toxicité aluminique. Le présent projet de thèse a pour objectif de caractériser la diversité génétique et sélectionner des variétés de blé hautement productives et tolérantes à la toxicité aluminique. En évaluant la diversité génétique au sein des 17 cultivars de blé tendre collectés au Cameroun, 11 marqueurs microsatellites ont révélé 77 allèles permettant de discriminer ces cultivars. Le nombre d'allèles par locus variait de 2 à 13 avec une moyenne de 7. L'ensemble des marqueurs microsatellites utilisés a montré un haut niveau de polymorphisme et suffisamment d'informations pour discriminer ces cultivars de blé hexaploïde. Ainsi, les variétés de blé camerounaise se sont regroupées en 5 groupes principaux en fonction de leur origine géographique, montrant un niveau de diversité génétique moyennement élevé mais, insuffisant pour faire face aux contraintes biotiques et abiotiques. A l'effet d'augmenter le niveau de diversité au sein des cultivars camerounais, plusieurs variétés de blé ont été collectées auprès des grands producteurs mondiaux et caractérisées suivant le protocole de génotypage par séquençage (GBS), qu'est une approche basée sur le séquençage de courts segments d'ADN (~ 100-150pb) dispersés dans tout le génome. Les séquences d'ADN ainsi obtenues (~ 2,4 millions par lignée) ont été analysées en utilisant le pipeline FastGBS. Cela a permis d'identifier 96667 marqueurs SNP répartis sur l'ensemble des 21 chromosomes du blé. L'analyse de la diversité génétique a révélé la présence de six groupes distincts dans cette collection internationale, avec forte structuration génétique. Des degrés très élevés de diversité moléculaire ont été observés chez les populations étudiées, avec faible niveau d'échanges de matériel entre le Cameroun et d'autres pays. Apres l'introduction des cultivars exotiques, on a noté une amélioration du niveau de la diversité génétique du blé au Cameroun. En utilisant un sous-ensemble de 71 lignées pour lesquelles des données de puce (90K SNP Wheat Infinium) étaient disponibles, nous avons trouvé un niveau élevé de concordance (> 95%) entre les deux technologies de génotypage. Afin de découvrir les régions génomiques contrôlant la variation de la taille des graines, une étude d'association pan génomique a été menée via l'approche GWAS-GBS. Au total, 7 SNP ont été associés aux deux caractères (longueur et largeur des grains), identifiant 2 locus de caractères quantitatifs (QTL) situés sur

les chromosomes 1D et 2D. Au voisinage du pic SNP sur le chromosome 2D, nous avons trouvé un gène candidat prometteur (D11), dont l'homologue avait précédemment été impliqué dans la régulation de la taille des grains chez le riz. Il a été rapporté que le gène D11 codait pour un nouveau cytochrome P450 avec une homologie avec les enzymes impliquées dans la biosynthèse des brassinostéroïdes. En outre, 102 variétés à haut rendement ont été sélectionnées. A l'effet de sélectionner de nouvelles sources génomiques pour la tolérance au sol acide, un total de 45 lignées de blé tendre a été phénotypés lors d'une expérience en serre et génotypés par GBS. L'efflux de citrate des apex racinaires excisées a été mesurés sur des plantules individuelles en utilisant aussi peu que quatre apex et sa concentration a été estimée avec des dosages enzymatiques couplés qui détectent la production ou la consommation de NADH. Le taux de croissance des racines de tous les génotypes a été réduit avec l'ajout d'Al aux pots et les génotypes de blé Al-sensibles et Al-tolérants ont été clairement identifiés. Les corrélations entre la longueur des racines et la hauteur des plantes, la biomasse fraîche des pousses / racines, la biomasse sèche des pousses / racines, le citrate libéré ou les indices de tolérance dans les conditions sol non-acide versus sol acide étaient très significatives ($P < 0,01$). Ainsi, 17 génotypes de blé ont été sélectionnés pour la tolérance au sol acide dans des conditions de toxicité de l'aluminium ; Finalement, nous avons pu sélectionner sept (7) variétés de blé à haut rendement, qui sont également tolérantes à la toxicité aluminique (ayant à la fois les gènes fonctionnels MATE1B et D11). Le présent projet a permis de mettre en évidence le niveau de diversité génétique disponible chez les lignées de blé camerounaises et dans une collection internationale des lignées de blé, de sélectionner des variétés de blé hautement productives et tolérantes à la toxicité aluminique, qui pourraient être utiles pour booster la productivité de blé dans les programmes d'amélioration.

Mots-clés : Blé tendre (*Triticum aestivum* L.), Diversité génétique, Cartographie QTL, Génotypage par Séquençage, tolérance aluminique

CHAPTER I: GENERALITIES

I.1. Introduction

Context and justification

World demand for wheat is expected to increase by 50% in 2050, to feed the growing world population. Hexaploid wheat (*Triticum aestivum* L.) is one of the three main cereal crops in the world and is derived from hybridization between *Aegilops tauschii* (diploid) and *Emmer* wild *Triticum turgidum* ssp. Dicoccoides, tetraploid (Matsuoka, 2011). Hybridization, domestication, and strong selection pressure have reduced the level of genetic diversity available to wheat breeders, and this lack of diversity is widely recognized as a limiting factor in the selection of high yielding varieties and could increase the inability of plants to respond to biotic and abiotic constraints (Tanksley et McCouch, 1997, Allen et al., 2017). This hexaploid or bread wheat has few pairs of chromosomes than humans, but has more genes, 164,000 to 334,000 genes, compared to 20,000 to 25,000 genes for a human. It obtained two copies of each set of chromosomes from each of his ancestors (diploid), and each of these copies is different. Wheat therefore has the largest genome among commonly grown agricultural crops (Colorado Wheat, 2013). Each set of three wheat chromosomes is almost twice as large as the human genome and the entire rice genome is only half of one of its 21 chromosomes. So, we understand that this species is the owner of a genome of immense size.

In Cameroon, the cultivation of wheat is a livelihood for many families. Its yield is very low in the country because most of the national production (90%) is provided by small family farms for home consumption. With areas of less than 0.5 ha, these farmers (98%) use family labor for all agricultural operations and produce only 1.5 $t.ha^{-1}$. As a result, Cameroon is forced to import more than 550,000 tons of wheat annually to meet the needs of its growing population (USDA, 2014). Thus, Yamdjeu (2013) claimed that in 2012, about 103 billion FCFA was used to import 518,000 tons of wheat flour to Cameroon. However, wheat cultivation has started in Cameroon since 1975, through the Development Corporation for Wheat Cultivation and Processing (SODEBLE). Based in the Adamawa region (Wassandé), SODEBLE had grown rainfed durum wheat, processed wheat into flour, marketed and experimented to improve production technics. Its aim was to reduce the country's dependence on imports of wheat flour. Before its closure in 1982, this company had produced several varieties, some of which are not available to Cameroonian farmers today. Indeed, SODEBLE had produced high-yielding wheat lines resistant to major fungal diseases (Monthé Biris and Habas, 1980). Twelve of these wheat varieties were evaluated for agronomic traits in the Northwest region, to identify high-yielding varieties for Bui and other high agro-ecological zones in Cameroon (Ayuk-Takem, 1984). The

study showed that the local variety (IRAB-1) had the highest yield (4.1 t / ha), but with a non-significant difference between the varieties Chris Mutageneuse (3.5 t / ha) and the Blé Blésil 430 (4 t / ha). However, this yield was significantly better compared to all the other tested varieties. These varieties were also the subject of various agronomic trials in 1985/86. Some of these varieties had not been made available to Cameroonian farmers. To date, no studies have ever been conducted in Cameroon to provide information on the level of genetic variability of accessions grown by farmers.

In the Cameroonian breeding programs, some higher quality cultivars have been lost because farmers' preferences are not generally taken into account in breeding processes. An effective selection should be based on a clear identification of the constraints perceived by farmers and their preferences for cultivars through participatory research (interaction between researchers and farmers). The Northern, Adamawa and Northwestern regions of Cameroon have seen in the recent past the integration of wheat into their cropping system. With the inability of these farmers to solve all their agricultural problems, it has become imperative to work in collaboration with researchers to develop new programs and select suitable wheat varieties in their environment. It is therefore necessary for researchers to identify farmers' preferences in terms of adapted varieties and specific needs for the crop. The socio-economic works carried out to examine Cameroonian farmers' perceptions of wheat varieties, production constraints and preferences for the selection of cultivars to stress-environment showed that farmers usually grow wheat for domestic use and for sale. Wheat production is faced with several constraints such as the problem of low grain yield, soil acidity, late varieties, high fertilizer costs, poor seed conditioning, diseases and pests, and non-use of adapted agricultural practices. It is therefore necessary to assess genetic diversity within local varieties because the ability to fully assess and utilize the genetic diversity present in germplasm collections will contribute to breeding efforts, increasing potential yields, and it has been recognized in recent years that national efforts should be coordinated to maximize the progress of wheat selection (Wheat Initiative, 2011). However, varietal selection can be achieved through conventional and molecular enhancement. But the use of molecular markers to assess genetic diversity is necessary because, unlike phenotypic markers, they are independent of environmental effects (Reza et al., 2015). In this context, the use of SSRs (Simple Sequence Repeats) combines many properties of desirable markers such as abundance, high levels of polymorphism (unlike RFLP), very good reproducibility (compared to RAPD) and the co-dominance (unlike AFLP) for which co-

dominance is not exploitable), but also a uniform coverage of the genome and the specificity of amplification (Tekeu et al., 2017).

Until recently, Single Nucleotide Polymorphisms (SNP) markers are largely replacing microsatellite markers (SSRs) as DNA markers of choice for applications in genetics and plant breeding because they are more abundant, more stable and amenable to automation (Kumar et al., 2012). Indeed, SNP markers represent the most abundant form of genetic variation in eukaryotic genomes because they occur in both coding and non-coding regions of nuclear DNA (Paux et al., 2012). Thus, the use of SNP markers is necessary for the analysis of the wheat genome (large and complex genome) via new sequencing technologies.

Nowadays, Next Generation Sequencing (NGS) has revolutionized research in plants and animals in many ways, including the development of new high throughput genotyping methods to highly accelerate the study of genome composition and their functions (Davey et al., 2011). The ability to characterize germplasm on a common genotyping platform will facilitate exchange of material between countries for introduction and mobilization of new genetic diversity. However, the Genotyping-by-Sequencing (GBS) approach offers a much simpler and more manageable library production procedure for a large number of individuals (Elshire et al., 2011). Thus, GBS is a fast and inexpensive tool for genotyping segregated populations and fixed lined, allowing breeders to implement QTLs mapping studies by association, genetic diversity, genetic linkage analysis, gene discovery and genomic selection.

So, increasing the production to achieve self-sufficiency in wheat depends on the resolution of the main constraints such as low yields and soil acidity which are the major limiting factors of agricultural productivity in forest areas of Cameroon (Ngonkeu, 2009).
For this purpose, grain size (GS) is a key agronomic trait that contributes to grain yield in wheat. As a quantitative trait, GS is mainly and tightly controlled by genetic factors, while grain filling is controlled by both genetic and environmental factors (Sakamoto and Matsuoka, 2008). Several quantitative trait loci (QTL) for grain size have been identified and characterized molecularly in rice and Arabidopsis. Many of these QTLs and genes have been heavily selected to improve rice productivity during domestication, providing an excellent model for future studies (Meyer and Purugganan, 2013).

On the other hand, acidic soils are a global constraint for wheat production, as about 50% of the world's potentially arable soil is acidic. In the humid forest zones of Cameroon, which cover 21.7 million hectares, are made up of 80% of acid soils (Von Uexkull and Mutert, 1995). The

lack of essential nutrients in the soil and the presence of toxic aluminum (Al) cation in the root zone of plants are the main cause of acid soil toxicity. Chemical analyzes of soils sampled in agricultural plantations in the northwest of Cameroon showed very acidic soils (pH <5.5). The acidity of these soils results from the presence of the basic materials which are acidic and which are characterized by an excess of (Al^{3+} + H^{+}) and Mn^{2+}, with deficiencies of Ca^{2+}, Mg^{2+} and $P04^{+}$, a low level of organic matter and a reduced rate of cation exchange. The Upper Farm village, which is home to major agricultural production, is characterized by aluminum toxicity (1.28 cmol / kg) with a saturation rate of 35%. In wheat, Al tolerance is mainly controlled by two genes: TaALMT1, which codes for a malate transporter on chromosome 4D, is constitutively expressed on root apices (Sasaki et al., 2004); TaMATE1 would respond to Al stress based on citrate efflux (Ryan et al., 2009).

Research questions

In order to dynamically revive the wheat production in Cameroon, the research questions were formulated as follows in this project:

- Is the level of genetic diversity of cameroonian wheat sufficient to face the actual agricultural production constraints?
- Which approaches could be used to provide the basis of wheat genetics improvement in Cameroon?
- Are there in an international collection of wheat some quantitative trait loci (QTLs) and promising candidate gene that could be used to improve the productivity of wheat?
- Are there new genomic resources of tolerance to aluminum toxicity in wheat lines introduced in Cameroon?

The hypotheses formulated to bring a light on these research questions are as follows:

Research hypotheses

1. An important genetic structuring can be found within the population of wheat accessions grown in Cameroon.
2. It is possible to enrich the genetic basis of wheat in Cameroon by introducing exogenous cultivars followed by Genotyping-By-Sequencing.
3. It is possible to identify QTLs and candidate gene associated with important grain yield components in an international population of bread wheat lines by association mapping.
4. New sources of tolerance to aluminum toxicity are present in wheat cultivars introduced in Cameroon.

Consequently, the objectives pursued to test these research hypotheses are as follow:

Objectives

The general objective is to characterize the genetic diversity and select high-yielding and tolerant wheat varieties to aluminum toxicity.

More specifically, it will:

1- assess genetic diversity in Cameroonian bread wheat cultivars through the use of microsatellite markers;

2- Characterize the genetic structuring in a local and exotic wheat collection via the Genotyping-By-Sequencing approach;

3- identify QTL and candidate gene for essential grain yield components in bread wheat;

4- select new genomic sources of wheat for the tolerance to aluminum toxicity.

I.2. Literature review

I.2.1. Wheat

I.2.1.1. Domestication of wheat

Wheat is a species divided into three taxonomic groups according to the level of ploidy: the diploid species (2n = 14), tetraploid (2n = 28) et hexaploid (2n = 42) (Feldman, 2001). Thanks to the repeated variation of the nucleotide sequences of the A genome of the diploid wheat species, *Triticum urartu* (2n = 2x = 14, AA) is the main donor of the A genome of all polyploid wheat species (Dvorak et al., 1988). The wheat Einkorn (*T. monoccum*, A^mA^m) is the only wheat grown with diploid husks. The B genome has undergone major changes in DNA when it has formed polyploid wheat species. One thinks that *Aegilops*. *speltoides* is the maternal parent of tetraploid and hexaploid wheat according to cytoplasmic analysis (Wang et al., 1997). The wild species Emmer (*T. dicoccoides*, 2n = 4x = 28, BBAA) is the progenitor of all modern wheat species cultivated (tetraploid and hexaploid). The Durum species (*T. turgidum*) is the most important cultivated tetraploid wheat used for the manufacture of pasta and breads. Among all wheat species, bread wheat (*T. aestivum*, 2n = 6x = 42, BBAADD), which is derived from *T. dicoccoides* and *Ae. Tauschii*, is the most widely grown wheat in the world (Feldman, 2001) (Fig. 1).

I.2.1.2. Origin and evolution of wheat

Wheat has a long history of cultivation and domestication. The Fertile Crescent in Southwest Asia is considered as the center of origin and the current center of distribution and diversity of wheat (DePauw et Hunt, 2001). This area extends to the western Mediterranean, northern Turkey, eastern Iran and south of the Syrian-Arab desert and includes 4 species of wild wheat and 17 species d'*Aegilops*, closely related to B and D genomes (Feldman, 2001). About 13,000 years ago, humans started harvesting wild grain cereals and growing einkorn and emmer wild wheat. Around 7000 years ago, domesticated tetraploid wheat was formed with non-brittleness and free threshing characteristics. It hybridized with *Ae. Tauschii* to become common wheat (bread wheat). About 8,000 years ago, wheat in South-West Asia migrated through Greece to Europe, Egypt to Africa, and northern Iran to East Asia (Feldman, 2001).

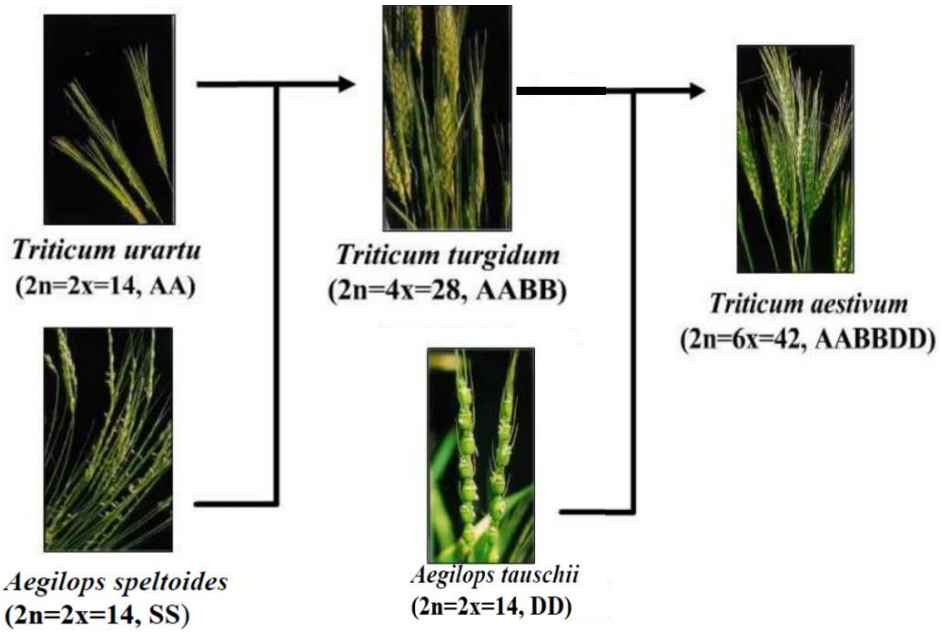


Fig.1. Wheat domestication events (Matsuoka, 2011)

I.2.1.3. Genome complexity in wheat

As in cells of every specie (plant or animal), DNA is organized into long structures called chromosomes. Because of the way on how different species have developed, some of them have more chromosomes than others. Different organisms have different numbers of chromosome copies. While a human cell (diploid) has two copies of 23 chromosomes for a total of 46 chromosomes, a bread wheat cell (hexaploid) has six copies of its seven chromosomes, or 42 chromosomes in total (Fig. 2). Maize has 20 chromosomes in total, and rice has 24 chromosomes (Colorado Wheat, 2013).

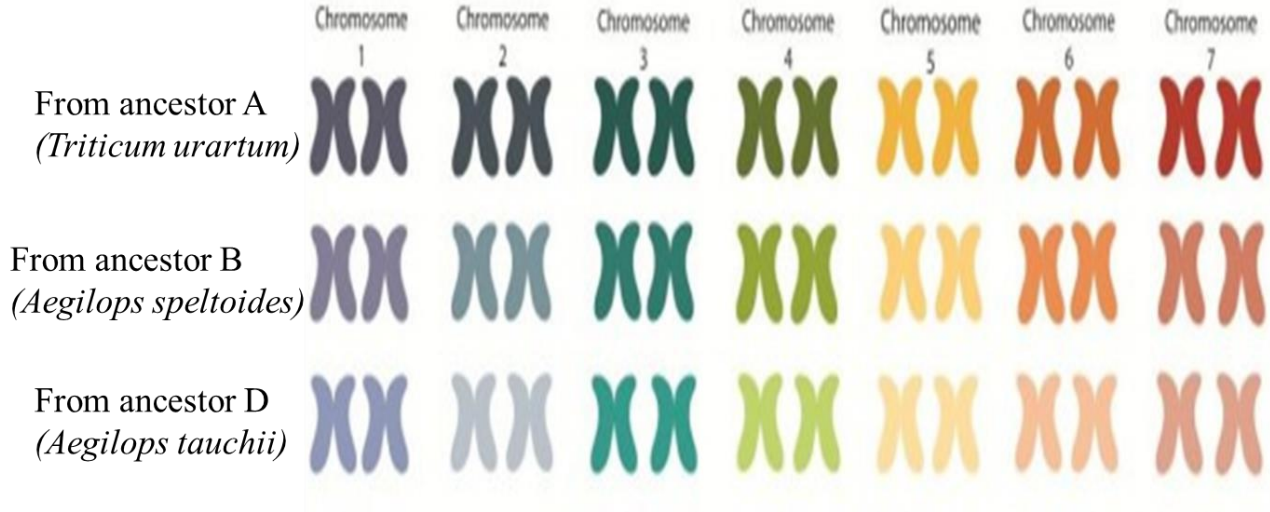


Fig. 2. The three genomes (A, B and D) of bread wheat with each of the seven sets of chromosomes of each ancestor.

Wheat has less pairs of chromosomes than humans but has more genes (164,000 to 334,000 genes), compared to 20,000 to 25,000 genes for a human. A gene is a segment of DNA that contains specific instructions for the structure or function of an organism.

Wheat obtained two copies of each set of chromosomes from each of its ancestors, as the ancestors were each diploid and came together to produce hexaploid wheat (Fig. 3). Each of these copies is different. They were all herbaceous species, so there is some similarity between them, but there are also different.

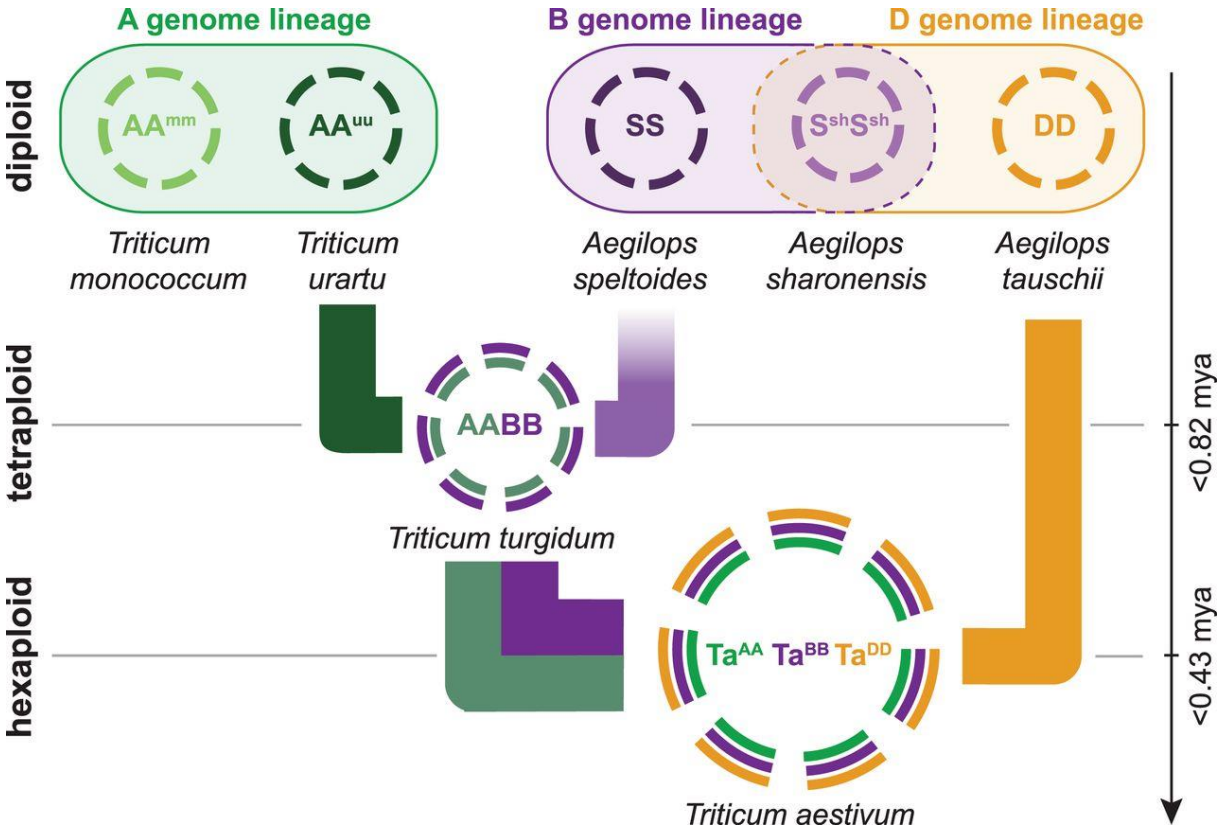


Fig. 3. Schematic diagram of the relationships between wheat genomes and the history of polyploidization and genealogy (IWGSC, 2014).

A genome is all the genetic material of specie, the set of instructions for the creation and function of an organism. Wheat has the largest genome among commonly grown agricultural crops. Each set of three wheat chromosomes is almost twice as large as the human genome and the entire rice genome is only half of one of the 21 wheat chromosomes. Thus, we understand that this specie is the owner of a genome of biggest size, unequaled to that of other species. Its size is about 5 times larger than the human genome, 40 times larger than that of rice and 128 times larger than that of *Arabidopsis thaliana*, which is a model plant (Fig. 4).

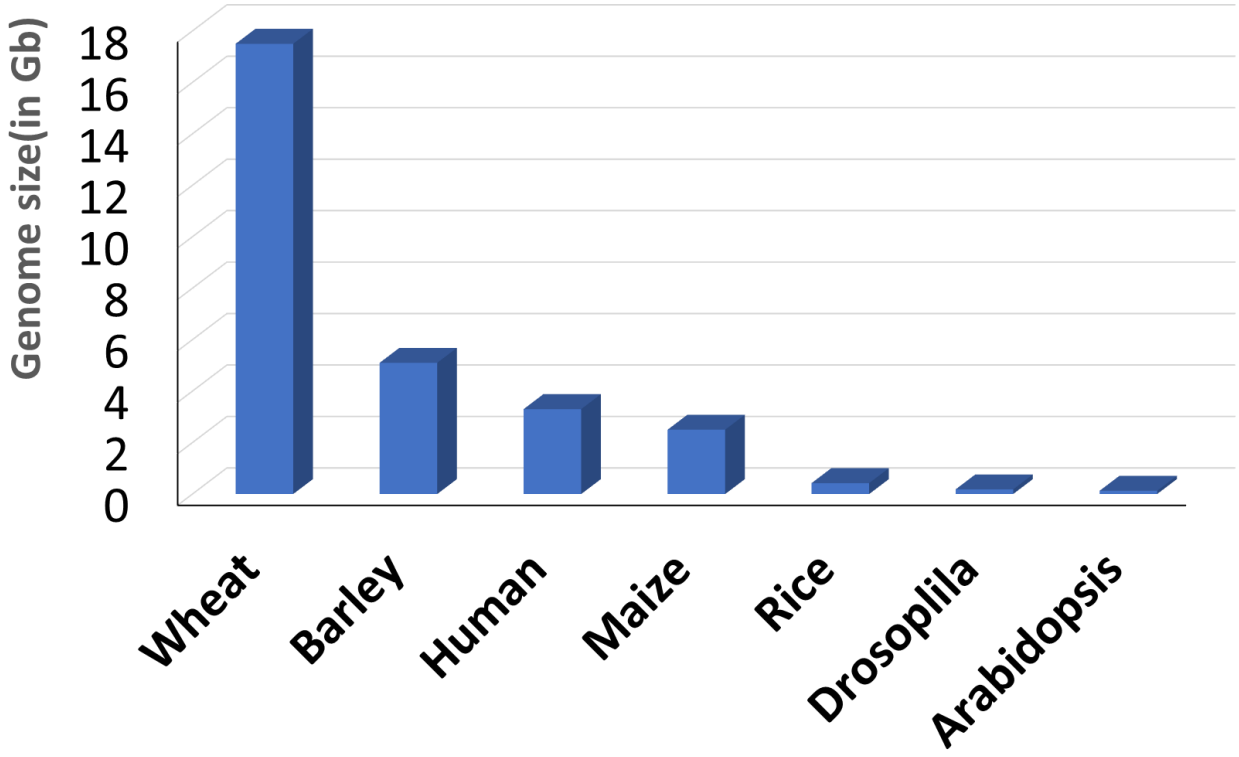


Fig. 4. Relative genome sizes of various species

I.2.1.4. Evolution of genome sequencing in wheat

In 2005, the International Consortium on Wheat Genome Sequencing (IWGSC) was created to facilitate and coordinate an international initiative to accelerate wheat breeding by providing plant breeders and scientists with a genome reference sequence anchored to genetic maps. The strategy adopted by the IWGSC focused on the production of physical maps anchored to genetic maps for each of the 21 individual bread wheat chromosomes and the subsequent sequencing of the minimal pathways of the mapped BACs. To complement the BAC-based sequences, IWGSC has produced other genomic resources specific to the chromosome (Fig.5).

IWGSC (2014) produced 17 gigabase hexaploid of wheat genome sequence via the project by sequencing isolated chromosome arms. They annotated 124,201 gene loci distributed almost equally between homologous chromosomes and sub-genomes, including a genome chromosome-based study sequence that provided early access to gene sequences and, more recently, an entire genome assembly using NRGene DeNovaMAGICTM software. By integrating new data resources with BAC-based resources (sequences, physical maps, WGP tags), IWGSC completed the sequencing of the wheat reference genome v1.0 (www.wheatgenome.org, 2017).

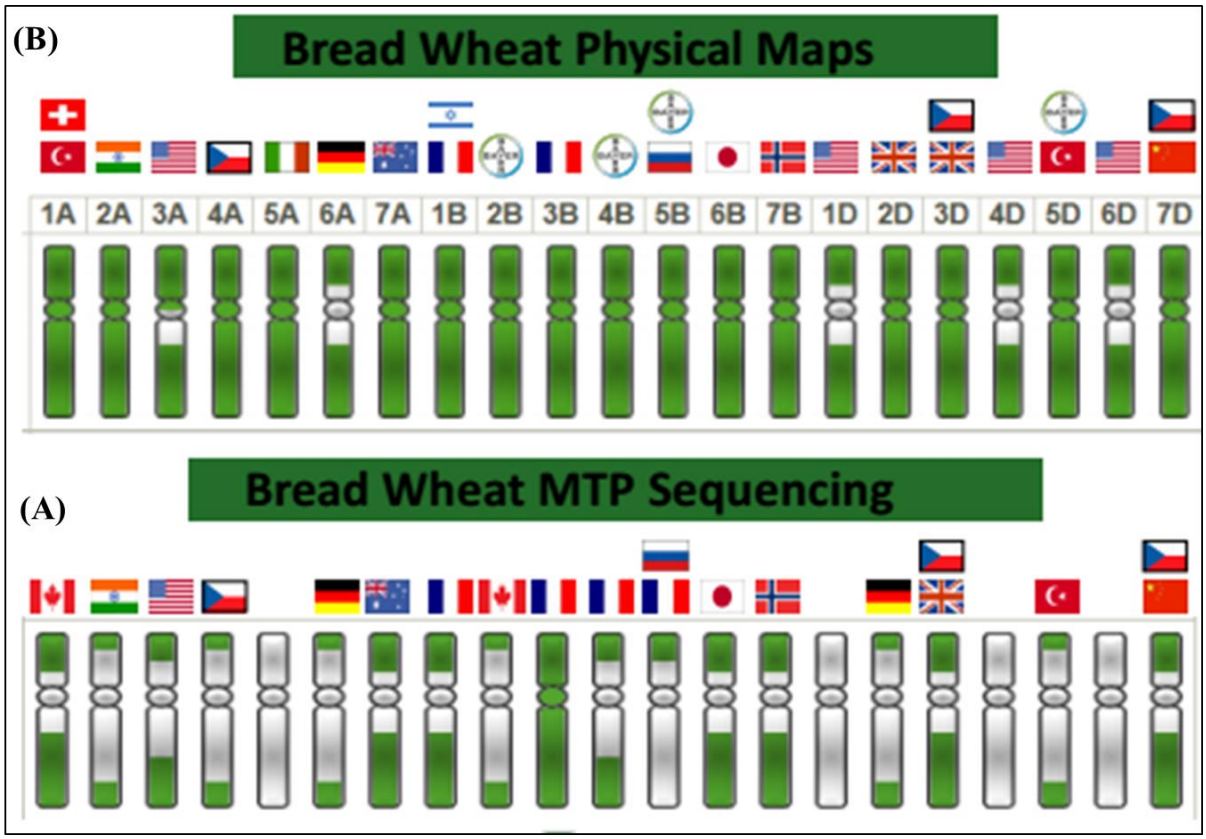


Fig. 5. Evolution of hexaploid wheat genome sequencing from 2005 (A) to 2014 (B). Green color represents the sequenced part of chromosomes and the non-sequenced parts is represents by white color

I.2.2. Importance and production of wheat

Wheat is a grain crop that can be classified into five broad classes. These five categories of wheat include: hard red winter, hard red spring, bread red winter, white wheat and durum wheat. Each class has a different end use and the culture tends to be region specific. Hard red winter wheat is grown mainly in the Great Plains region from Montana to Texas. This type is mainly used for making bread flour. Hard red spring wheat is grown primarily in the Northern Plains region. Their ears of wheat are mainly used for protein blends. Durum wheat, grown primarily in North Dakota and Montana, is known for its excellent pasta production. The wheat class that everyone knows about their breakfast cereals is known as white wheat. Almost all major powers are involved in agricultural wheat production. The largest 10 top global and African wheat producing countries in 2016 are showed (Fig. 6):

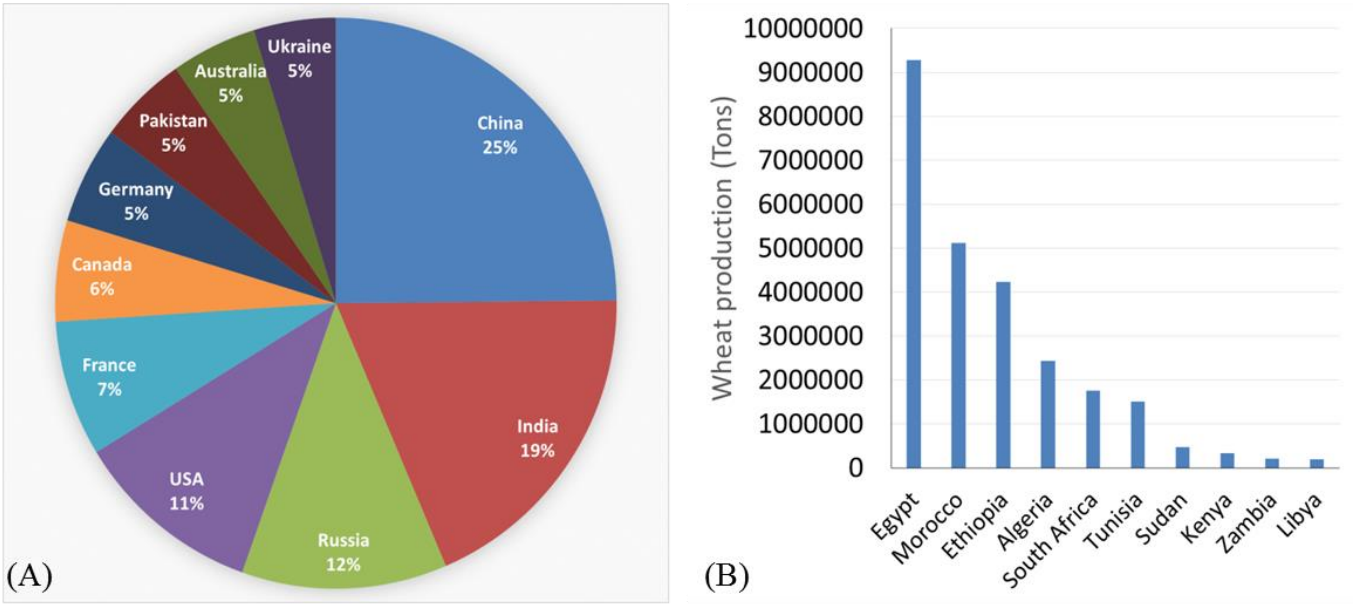


Fig. 6. Top 10 global (A) and African wheat producers (B) in 2016.

I.2.3. Wheat production in Cameroon

World demand for wheat is expected to increase by 50% in 2050 to feed the growing world population (Matsuoka, 2011). To meet this demand, wheat production can be increased by agricultural intensification in growing areas. To this end, Rajaram and Hettel (1994) delineated 12 mega environments (ME) for wheat cultivation, three (ME2, ME3 and ME4) of which correspond to certain agroecological zones in Cameroon. Among them, the main favorable zones for this crop in Cameroon exist in the North, North-West and Adamawa regions (Fig.7). ME2 is characterized by Spring Wheat High Rainfall (10 mill. ha; 75% Bread Wheat), temperate environments with an average of more than about 500 mm of rainfall: during the cropping cycle. The areas are concentrated in West Asia, Central Africa and North Africa, the Highlands of East Africa and Central America, plus the Southern Cone and Andean Highlands of South America. Typical in Toluca, Mexico; Sevilla, Spain and Cameroon (Adamawa, North-West). Rusts, septorias, fusarium, BYDV are the mains diseases:

ME3 is characterized by Spring Wheat Acid Soils (1.7 mill. ha; 100% Bread Wheat), Soils have pH< 5.5. Temperate environments with an average of more than about 500 mm of rainfall during the cropping cycle. Areas are concentrated Mostly in Brazil, the Himalayas, and Central Africa (Cameroon, North-west), Typical in Cruz Alta, Brazil. Rusts, septorias, fusarium, BYDV are the mains diseases and the unavailability of phosphorus, and toxic levels of aluminum and manganese are major abiotic constraints.

ME4 is characterized Spring Wheat Low Rainfall (21.6 mill. ha; 67% Bread Wheat), less than 500 mm of water are available for the crop. There are three major sub-Mes: CIMMYT's wheat breeding mega-environments: ME. A: (10 mill. ha; 53% Bread Wheat) Winter rain followed by late, Mediterranean-type drought. Typical in Aleppo, Cameroon (North, Centre) and Syria. Abiotic constraints are Post flowering moisture and heat stress; ME. B: (5.8 mill. ha; 100% Bread Wheat), Early, winter drought followed by late summer rain. Typical in Marcos Juarez, Argentina. Mostly in Southern Cone. Abiotic constraints predominantly pre-flowering water stress; ME,C: (5.8 mill. ha; 74% Bread Wheat), Crop growth depends largely on soil-stored moisture after monsoon rains. Typical in Dharwar and India. Abiotic constraints are water stress throughout cycle, increasing toward end.

Fig. 7. Twelve Mega environments (MEs) or favorable zones for wheat cultivation in Africa, of which three are exploitable in the Nord-west (A), North (B) and Adamawa (C) regions of Cameroon (Photo from Tekeu, 2015).

Wheat farming is a livelihood for many families in Cameroon. Its yield is very low in the country because most of the national production (90%) is provided by small scale farms for family consumption. With an area of less than 0.5 ha, these farmers (98%) use family labor to carry out all agricultural operations and produce only 1.5 t.ha-1. As a result, Cameroon is forced to import nearly 300 000 tons of wheat per year to meet the needs of its growing population from 2000 to 2007 (USDA, 2014). Due to the food crisis in 2008, the importation has doubled (Fig. 8). Thus, Yamdjeu (2013) indicated that in 2012, about 103 billion FCFA was used to import 518,000 tons of wheat flour to Cameroon.

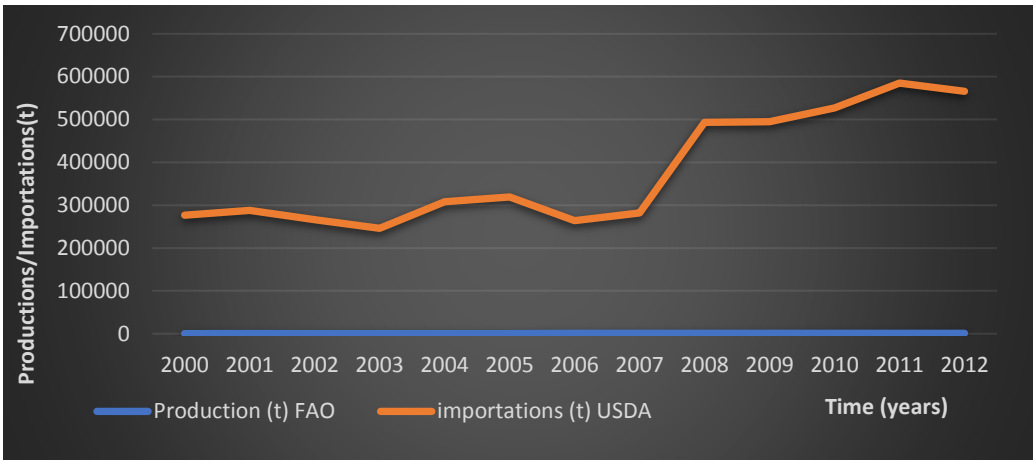


Fig. 8. Evolutions of productions and importations of wheat in Cameroon (USDA, 2014)

In addition, wheat cultivation has started in Cameroon since 1975, through the Development Corporation for Wheat Cultivation and Processing (SODEBLE). Based in the region of Adamaoua (Wassandé), SODEBLE had grown durum wheat, wheat processing into flour, marketing and experimentation for the improvement of techniques of production. Its aim was to reduce the country's dependence on imports of wheat flour. Before its closure in 1982 (Fig. 9), this company had produced several varieties, some of which are not available to Cameroonian farmers today. Indeed, SODEBLE had produced high-yielding wheat lines resistant to major fungal diseases (Monthé Biris and Habas, 1980). Twelve of these wheat varieties were evaluated for agronomic traits in the Northwest Region (Ayuk-Takem 1984). In addition, Ayuk-takem (1984) assessed the agronomic characteristics of twelve varieties of bread wheat in the Northwest region to identify high-yielding varieties for Bui and other high agro-ecological zones in Cameroon. The study revealed that the local variety (IRAB-1) had the highest yield (4.1 t / ha), but with a non-significant difference between the varieties Chris Mutageneuse (3.5 t / ha) and the Ble bresil 430 (4 t / ha). However, this yield was significantly better compared to all the other tested varieties. These varieties were also the subject of various agronomic trials in 1985/86. In doing so, some varieties were not made available to Cameroonian farmers. However, no studies have ever been conducted in Cameroon to provide information on the level of agro-morphological and genetic variability of genetic material cultivated by farmers.

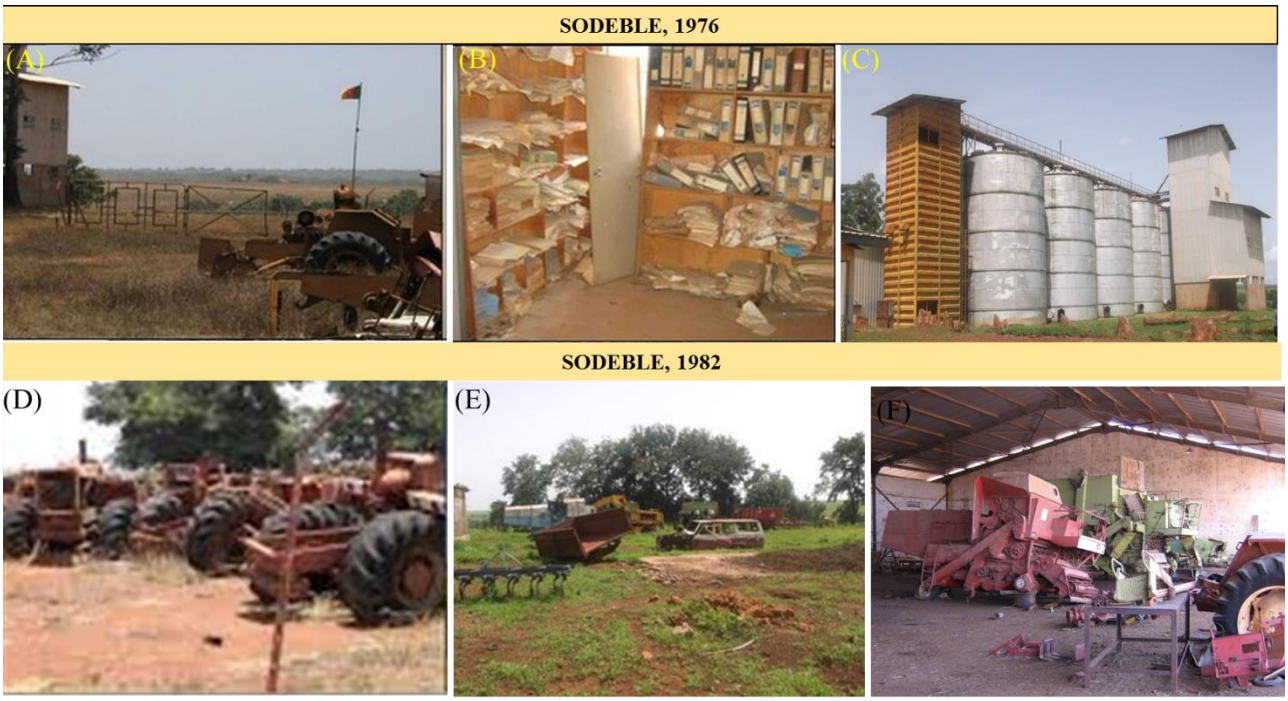


Fig. 9. SODEBLE from 1976 to 1982 in Wassandé (Adamawa, Cameroon). SODEBLE had grown wheat (A), conserved the seeds in gene bank (B), keep the seeds in silos to turn into flour (C), marketing and experimentation for the improvement of techniques of production. During its closure in 1982, sowing tractor (D), agricultural fields (F) and machines for managing wheat into flour (F) were lost.

In the Cameroonian breeding programs, some higher quality cultivars have been lost because farmers' preferences are not generally taken into account in breeding processes. Breeder tended to set production targets for the needs of the commercial agriculture sector. An effective selection should be based on a clear identification of the constraints perceived by farmers and their preferences for cultivars through participatory research (interaction between researchers and farmers). Some breeders focus more on the selection of varieties for their agronomic values and farmers perceive little benefit in growing these varieties because they are not designed for their needs (Reeves and Cassaday, 2002) and sometimes because they are unaware of the existence of the development of these new improved varieties (Tandzi, 2015). For effective selection, farmers' preferences for varieties should be clearly identified through interaction and collaboration between researchers and farmers (Sibiya et al., 2013).

Danial et al. (2007) showed that breeders who involve farmers in breeding programs learn more about the most important selection criteria for growers, producing preferred cultivars in rural growing environments. This idea motivated the use of locally adapted cultivars, which resulted in less dependence on foreign matter. Such a strategy has led to the selection and development of new cultivars of wheat, barley, beans, quinoa, potatoes and maize in various parts of the

world (Danial et al., 2007). For example, breeders should be familiar with farmers' preferences, such as requirements for agronomic, storage, processing and marketing characteristics, if the adoption rate is high (Danial et al., 2007).

The regions of North, Adamawa and North-West Cameroon have seen in the recent past the integration of wheat into their farming. With the inability of these farmers to solve all their agricultural problems, it has become imperative to work in collaboration with researchers to develop new programs and select suitable wheat varieties in their environment. It is therefore necessary for researchers to identify farmers' preferences in terms of adapted varieties and specific needs for the crop. A participatory rural appraisal, including a group discussion, was conducted to analyze farmers' perceptions of wheat cultivars and identify constraints and their needs.

I.2.4. Genetic diversity in hexaploid wheat

Hexaploid wheat (*Triticum aestivum* L.) is one of the three main cereal crops in the world and is derived from the hybridization of Aegilops *tauschii* diploid with wild *Emmer* tetraploid, *Triticum turgidum* ssp. dicoccoides (Matsuoka, 2011). Hybridization, domestication and strong selection pressure have reduced the level of genetic diversity available to wheat breeders, and this lack of diversity is widely recognized as a limiting factor in the selection of high yielding varieties, particularly in response to biotic and abiotic challenges (Tanksley and McCouch, 1997, Allen et al., 2017). The ability to assess and make full use of the genetic diversity present in the germplasm collections will contribute to breeding efforts to increase potential yields, and it has been recognized in recent years that national efforts should be coordinated to maximize the progress of wheat selection (Wheat Initiative, 2011).

Several studies have shown that narrow genetic diversity can increase susceptibility to diseases and pests, as well as the ability of plants to respond to changing environmental conditions (Gorji and Zolnoori, 2011). It is therefore fundamental to understand the levels and distribution of genetic diversity in hexaploid wheat, as a basis for developing resource management and exploitation strategies in Cameroon, in order to increase the national economy through the agricultural sector, improve farmers' livelihoods, reduce the use of chemical inputs, increase basic seed production and improve seed availability.

In addition, varietal selection can be achieved through classical and molecular breeding. But the use of molecular markers to assess genetic diversity is necessary because, unlike phenotypic markers, they are independent of environmental effects (Reza et al., 2015).

In this context, varietal improvement has progressed so rapidly that several types of molecular markers have been developed and used for decades. The RFLP was first applied as DNA markers in plant genotyping (Botstein et al., 1980). With the evolution in biotechnology, several types of PCR-based markers have been developed and used in plant breeding programs and these markers mainly include RAPDs, SCAR, CAPS, SSR; AFLP; LRAD. Compared to RFLP, all of these PCR-based markers are amplified from small-scale, relatively inexpensive and shorter-length individual genomic sequences (Jiangfeng et al., 2014).

I.2.5. Technologies of genomic sequencing

The use of SNPs as DNA markers for plant genotyping has increased the potential for variation in specific DNA fragments (Lander, 1996). In addition, information on potential millions of genomic SNPs or small deletions-insertion and their surrounding sequences establishes the basis for high-throughput genotyping across a wide range of genomic scales (Zhu et al., 2003). Today, strong demand for low cost sequence data has led to the development of high throughput sequencing (NGS) technologies that can produce 1000 or millions of sequences simultaneously (Jiangfeng et al., 2014). The ability to characterize germplasm on a common genotyping platform will facilitate exchange of material between countries for introduction and mobilization of new genetic diversity. High throughput genotyping of hexaploid wheat has been made possible in recent years by the advent of next generation sequencing for sequencing genotyping (GBS, Rife et al., 2015) and SNP marker discovery (Winfield et al., 2012). These range from flexible and scalable PCR-based simple assays such as KASP (Allen et al., 2011) and TaqMan to high-density fixed content arrays, such as the Illumina 90k iSelect Matrix (Wang et al. , 2014). Recently, Winfield et al. (2015) reported the generation of an ultra high density Affymetrix Axiom matrix containing 820k SNP markers.

Similarly, several NGS platforms such as Roche 454 FLX Titanium (Thudi et al., 2012), Illumina MiSeq and HiSeq2500 (Bentley et al., 2008), Ion Torrent PGM (Torkamaneh and Belzile, 2015) have been developed. Up to date, NGS progress has resulted in DNA sequencing costs as GBS is now very accessible to assess species diversity at the genomic scale. The GBS method offers a much simpler library production procedure that can be used for a large number of individuals (Elshire et al., 2011). A two-enzyme GBS (PstI / MspI) protocol, which provides a greater degree of complexity reduction and a uniform library for sequencing, than the original protocol (using the ApeKI enzyme) has now been developed and applied to wheat and barley (Poland et al., 2012a). Sonah et al. (2013) described a modified library preparation protocol, in which selective amplification is used to increase both the number of SNPs called and their depth

of coverage, resulting in a high yield allowing a significant reduction in cost per sample. Two different GBS strategies have been developed with the PGM Ion system (Poland et al., 2012a). The first, called restriction enzyme digestion, in which no specific SNPs were identified and ideal for discovering new markers for marker assisted breeding programs. The complexity of the genome under this approach is reduced by digesting the DNA with one or two selected restriction enzymes prior to ligation. The construction of GBS databases is based on the reduction of genome complexity with restriction enzymes (RE, Elshire et al., 2011). This approach is simple, fast, highly specific, highly reproducible and can reach large areas of the genome inaccessible to sequential capture approaches. By selecting appropriate ERs, repetitive regions of genomes can be avoided, and lower copy regions can be targeted with two to three times higher efficiency (Gore et al., 2007). To confirm the accuracy of the imputed data, Torkamaneh and Belzile (2015) explored the use of imputation to replace the missing data proportions of soybean genotypes in the GBS dataset. Thus, the GBS is a fast and inexpensive tool for genotyping segregated populations and fixed lines, allowing breeders to implement QTLs association studies, study genomic diversity, analyze genetic links, discover molecular markers and genomic selection. It is not necessary to know priory the genomes of the species, because the GBS method has proven robust for a range of species whose discovery and genotyping of SNPs are completed (Narum et al., 2013). Initially, the GBS approach was developed in three species (maize, wheat and barley) and the results were very interesting. Indeed, Elshire et al. (2011) reported more than 25,000 SNPs in an F2 maize population. In barley and wheat more than 34,000 and 20,000 SNPs were obtained respectively (Poland et al., 2012). The reduction of genome complexity is achieved using restriction enzymes. The use of methylation-sensitive restriction enzymes is an approach that largely avoids the heterochromatic regions of the genome that are rich in repetitive and gene-poor sequences. The entire protocol follows the steps (Fig. 10): (i) extraction of high quality DNA, (ii) selection of appropriate enzymes and adapters, (iii) preparation of GBS libraries for sequencing, (iv) sequencing multiplexed single-end libraries, (v) filtering / evaluating sequence quality, (vi) sequence alignment, (vii) and SNP calling.

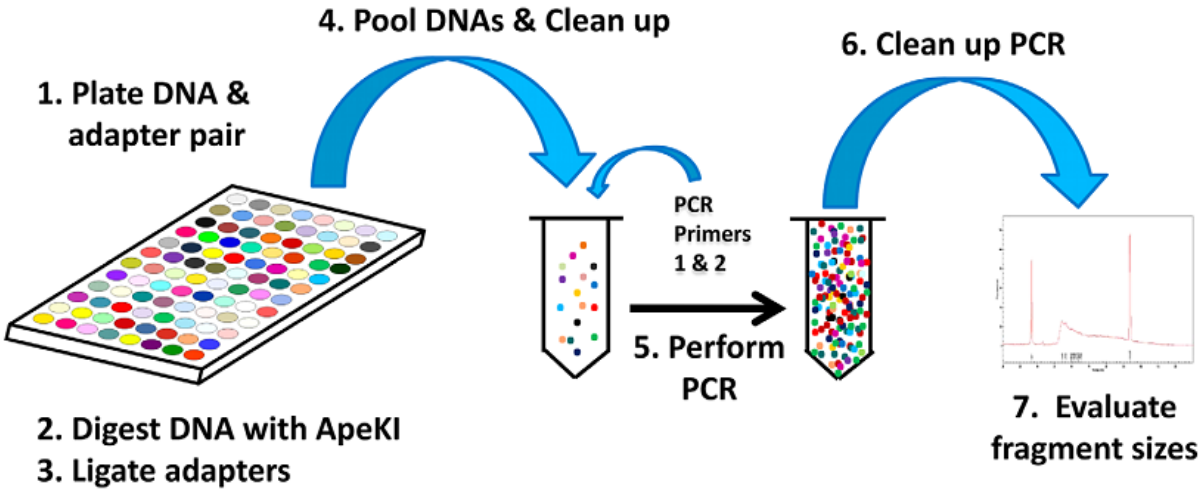


Fig. 10. Steps for preparation of GBS' libraries (Elshire et al., 2011).

I.2.6. Bioinformatics Algorithms and Pipelines for DNA Sequence Analysis

Understanding the application of GBS across different sequencing platforms will help to adapt this approach for future technological developments. Keeping in mind the advancement technique in DNA sequencing, new bioinformatics algorithms and pipelines for sequencing analysis are also under development. Generally, GBS data processing pipelines are classified into two groups; based on novo and based on the reference. When a reference genome is available, reduced representation sequencing reads can be mapped to that genome and SNPs can be called as for whole genome resequencing projects (Li and Durbin, 2009, Nielsen et al., 2011). To date, several reference-based GBS analysis pipelines have been developed. The most widely used GBS benchmark test pipelines are: TASSEL-GBS, Stacks, IGST and Fast-GBS (Sonah et al., 2013, Bradbury et al., 2007, Catchen et al., 2013). In the absence of a reference genome, almost identical pairs of readings (assumed to represent alternative alleles of a locus) must be identified (Torkamaneh et al., 2016). The most used pipelines for such a de novo approach are UNEAK and Stacks (Catchen et al., 2013, Lu et al., 2013). Thus, we compared the UNEAK and FastGBS analysis pipelines on the basis of the number of SNPs called, the precision of the resulting genotypes. Then, the panel of polymorphic SNPs obtained was used to characterize the genetic diversity in a global hexaploid wheat collection.

I.2.7. Genetic mapping

Genetic mapping involves positioning one or more genes on a genetic linkage map in relation to molecular markers that serve as reference points. In the case of simple genetic traits, which are controlled by a single or very small number of genes, it will be relatively easy to identify the genomic regions where the genes in question are located. On the other hand, for complex or quantitative characters, it will be several QTL ("quantitative trait locus") that will be

responsible for the observed variation. Each of these QTLs is presumed to correspond to a gene (Elmer, 2014).

In 2012, genetic mapping in plants relied entirely on an approach called "two-parent mapping", based on the analysis of offspring from a cross between two individuals. Then, a new approach called "Association Mapping" has been proposed in plants (Soto-Cerda and Cloutier, 2012). Both approaches are based on the principle that genes and QTL can be tracked through their genetic linkage with molecular markers (Bastien, 2013).

I.2.7.1. Bi-parental (two-parent) mapping

Bi-parental mapping, also called conventional mapping, is based on the association between genes / QTLs that govern a trait and genetic markers that are located nearby in a segregated population, developed by crossing two parents with different values contrasting phenotypes for a trait of interest (Zhu et al., 2008). Bi-parental mapping can be performed with populations of various natures such as second generation (F2) individuals, Recombinant Inbred Lines (RILs), isogenic lines (NILs), double haploid lines (HD) or backcross populations (Mackay and Powell 2007). However, it is necessary to develop experimental populations using parents who contrast for the traits being studied (such as two parents or one is tolerant and the other sensitive to biotic or abiotic stress). Bi-parental mapping uses polymorphic markers between parental genotypes to identify markers that are associated with the locus responsible for variation in phenotypic expression of the trait being assessed in multiple environments (Collard et al., 2008).

This biparental mapping approach requires the creation of a genetic map, which can be considered as a "road map" of chromosomes from two different parents (Elmer, 2014). Genetic maps indicate position and relative genetic distances between markers throughout the chromosomes. QTL mapping is based on the principle that genes and markers are segregated by recombination of chromosomes (called "crossing-over") during meiosis, thus allowing their analysis in offspring (Semagn et al., 2010; Mammadov et al. al., 2012).

The four methods that are commonly used to detect QTLs are punctual analysis, simple interval mapping, compound interval mapping, and inclusive composite interval mapping (Wang, 2013).

I.2.7.1.1. Punctual analyses

It is the simplest method for detecting QTLs associated with markers. Statistical methods used to analyze a single marker include the t test, analysis of variance and linear regression. Linear regression is the most commonly used because the coefficient of determination of the marker (R^2) explains the part of the phenotypic variation resulting from QTL that can be linked to the

marker. This method does not require a complete link map and can be performed with basic programs found in any statistical analysis software.

However, the major disadvantage of this method is that the distance between a marker and a QTL greatly affects the probability of detecting the QTL and measuring the effect of this QTL (Collard et al., 2005). Indeed, in such an approach, it is not possible to distinguish whether our significant marker is close to a QTL with a low phenotypic impact or is far from a QTL with a greater impact.

I.2.7.1.2. Simple interval mapping (SIM)

SIM uses linkage maps and analyzes the intervals between pairs of adjacent markers all along a chromosome at the same time, instead of analyzing the markers individually (Semagn et al., 2010). The use of linked markers in such an analysis offsets the recombination between markers and QTL and is considered statistically more powerful as compared to punctual analysis (Li et al., 2007, Eeuwijk et al., 2010). Many QTL mapping software has been developed for SIM analysis, such as MapMaker / QTL (Lincoln et al., 1993b) and QGene (Nelson, 1997).

I.2.7.1.3. Composite interval mapping (CIM)

CIM has become popular recently for QTL mapping. This method combines interval mapping with linear regression. It includes additional genetic markers in the statistical model, in addition to a pair of linked markers for interval mapping (Li et al., 2007; Wang, 2009). The main advantage of the CIM is that it is more accurate and efficient in locating QTLs as compared to single point analysis and interval mapping, especially when related QTLs are involved. Many software packages are available for such analyzes: QTL Cartographer (Basten et al., 1994, 2004), QTX MapManager (Manly et al., 2001) and PLABQTL (Utz and Melchinger, 1996).

I.2.7.1.4. Inclusive Composite Interval Mapping (ICIM)

ICIM is a new statistical method, consisting of two stages. In a first step, stepwise regression is applied to identify the most significant regression variables. In the second step, one dimensional interval analysis is done to detect additive (and dominant) QTLs. ICIM retains all the benefits of CIM on interval mapping (IM), and avoids a possible increase in sampling variance and the complicated process of selection of background markers in CIM (Elmer, 2014). In-depth simulations using two genomes and various genetic models indicated that ICIM increases detection power, reduces the rate of false detections, and produces less biased estimates of QTL effects. ICIM provides intuitive statistics for testing additive, dominance and epistasis effects, and can be used for most experimental populations from two parental lines (Wang, 2013). Bi-parental mapping nevertheless presents certain challenges and shortcomings, namely:

- the production of mapping populations; which may require several years of work depending on the type of population (Fig. 11a and 11b).

- the QTLs identified at the end of such work are only valid for this population and are not necessarily useful in other populations;

- the resolution of such an analysis is normally insufficient for anyone who would like to identify possible candidate genes because the genomic regions in which the QTLs are located are usually very large (several million of base pairs).

Thus, it may be interesting to consider association mapping

I.2.7.2. Association Mapping

Association mapping (AM) is another approach of genetic mapping that is very distinct from conventional bi-parental mapping. Indeed, it is based on the characterization of distinct lines, not coming from a cross (Fig. 11). It makes it possible to use existing collections of lines characterized for many different phenotypes, and also to take advantage of historical data of phenotypes measured during the development of new varieties (Waugh et al., 2009, Zhu et al., 2008). It allows to exploit a greater allelic diversity and to benefit from a higher resolution due to more historical recombinations (Zhu et al., 2008). Association mapping is based on linkage disequilibrium between markers and QTLs. Linkage disequilibrium (LD) is the non-random association of alleles at two or more loci in a population (Soto-Cerda and Cloutier 2012). Generally, the LD tends to be higher between the locus alleles that are located close to each other. The more distant the loci, the more likely they will be out of linkage disequilibrium because of the genetic mixing that occurs during meiotic genetic recombination.

Several measures can be used to describe the LD such as the normalized imbalance coefficient (D') or the square of the allelic correlation between two loci (r^2) (Slatkin, 2008, Zhao et al., 2007). However, r^2 would be the most relevant measure of LD for the identification of markers strongly associated with phenotypic variation in plants (Zhu et al., 2008). The illustration of the LD between two loci, locus 1 and locus 2 showed that in the first case (a), the loci do not show a statistical correlation with the phenotype (seed color). Unlike in the second case (b), the loci show a significant covariance with the phenotype and this is considered, as evidence of association. Thus, LD results in an association between a marker allele and the phenotype studied (Fig. 12).

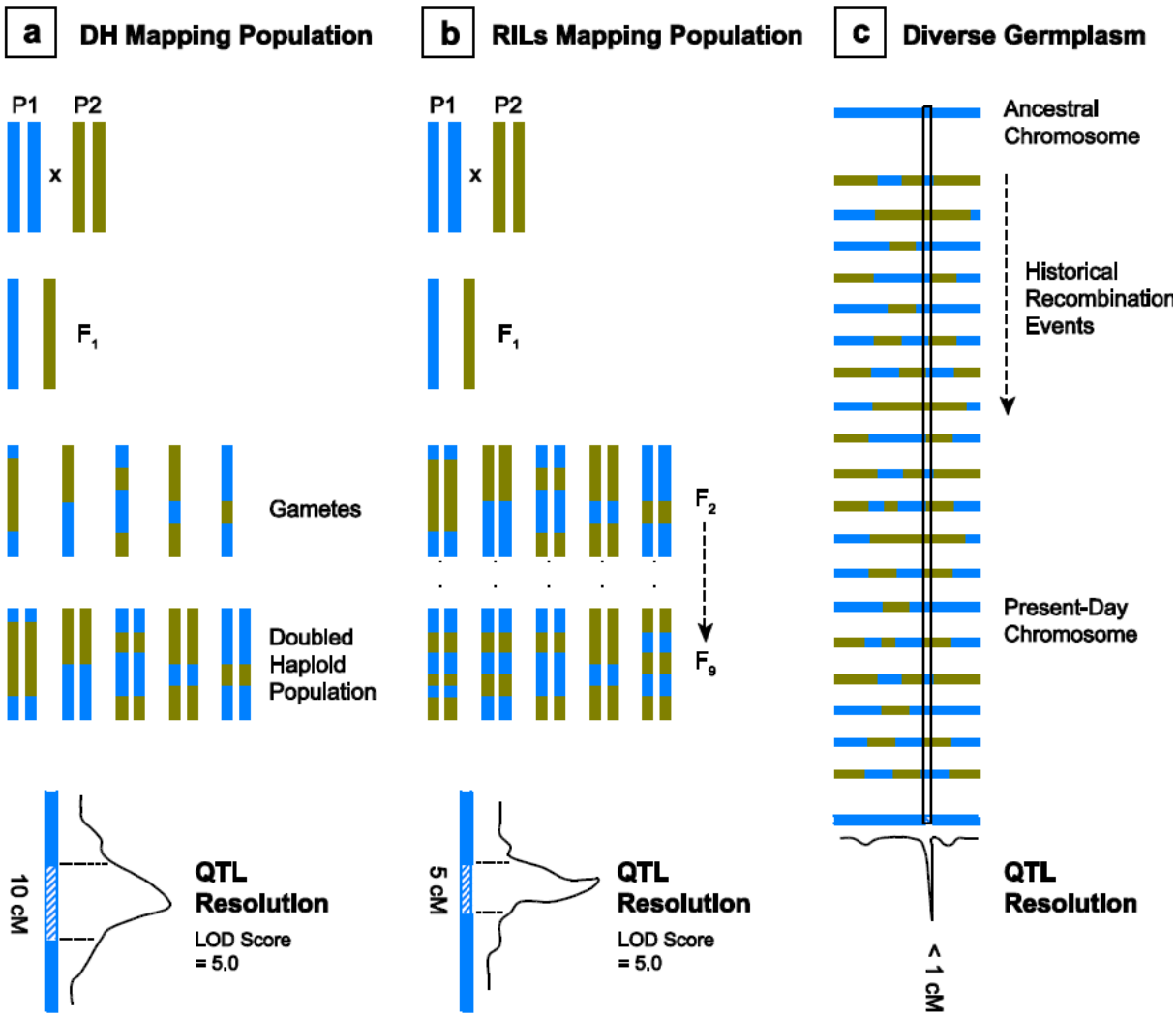


Fig. 11. Illustration of the two genetic mapping approaches (Soto-Cerda and Cloutier, 2012).

The resolution obtained and the density of markers required for association mapping are determined by the extent of genomic LD within a collection of analyzed individuals. If the LD extends a short distance, the resolution of the mapping will be high, but a very large number of markers will be required (Chen, 2013, Zhu et al., 2008). On the other hand, if the LD extends over a long distance, then the resolution of the mapping will be small, but a smaller number of markers will be sufficient (Bradbury et al., 2007). Therefore, a thorough understanding of the extent of genomic LD is one of the keys to the design and implementation of association mapping. The extent of LD is affected by many factors such as allelic frequency and recombination rate, population structure, mutation, and mode of reproduction (Flint-Garcia et al., 2003). In general, the LD extends over a much larger distance in inbred species such as soybeans than in cross-pollinated species such as maize. Similarly, the presence of structure (subdivision) within the population may increase the LD and thus lead to the detection of false

associations between the genotype and the trait of interest. Thus, it is important to take into account such a structure of the population during the analyzes (Falush et al., 2007, Yu et al., 2006, Zhao et al., 2007b, Yan et al., 2011).

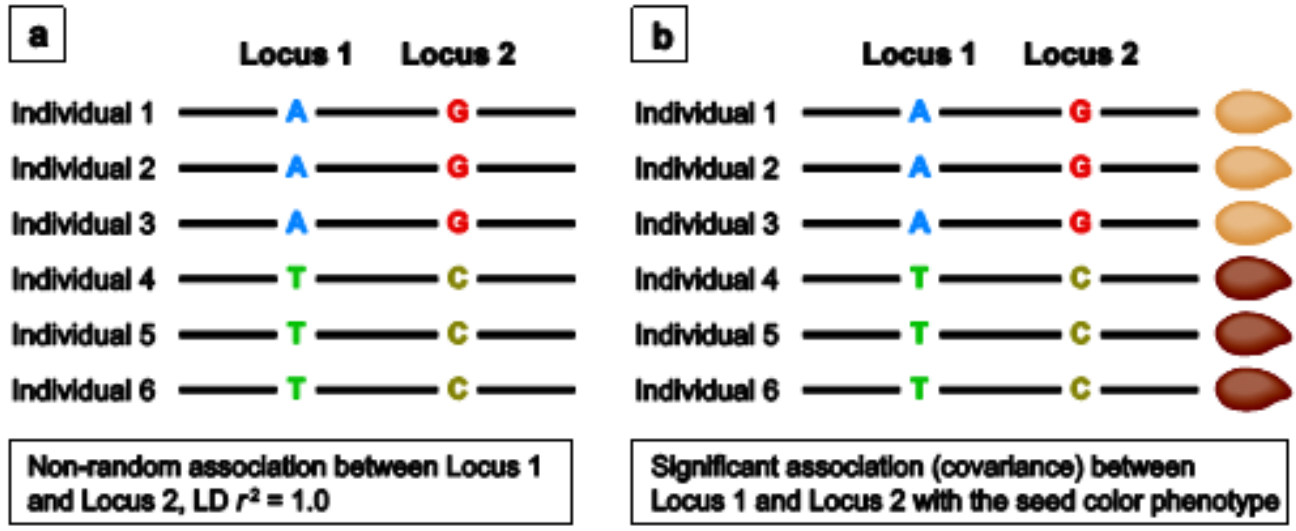


Fig. 12. Principles of linkage disequilibrium (LD) and association mapping.
(a) Linkage disequilibrium. Locus 1 and Locus 2 present an unusual pattern of association between alleles A-G and T-C, which deviate from Hardy-Weinberg expectations, but without any statistical correlation with a phenotype. (b) Association mapping. Locus 1 and Locus 2 are in LD. Significant covariance with the seed color phenotype is considered evidence of association.

The general approach to an association mapping includes the following steps:

1.- Selection of a collection of various genotypes that may include more or less wild accessions, varieties of countries, elite cultivars, etc. On the other hand, there is a cost associated with genotyping and phenotyping large populations, especially for traits requiring large field trials. As a result, the size of the population, the number of repetitions and sites for phenotyping is frequently a limitation (Semagn et al., 2010).

2.- Complete and accurate phenotyping of the population for the traits of interest. On the other hand, a crucial factor determining the success of the analysis mapping is the intensity of phenotyping. Strong heritability is key to achieving reliable results (Würschum, 2012).

3.- Very dense genotyping with adequate molecular markers. This will depend on the size of the genome and the deterioration ("decay") of the LD. For example, in rice that has a slow deterioration of LD, thousands of markers are needed. In wheat with a very large genome and rapid deterioration of LD, thousands to millions of markers are needed.

4.- An analysis of the structure of the population and kinship relations to limit false positives. Population structure is due to the presence of two or more major subpopulations, while family structure refers to different levels of kinship relationship between individuals (Würschum, 2012).

5.- The measurement of the LD. The two ways to visualize the extent of LD between a pair of markers are the graphs of LD deterioration and the imbalance matrices (Semagn et al., 2010).

6.- Finally, joint analysis of genotype and phenotype using appropriate software to identify markers that have a significant association with the trait (Al-Maskri, 2012).

All the above different components are integrated into an association analysis approach (Fig. 13). A group of unrelated individuals ("germplasm") is characterized for both phenotype and genotype. Genotypic data is then used to characterize population structure and kinship relationships between lineages. Finally, all this information is exploited in the association analysis, which identifies QTLs.

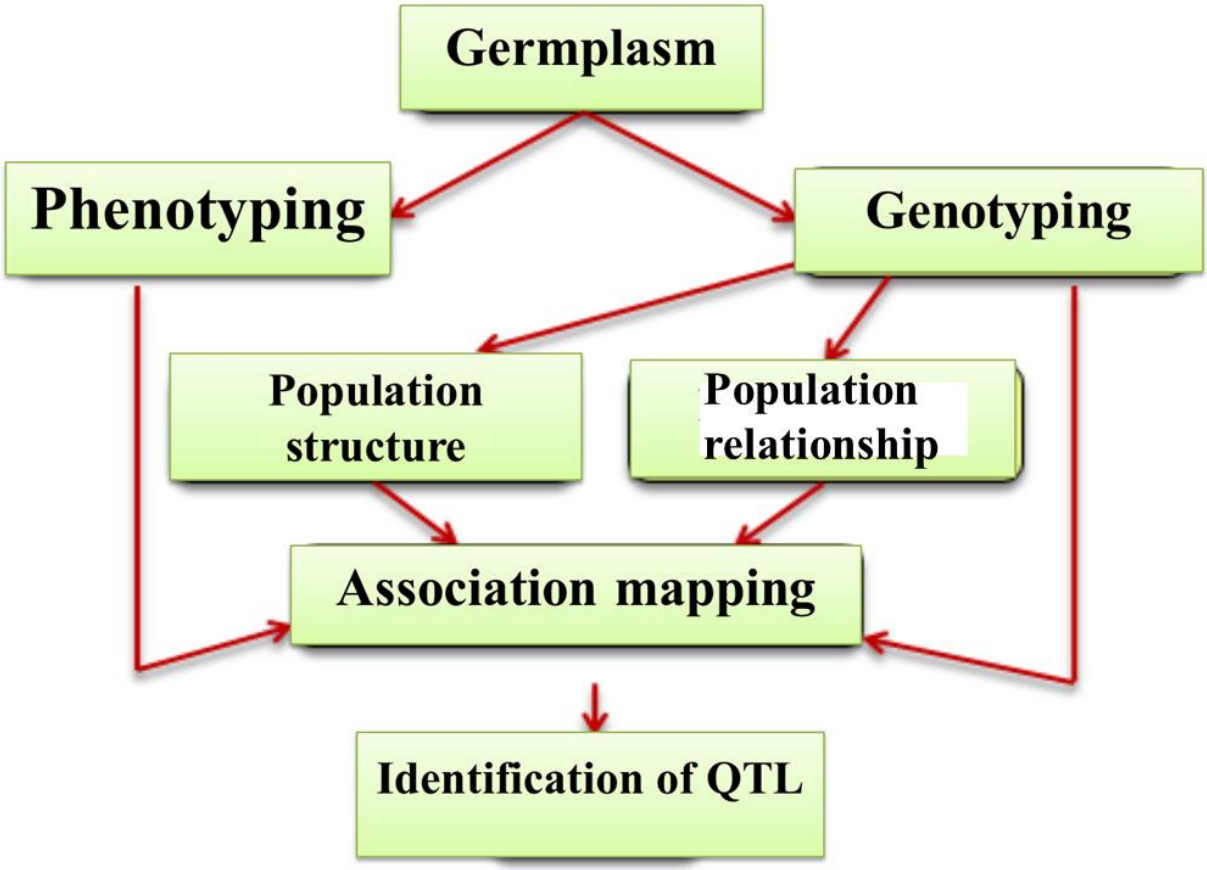


Fig. 13. Genetic mapping by Association Mapping approach.

In addition to information on the phenotype and genotype of each individual, a mixed model integrates information on population structure (P-matrix) and kernel relationships (K-matrix) is required to identify QTLs.

Up to now, various statistical models are available to perform these association mapping analyzes and the quality of the data obtained can be evaluated using a quantile-quantile diagram ("Q-Q plot", Fig. 14). Such a graph indicates the proportion of observed p-values that are less than or equal to different values of p between 0 and 1. The p-values referred to here are a measure of the significance (in the statistical sense) of the association between a marker and the measured phenotype. A marker with a significant association with the phenotype will be presumed to be close to a QTL that helps determine this phenotype.

If a model takes into account all the variables described above, we obtain a line of slope 1. In such a case, only 5% of the observed p-values would be <0.05, 10% would be <0.01. A general linear model (GLM) that does not take into account information on the structure of the population or kinship relations in the analysis often has poor performance (Pasam et al., 2012), such a model would deliver an excessive number (~ 50%) of values p <0.05. The vast majority of markers deemed "significant" would likely be "false positives". In contrast, the Q and P models all include population structure information, measured either with a Bayesian approach (Q-matrix) or a principal-component analysis (P-matrix), and they perform significantly better than the model Naive. Other models, such as the Q + K or P + K models, include information on both the population structure (matrix Q or P) as well as information on kinship relationships between individuals (matrix K). It is often these models which are the most efficient and whose distribution of the observed p-values suffers the least from inflation at the level of the low values of P.

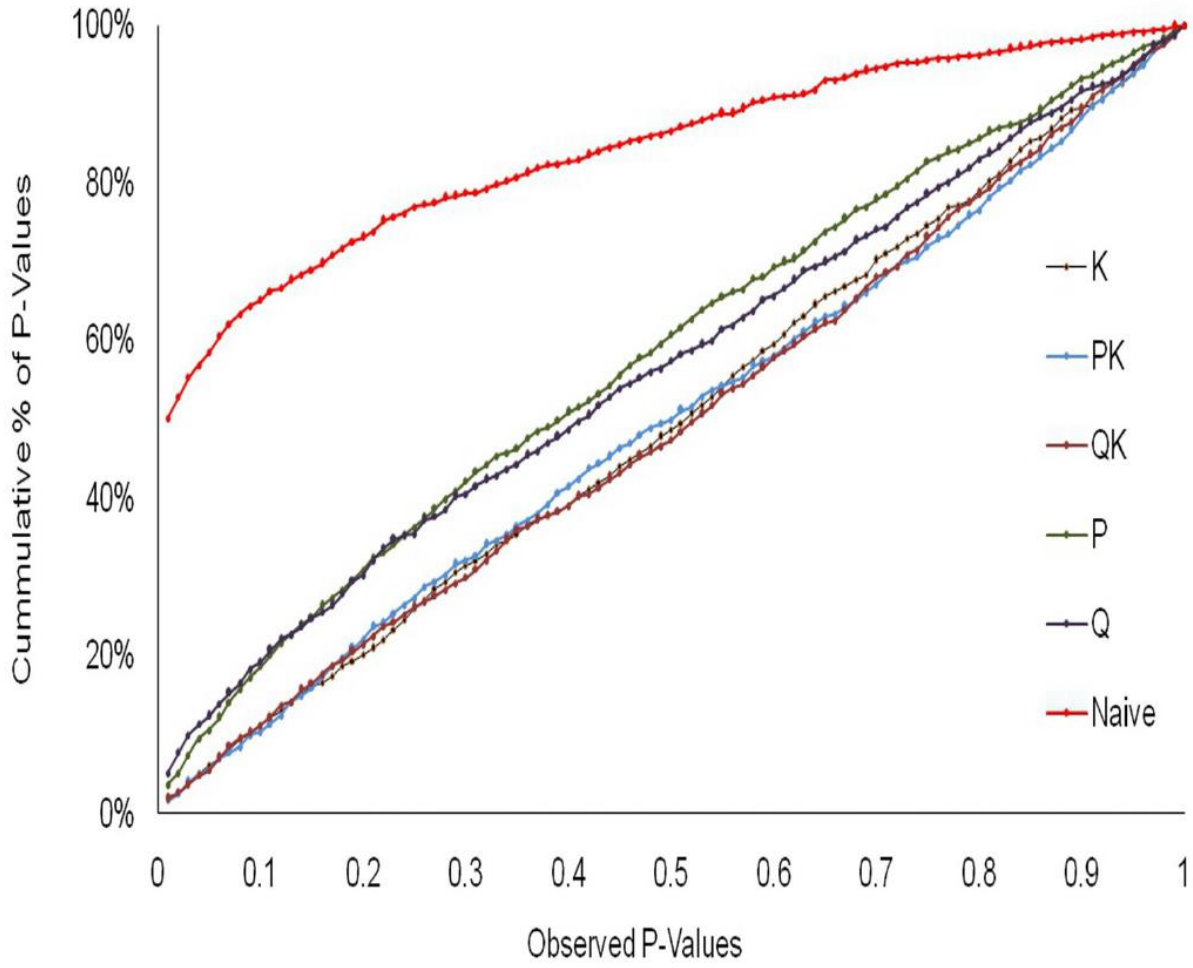


Fig. 14. Comparison of different mapping models by association analysis.

Different models presented (Fig. 14) are described as follow: "Naive" = the model does not take into account either the structure of the population or kinship relations; "P" = the model only takes into account the population structure obtained by the principal component analysis (matrix P); "Q" = the model takes into account the structure of the population obtained by Structure (matrix Q); "QK" = the model takes into account the structure of the population (matrix Q) as well as kinship relations (matrix K); "PK" = the model takes into account the structure of the population (matrix P) and kinship relations (matrix K). (Pasam et al., 2012).

Several software packages are available for Association Mapping analysis: TASSEL ("Analysis by Association, Evolution and Linkage", Bradbury et al., 2007) and GAPIT ("Genome Association and Prediction Integrated Tool", Lipka et al., 2012) are two commonly used software for Association Mapping in plants. These programs offer the possibility of using different regression models such as the general linear model (GLM) and mixed linear models (MLM) that can take into account the relationship between individuals or the structure of the population. TASSEL and GAPIT software's can also be used for calculating and graphing LD

statistics, for analyzing population structure using Principal Component Analysis (PCA) or for measuring kinship. ("Kinship") between the lines (Soto-Cerda and Cloutier, 2012, Lipka et al., 2012).

The mapping-by-association (Association Mapping) approach is advantageous because it potentially offers higher resolution. The availability of greater allelic diversity of germplasm provides greater allele coverage and time and cost savings over the resources used to establish bi-parental mapping populations. In addition, it potentially offers higher resolution due to the exploitation of more meiotic events throughout the history of germplasm development. Finally, the Association Mapping also offers the possibility of using historical phenotypic measurement data (Kraakman et al, 2004, Kraakman et al, 2006, Al-Maskri, 2012). On the other hand, AM also presents some challenges. The detection power of QTLs will be limited if there is not a strong LD between a neighboring marker and the QTL (Van Inghelandt et al., 2011). Incomplete LD between a marker and a QTL will lead to an underestimation of the variance explained by this QTL or the inability to detect it (Würschum, 2012). To maximize the LD between markers and QTL, it will be necessary to ensure a very dense coverage of the genome, which may require a very large number of markers.

I.2.8. Genetic mapping for grain size

The size and shape of grains, which are associated with yield and milling quality, are two of the most important characteristics of wheat domestication and selection (Breseghello and Sorrells 2007). Grain size is mainly characterized by grain weight and area, while grain shape is generally estimated by length, width, vertical perimeter, sphericity and proportion of horizontal axes (Breseghello and Sorrells 2007) . Modern wheat varieties have higher grain width and lower grain length than ancestral wheat varieties with greater variability in grain size and shape (Gegas et al. 2010).

Larger grains may have a positive effect on seedling vigor and increase yield (Gan and Stobbe 1996). Geometric models have shown that changes in grain size and shape can result in an increase in flour yield of up to 5% (Marshall et al., 1984). Many studies have identified quantitative trait loci (QTLs) for grain size and shape in common wheat cultivars and QTLs have been assigned to various chromosomes (Breseghello and Sorrells 2006, 2007, Gegas et al 2010, Sun et al., 2009, 2010, Tsilo et al., 2010, Williams et al., 2013). Okamoto et al. (2013) performed QTL analyzes for grain size and shape-related traits using four synthetic wheat F2 populations to identify the genetic loci responsible for grain size and shape variation in the hexaploid background. On the chromosome 2DS, QTL for grain length parameters were found

in the Ldn / PI476874 // Ldn / KU-2069 population. A QTL for grain length (qGL3), which encodes a putative protein phosphatase with a Kelch-like repeat domain (OsPPLK1), was found in the centromeric region of rice chromosome 3 (Qi et al., 2012). Recently, QTL of thousand grain weight has also been identified in distal long arms on rice chromosomes 5 and 12 (Marathi et al., 2012). Thus, genes or QTLs associated with grain shape and size are of interest for domestication and enhancement (Simons et al., 2006, Williams et al., 2013).

Many studies have revealed that in the genome-D of common wheat donor, *Ae. Tauschii* provides a great source of genetic variability and improves agronomic traits (Reynolds et al., 2007, Rana et al., 2013). As a result, many synthetic allohexaploid wheat lines were produced by hybridization of tetraploid wheat and *Ae. tauschii* and successfully used in wheat cultivation (Trethowan and Mujeeb-Kazi, 2008, Takumi et al., 2009). At the genomic level, several QTL analyzes were used to identify D-genomic regions of synthetic allohexaploid wheat lines associated with grain size and shape (Okamoto et al., 2013, Yu et al., 2014). Notably, Tg-D1 on the 2D chromosome is one of the well-known loci that have been recruited for the domestication of wheat grain size and shape. At the allohexaploid wheat speciation, a dramatic change in grain shape occurred due to the mutation in the Tg gluteal toughness gene (Nalam et al., 2007, Dvoraket al., 2012).

To reveal genetic factors on the D genome that control variation in grain size and shape, Yan et al. (2017) analyzed quantitative trait loci (QTL) using F2 and F2: 3 populations derived from a common allohexaploid wheat TAA10 and synthetic allohexaploid wheat XX329, which have almost identical AABB genomes and different DD genomes. Based on genotyping using 660K wheat single nucleotide polymorphism (SNP), a total of nine stable QTLs associated with grain size and shape were mapped to 2D and 7D chromosomes and verified using near isogenic lines (NILs) with allohexaploid synthetic wheat XX329 contributing favorable alleles. Notably, a new QTgw.cau-2D QTL controlling grain weight was first identified from synthetic allohexaploid wheat, which could be a more desirable target for genetic improvement of wheat. These authors identified the candidate genes for seed size (GIF1 and PFPβ), orthologs in rice. However, these results provide further information on the genetic factors that shaped grain morphology during the evolution and domestication of wheat.

In addition, recent advances in genomic technologies have allowed for a better understanding of the genetic basis of variation in large sets of genetic material using Genome wide association studies (GWAS). GWAS is one of those approaches that can be used for the identification and high-resolution mapping of genetic variability useful from sets of genetic material that have resulted from many historical recombination cycles (Yu and Buckler, 2006). Arora et al. (2017)

conducted GWAS in a collection of *Ae. Tauschi*i accessions for grain length, width, and weight using single-nucleotide polymorphic markers (SNPs) based on genotyping-by-sequencing (GBS). So, genetic similarity was calculated between accessions and GWAS was performed using 114 non-redundant accessions and 5249 SNP markers. A total of 17 SNPs associated with granulometric characteristics distributed over all seven chromosomes was revealed, with particularly the chromosomes 6D, 5D and 2D harboring the most important marker-trait associations. Most studies on germplasm of hexaploid wheat have focused on understanding the genetic and morphological diversity of this species. Few studies have used GWAS in hexaploid wheat for economically important traits such as grain yield and its components.

I.2.9. Impact of acid soils on wheat

I.2.9.1. Characteristics of acid soils

The tropical zone is generally covered with savannahs and tropical forests whose soil type is largely ferralitic. In addition, 80% of the total area of these forests is dominated by oxisols and utisols (Von Uexküll and Mutert, 1995). The homogeneous texture of the clay horizon is relatively constant and ranges between 50 and 70% of clay soils (Robain, 1993).

Lateritic soils consist of hardened materials and iron oxyhydroxides. They have a low cation exchange capacity (10 meq / 100g) and total cation reserves (Fig.15).

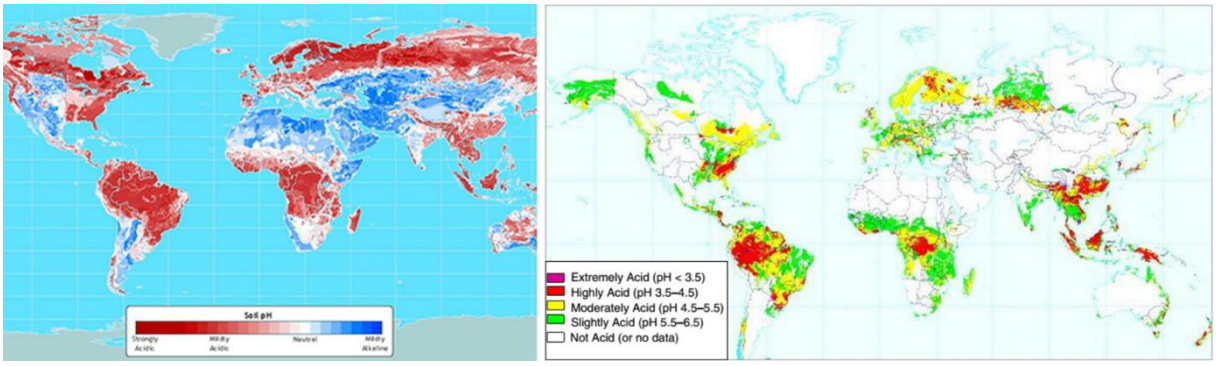


Fig. 15. Global distribution of acid soils (Von Uexkull et Mutert, 1995; http://websoilsurvey.nrcs.usda.gov/).

The soil is acidic when the quantity of free hydrogen ions (H +) is greater than the quantity of alkaline or alkaline-earth cations (Lozet and Mathieu, 1986). Acid soils, mainly found in the tropical and subtropical humid regions of the world, are generally characterized by Al^{3+}, Mn^{2+} and H^{+} , with deficiencies of Ca^{2+}, Mg^{2+} and $PO4^{3+}$ (Krstic et al., 2012), low pH (pH <5.5), low levels of organic matter, reduced cation exchange rate and excessive leaching of elements in areas with high rainfall. They may be due to the toxicity of Al, Fe and Mn, and / or they may

be caused by the absence or absence of N, P, Ca, Mo, Si and Mg, inhibiting root development and reduction. water absorption (Mossor-Pietraszewska, 2001, Velasquez et al., 2008). The soil acidity complex creates a chemical barrier that reduces root growth and development; these roots become ineffective at absorbing nutrients and water (Reynolds et al., 2001). Low availability of P and highly exchangeable Al inhibit root growth with concomitant losses in grain yield and quality (Von Baer, 2007).

Soil acidification is mainly caused by the nature of the bedrock. Harter (2007) has described other processes that contribute to increasing the amount of hydrogen ions (H^+) in soil solutions and creating more acidic conditions.

Acid soils are phytotoxic as a result of nutritional disorders: deficiency or unavailability of essential nutrients such as Ca, Mg, Mo and P, and toxicity of Al and Mn (Jayasundara et al., 1998). The solubility of soil compounds, and hence the availability of nutrients for plants, is related to soil pH (Fig. 16).

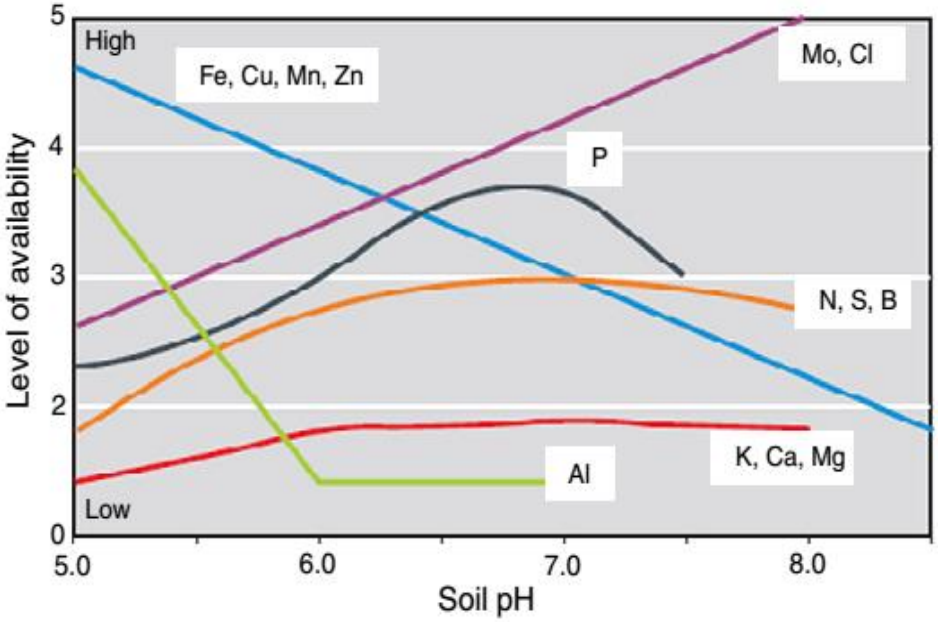


Fig.16. Relationship between element availability and soil pH (Goedert et al., 1997).

I.2.9.2. Impact of aluminum toxicity

In a neutral or basic environment, Al is insoluble and is in the form of monovalent and divalent oxides or aluminum hydroxide Al(OH)3, and is a normal part of any soil, useful for plants (Harter, 2007). In an acid environment, it is soluble and is in trivalent form with the Al^{3+} ion becoming the dominant species (Krill et al., 2010, Kochian et al., 1995). The Al^{3+} ions disrupt many physiological processes in plants by both apoplastic and symplastic interactions. Al delays the growth of the primary root and inhibits lateral root formation (Krstic et al., 2012).

This inhibition of root growth is the main symptom of toxicity in Al (Delhaize and Ryan, 1995, Tabuchi and Matsumoto, 200, Krill et al., 2010). The reduction in root growth and function leads to increased sensitivity to other constraints, mainly mineral deficiencies and drought, due to the limited capacity of Al's poisoned roots to acquire sufficient water and nutrients (Krill et al., 2010).

The common responses of plants to Al toxicity are mainly structural and ultra-cellular changes in the leaves, increased rates of diffusion resistance, reduced stomatal opening, reduced photosynthetic activity leading to leaf and size leaf chlorosis and necrosis; and decreased biomass (Thornton et al., 1986; Tekeu et al. 2015).

The roots of the plants subjected to Al toxicity have a spatula appearance, the truncated ends turning brown and the root system being entirely coronoid and having several truncated and filiform lateral roots, without shoots.

Al toxicity predisposes the plant root to fungal attack (Ota, 1968). In the roots, Al interferes with cell division, reduces DNA duplication by increasing the rigidity of the double helix (Rout et al., 2000). Al binds phosphorus in assimilated form, interferes with certain enzymes governing the deposition of polysaccharides in the cell membrane and increases their rigidity. Young plants are generally more affected by the toxicity of aluminum than older plants (Rout et al., 2000). Al toxicity is the most important factor limiting agricultural productivity on 67% of the total area of acid soils (Eswaran et al., 1997).

I.2.9.3. Identification of functional genes for aluminum tolerance

Al tolerance genes belong mainly to the MATE (multidrug extrusion and extrusion of toxic compounds) and ALMT (aluminum-activated malate transporter) families. MATE genes code for transporters excreting a wide range of metabolites and xenobiotics in eukaryotes and prokaryotes (Weston et al., 2012), and members of the ALMT family code for vacuolar malued channels (Kovermann et al., 2007).). In wheat, Al tolerance is mainly controlled by two genes: TaALMT1, which encodes a malate transporter on chromosome 4D, is constitutively expressed on the root apex (Sasaki et al., 2004). TaMATE1 would respond to Al stress based on citrate efflux (Ryan et al., 2009, Kochian et al., 2015). The general model illustrating the mechanisms of Al resistance (Al exclusion and Al tolerance / detoxification) that wheat plants use as Al^{3+} is taken up by the root from the soil solution and a portion of the absorbed Al. is transferred and stored in the shoot. At the top right, the exclusion of Al and the internal mechanisms of detoxification of Al (tolerance) in the root are based on the chelation of Al^{3+}, mainly by organic acid anions (OA), (Fig. 17).

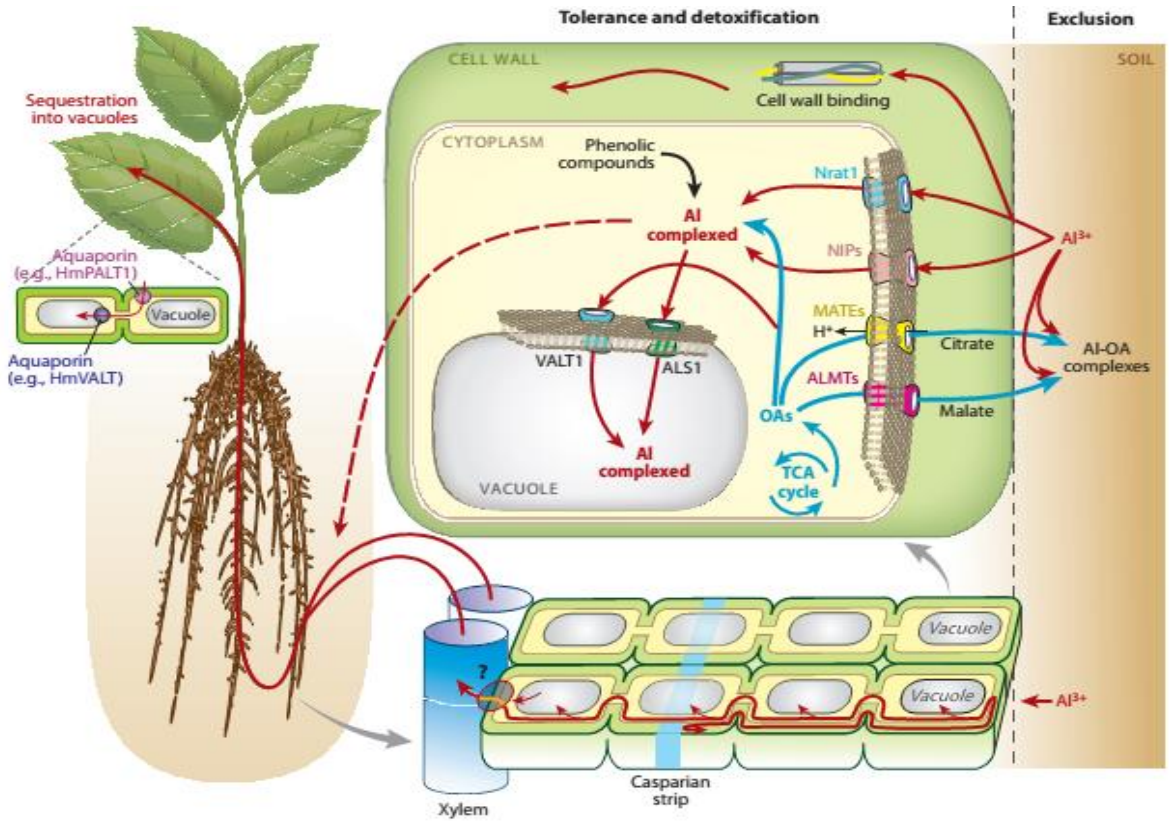


Fig.17. Methods illustrating Al exclusion, tolerance and detoxification mechanisms (Kochian et *al*., 2015).

I.2.9.4. Development of molecular markers for assisted selection

Different types of markers, including RFLPs, AFLPs, RAPDs, SSRs, DArTs and SNPs, have been developed and used in Al tolerance studies. With the identification and sequencing of genes for Al tolerance in plants, PCR-specific gene markers developed from genetic sequencing are preferred in marker-assisted selection (MAR) for easy identification, polymorphism and their good reproducibility (Poczai et al., 2013). An RFLP marker bcd1230 co-segregated with a major Al tolerance gene on wheat chromosome 4DL accounted for 85% of the phenotypic variation in Al tolerance (Riede and Anderson, 1996).

Miftahudin et al. (2002) found that 5 AFLP markers (AMAL1, AMAL2, AMAL3, AMAL4 and AMAL5) were closely related to, and flanked Alt3 on the long arm of chromosome 4R. After screening 35 accessions of Al-tolerant wheat varieties using 10 combinations of AFLP primers, Stodart et al. (2007) found that these accessions had different genetic backgrounds and were therefore valuable germplasm for Al tolerance selection. Ma et al. (2005) reported SSR markers (Xwmc331 and Xgdm125) flanking the ALMT locus and indicated that these markers could be used for MAS in the selection of Al tolerant wheat cultivars.

Raman et al. (2006) developed SSR, ALMT1-SSR3a and ALMT1-SSR3b markers and a CAPS marker from repetitive InDels and the TaALMT1 gene substitution region. These PCR-based markers segregating with the tolerance locus should be effective tools for MAS (Raman et al., 2008). Mantovani et al. (2008) used DArT markers for Al screening in wheat. It is particularly notable for its broadband, speed, high reproducibility and low cost (Semagn et al., 2006). Hundreds to thousands of polymorphisms can be detected very quickly (Sansaloni et al., 2010).

I.2.9.5. QTL mapping and inheritance of Al tolerance

In 2006, Ma et al. (2006) evaluated a set of ditelosomal lines derived from the moderately Al-resistant wheat cultivar, Chinese Spring for Al resistance. A genetic linkage map consisting of 381 AFLP markers and 168 SSR markers was constructed to determine the genetic effect of quantitative trait loci (QTL) for Al resistance in Spring Chinese. Three QTLs (Qalt.pser-4D, Qalt.pser-5A and Qalt.pser-2D) have been identified as enhancing Al stress root growth, suggesting that inheritance of Al resistance in Chinese Spring is polygenic. The QTL with the greatest effect was flanked by markers Xcfd23 and Xwmc331 on chromosome 4DL and most likely multi-allelic with major QTL identified in Atlas 66. Two others (QTL, Qalt.pser-5A and Qalt.pser-2D) on 5AS and 2DL chromosomes, respectively, were also detected with marginal significance in the population.

Genetic mapping refers to the mapping of genes / loci to specific chromosomal locations using linked genetic markers (Semagn et al., 2006). A single dominant gene is responsible for Al tolerance in wheat (Delhaize et al., 1993); However, there are exceptions in some cultivars (Tang et al. 2002). Using different populations, genes / locus for Al tolerance were mapped to different wheat chromosomes. Unique loci for Al tolerance were identified on 4DL, 4D, 4BL or 3BL chromosomes, with phenotypic contributions of 85% (4DL locus), 50% (4D), 50% (4BL) and 49% % (3BL) (Ryan et al., 2009, Riede and Anderson, 1996, Ma et al., 2005, Navakode et al., 2010). In addition, the genes / loci on the 6AL, 7AS, 2DL, 5AS, 3DL and 7D chromosomes play a role in Al tolerance in wheat (Aniol and Gustafon, 1984, Aniol, 1990). The complex inheritance of Al tolerance has been found in wheat. Zhou et al. (2007) identified a secondary QTL for Al resistance on chromosome 3BL in Atlas 66, which was effective only when the epistatic gene on 4DL was absent. Cai et al. (2008) mapped three QTLs responsible for Al tolerance on wheat chromosomes 4DL, 3BL and 2A, which together accounted for 80% of the phenotypic variation.

I.2.9.6. Association Mapping Analysis for Al tolerance

According to Zhu et al. (2008), association mapping is based on associations between molecular markers and traits that can be attributed to the strength of linkage disequilibrium in large non-crossing populations. It differs from bi-parental QTL mapping which evaluates only two alleles. Associative mapping allows multiple alleles to be evaluated simultaneously and is useful for studying the inheritance of complex traits controlled by multiple QTLs (Buckler and Thomsberry, 2002). Ryan et al. (2009) used DArT markers to perform whole genome mapping in some Brazilian wheat cultivars validating the citrate efflux mechanism for Al tolerance.

Raman et al. (2010) used association mapping to identify genomic regions associated with Al3 + resistance using 1055 common wheat accessions from different geographic regions of the world and 178 DArT polymorphic markers. Genome-wide association analyzes detected markers significantly associated with Al^{3+} resistance on chromosomes 1A, 1B, 2A, 2B, 2D, 3A, 3B, 4A, 4B, 4D, 5B, 6A, 6B, 7A and 7B. Some of these genomic regions correspond to previously identified loci for Al3 + resistance, while others appear to be new.

Dai et al. (2013) reported a large genome association study with a limited number of regions identified by markers involved in wheat A1 resistance, including the TaALMT1-harboring region on chromosome 4DL, a previously QTL on chromosome 3B.

Navakode et al., (2014) reported the results of a genome-wide association mapping approach in wheat using 525 genotyped DArt markers in a core collection of 96 winter wheat accessions. Marker character associations (MTA) were detected using both the general linear model (GLM) and the mixed linear model (MLM). Five significant MTAs were identified as being shared by models on chromosomes 1A, 1D, 3B and 6A. MLTs have been identified by MLM on chromosomes 1D and 3B which could be new candidate loci for future studies.

In 2015, Kochian et al. (2015) suggest that the time has come to explore the potential for improving wheat Al resistance beyond the use of TaALMT1.

Recently, Froese and Carter (2016) genotyped two diverse populations of adapted northwestern US winter wheat (PNW) on 9K and 90K SNP markers developed and phenotyped in three low pH and Al toxicity environments. population included 459 accessions of soft white wheat and the second, which was phenotyped by hydroponics, included 401 accessions (soft white and some hard red). Genome-Wide Association Studies (GWAS) revealed a total of 55 loci, of which 15 were common to both populations. The molecular marker wmc331, linked to ALMT1, was analyzed in both populations, revealing only eight individuals with the favorable allele. They found that the unique germplasm examined in their study has new sources of low pH tolerance.

CHAPTER II: MATERIALS AND METHODS

II.1. Genetic diversity of Cameroonian bread wheat (*Triticum aestivum* L.) cultivars revealed by microsatellite markers

II.1.1. Plant material

The plant material consists of 17 cultivars of bread wheat (*Triticum aestivum* L.) collected in six villages located in the North-west (A), North (B) and Adamawa (C) Regions of Cameroon (Fig. 18). Among them, 11 accessions were collected in five villages of Northwest and 6 were collected from one village (Wassande) of Adamawa region. The cultivars of the North West are mainly local seeds, whereas those of Adamawa were originally given by the SODEBLE and some others were imported from Tchad. In our study, we collected all materials used by farmers in those regions (Table I).

Fig.18. Interview and collection of wheat cultivars from Cameroonian farmers in North-west (A), North (B) and Adamawa (C).

II.1.2. Genomic DNA extraction

An adjusted Doyle and Doyle (1990) protocol was used to extract genomic DNA (gDNA) from seedlings at the two to three leaf stage. Approximately 100mg of plant tissue were cut up and placed into a micro centrifuge tube. 500 µl of 2 % (m/v) CTAB extraction buffer [1.4 M NaCl, 20 mM Na2EDTA (pH 8), 100 mM Tris-HCl (pH 8)] and tow sterilized steel bearings were added to each sample. A Qiagen®TissueLyser (Qiagen (Pty) Ltd; local distributer: Southern Cross Biotechnology, Claremont, RSA) was used to grind the samples 2 times for 2 minutes

(min) at 30Hz. This mixture was then incubated in a water bath for 20 min at 60 ºC. 200 µl and 50 µl of chloroform:isoamyl-alcohol (C:I::24:1) were added and the solution was centrifuged for 5 min at 12000 rpm. The supernatant was transferred to a clean centrifuge tube and another 250µl of chloroform:isoamyl-alcohol was added, followed by centrifugation for 5 min at 12000 rpm. The supernatant was once again transferred to a clean centrifuge tube. 50 µl of 3M Sodium acetate (pH 5.0) was added, followed by 500µl of ice cold 100 % ethanol. The tubes were carefully inverted and the gDNA was precipitated. After incubation, the pellet underwent two wash steps with 500µl of 70% ethanol. Then, the tubes were centrifuged for 5 min at 12000 rpm. The supernatant was discarded and the pellet was left to air dry. The pellet was finally re-suspended in 30µl of DNase/RNase-free water and incubated at 60 ºC for 2 min. The extracted gDNA was quantified using a Nanodrop® ND-1000 spectrophotometer. The DNA was diluted with DNase/RNase-free water to a concentration of 100ng/µl and stored at the fridge (- 20ºC).

Table I. Wheat cultivars used and their origins in Cameroon

No	Samples' names	Local name	Village	Region
1	Ngm 2	Ngm 2	Wassande	Adamawa
2	Fuanb2	Fuanb2	Fuanentui	Northwest
3	Babankit	Babankit	Smal Babanki	Northwest
4	Alexander wonder	Alexander wonder	Boyo	Northwest
5	Fuanb1	Fuanb1	Fuanentui	Northwest
6	Sonalika	Sonalika	Wassande	Adamawa t
7	Fuanb3		Fuanentui	Northwest
8	Fuanb4		Fuanentui	Northwest
9	HGW	Hard wheat	Abongphen	Northwest
10	BBT2		Abongphen	Northwest
11	WASSANDE 2	WASSANDE 2	Wassande	Adamawa
12	Vrack	Vrack	Bambui	Northwest
13	Ngderem4		Wassande	Adamawa
14	Ngderem1	Ngderem1	Wassande	Adamawa
15	Ngderem3		Wassande	Adamawa
16	IRAT 10	IRAT 10	Bambui	Northwest
17	RIBA	RIBA	Boyo	Northwest

II.1.3. Microsatellite markers and PCR amplification

Eleven wheat microsatellite markers for 11 loci located in the chromosomes 1A, 2A, 2D, 3A, 3B, 4D, 5D, 6B and 7D, were used for genetic diversity analysis. Xgwm and Xwmc markers were obtained respectively from Röder *et al*. (1998) and Somers and Isaac (2004; Grain Genes). PCR reactions were carried out in 14µl reaction mixture of KAPA2GTM Fast Multiplex PCR Mix, 6.25 µM of each forward and reverse primer, 1 µl gDNA and dH_2O. The PCR cycling

conditions was set at 94 ºC for 3 min of denaturation, followed by 45 cycles of 1 min at 94 ºC, 1 min at the annealing temperature (Ta), 2 min at 72 ºC and then 72ºC for 10 min for extension. The PCR products were electrophoresed on 6 % denaturing polyacrylamide gels containing 1xTBE (Tris Borate EDTA). The amplified band sizes for each SSR locus were determined on the basis of their migration relative to the 50 bp marker.

II.1.4. Data analysis

The molecular diversity within all accessions was estimated for each SSR locus, using the PowerMarker 3.25 software (Liu and Muse 2005). To measure the informative character of the SSR markers, the Polymorphism Information Contents (PIC) for each marker was calculated using the formula of Nei (1973): $PIC = 1-\Sigma_{i=1}^{k} P_i^2$, where k is the total number of alleles detected per locus and Pi the frequency of the allele i in all 17 accessions.

Genetic similarity (GS; Dice 1945) was calculated as: $GS = 2N_{ij}/(N_i + N_j)$ where N_{ij} is the number of fragment common to individual i and j, and (Ni + Nj) is the total number of fragment in both individuals.

Genetic distance (GD) among group pairs was calculated following Nei and Li (1979), $(GD_{xy}) = 1-(2N_{xy}/N_x + N_y)$. The dendrogram was constructed using the method based on the genetic distance (SAHN method, UPGMA algorithm) of the 17 accessions using the software Statistica 12. To calculate allele frequency (A_{xy}) from one of variation to another in each locus, the formula of Khlestkina et al. (2004) was used: $A_{xy} = \Sigma IP_{xi} – P_{yi}IN_{xy}$, where P_{xi} and P_{yi} are the frequencies of the i^{th} allele in regions X and Y, respectively, and N_{xy} is the total number of alleles for the two groups X and Y. The allelic frequency variation was calculated separately for each of the 11 loci and then for all of them as an average. All fragments were used to generate a GS matrix for Principal Component Analysis (Sneath and Sokal 1973).

II.2. Population structure in a global accession of hexaploid wheat (*Triticum aestivum* L.) breeding lines using Genotyping by Sequencing approach

II.2.1. Plant material

An international collection of 288 wheat lines was used in this study (Table II). A total of 179 lines representing almost all the diversity present in Central Africa (Cameroon), North Africa, South Africa, East Africa (Kenya, Ethiopia), France, Mexico (Mexicali, Baja California) was phenotyped, followed by genotyping, using the Genotyping-by-Sequencing approach. Then, a sub collection of 108 Canadian lines (Quebec, Ontario) used for controls standard (out-group and accuracy verification), were genotype using 90K SNP Array Wheat approach.

Table II: List of cultivars and their origins

No	Name of accessions	Origin's zones	Origin's country
1	FUANB2	Fuanentui	Cameroon
2	RIBA	Boyo	Cameroon
3	ALEXANDERWONDER	Boyo	Cameroon
4	SONALIKA	Wassande	Cameroon
5	FUANB3	Fuanentui	Cameroon
6	HGW	Abongphen	Cameroon
7	VRACK	Bambui	Cameroon
8	IRAT10	Bambui	Cameroon
9	Kenya1	Baja California	Mexico
10	Kenya4	Baja California	Mexico
11	Nd643-2	Baja California	Mexico
12	Babax9	Baja California	Mexico
13	Mercato2	Baja California	Mexico
14	NGL	Baja California	Mexico
15	Babax2	Baja California	Mexico
16	Kenya2	Baja California	Mexico
17	Nd643-4	Baja California	Mexico
18	Babax5	Baja California	Mexico
19	Babax10	Baja California	Mexico
20	Mercato1	Baja California	Mexico
21	Babax16	Baja California	Mexico
22	Mutus1	Baja California	Mexico
23	Nd643-1	Baja California	Mexico
24	Babax12	Baja California	Mexico
25	Babax11	Baja California	Mexico
26	Premio2	Baja California	Mexico
27	Nd643-5	Baja California	Mexico
28	Mercato3	Baja California	Mexico
29	Waxwing1	Baja California	Mexico
30	Babax14	Baja California	Mexico
31	Pfau1	Baja California	Mexico
32	Attila4	Baja California	Mexico
33	Babax1	Baja California	Mexico
34	Premio4	Baja California	Mexico
35	Waxwing2	Baja California	Mexico
36	Babax6	Baja California	Mexico
37	Pfau2	Baja California	Mexico
38	Babax15	Baja California	Mexico
39	Sup152-1	Baja California	Mexico
40	Premio3	Baja California	Mexico
41	KronstadF2004-1	Baja California	Mexico
42	Babax7	Baja California	Mexico

43	Mino898	Baja California	Mexico
44	Babax17	Baja California	Mexico
45	Pfau3	Baja California	Mexico
46	Wbll3	Baja California	Mexico
47	Croc_1	Baja California	Mexico
48	Babax13	Baja California	Mexico
49	Ceta	Baja California	Mexico
50	Babax18	Baja California	Mexico
51	Pfau4	Baja California	Mexico
52	Pfunye1	Baja California	Mexico
53	Nd643-3	Baja California	Mexico
54	Babax8	Baja California	Mexico
55	Yuk	Baja California	Mexico
56	Premio1	Baja California	Mexico
57	ADAGIO	Baja California	Mexico
58	misr1	/	Egypt
59	misr2	/	Egypt
60	danda`a	/	Egypt
61	kakaba	/	Ethiopia
62	hidase	/	Ethiopia
63	Bermude	/	France
64	Apache	/	France
65	w11	Ontario	Canada
66	w52	Ontario	Canada
67	w145	Ontario	Canada
68	w160	Ontario	Canada
69	w243	Ontario	Canada
70	w269	Ontario	Canada
71	w282	Ontario	Canada
72	w342	Ontario	Canada
73	w416	Ontario	Canada
74	w473	Ontario	Canada
75	w12	Ontario	Canada
76	w55	Ontario	Canada
77	w149	Ontario	Canada
78	w163	Ontario	Canada
79	w244	Ontario	Canada
80	w270	Ontario	Canada
81	w283	Ontario	Canada
82	w346	Ontario	Canada
83	w431	Ontario	Canada
84	w16	Ontario	Canada
85	w56	Ontario	Canada
86	w150	Ontario	Canada

87	w167	Ontario	Canada
88	w253	Ontario	Canada
89	w271	Ontario	Canada
90	w286	Ontario	Canada
91	w347	Ontario	Canada
92	w445	Ontario	Canada
93	w18	Ontario	Canada
94	w127	Ontario	Canada
95	w151	Ontario	Canada
96	w210	Ontario	Canada
97	w255	Ontario	Canada
98	w274	Ontario	Canada
99	w287	Ontario	Canada
100	w348	Ontario	Canada
101	w446	Ontario	Canada
102	w20	Ontario	Canada
103	w128	Ontario	Canada
104	w153	Ontario	Canada
105	w218	Ontario	Canada
106	w256	Ontario	Canada
107	w275	Ontario	Canada
108	w288	Ontario	Canada
109	w349	Ontario	Canada
110	w447	Ontario	Canada
111	w29	Ontario	Canada
112	w131	Ontario	Canada
113	w155	Ontario	Canada
114	w229	Ontario	Canada
115	w259	Ontario	Canada
116	w276	Ontario	Canada
117	w291	Ontario	Canada
118	w350	Ontario	Canada
119	w467	Ontario	Canada
120	w30	Ontario	Canada
121	w134	Ontario	Canada
122	w156	Ontario	Canada
123	w236	Ontario	Canada
124	w266	Ontario	Canada
125	w277	Ontario	Canada
126	w340	Ontario	Canada
127	w405	Ontario	Canada
128	w470	Ontario	Canada
129	w51	Ontario	Canada
130	w140	Ontario	Canada

131	w157	Ontario	Canada
132	w239	Ontario	Canada
133	w268	Ontario	Canada
134	w280	Ontario	Canada
135	w341	Ontario	Canada
136	w408	Ontario	Canada
137	w471	Ontario	Canada
138	Attila3	/	Egypt
139	Katila15	/	Egypt
140	ZAKIA-5(CHECK-4)	/	Egypt
141	Star3	/	Egypt
142	Attila8	/	Egypt
143	Katila17	/	Egypt
144	REYNA-8	/	Egypt
145	Vee	/	Egypt
146	Cham-6	/	Egypt
147	Debeira(check)	/	Egypt
148	Seri1B-1	/	Egypt
149	WATAN-7-SEKHRAH-2	/	Egypt
150	Debeira2	/	Egypt
151	Katila8	/	Egypt
152	Seri1B-2	/	Egypt
153	NEJMAH-12(CHECK-6)	/	Egypt
154	Crows	/	Egypt
155	Attila5	/	Egypt
156	GOUMRIA-3(CHECK-2)	/	Egypt
157	Kauz	/	Egypt
158	Seri1B-3	/	Egypt
159	Attila6	/	Egypt
160	Florkwa-2	/	Egypt
161	P1871	/	Egypt
162	Seri1B-4	/	Egypt
163	Attila1	/	Egypt
164	HIJLEEJ-1	/	Egypt
165	REYNA-25	/	Egypt
166	HAALA-34(CHECK-5)	/	Egypt
167	Attila7	/	Egypt
168	Hubara-2	/	Egypt
169	REYNA-29	/	Egypt
170	Seri1B-5	/	Egypt
171	Kenyawren	/	Kenya
172	robin,ken	/	Kenya
173	Kenyasunbird	/	Kenya
174	Kenyatae	/	Kenya

175	Kenyaeagle	/	Kenya
176	Kenyakingbird	/	Kenya
177	Mutus1	Mexicali	Mexico
178	Kachu4	Mexicali	Mexico
179	Pastor1	Mexicali	Mexico
180	Kachu7	Mexicali	Mexico
181	Trap	Mexicali	Mexico
182	Sup152-2	Mexicali	Mexico
183	Kiskadee	Mexicali	Mexico
184	Kachu5	Mexicali	Mexico
185	Kachu8	Mexicali	Mexico
186	BAJ#1	Mexicali	Mexico
187	Trch	Mexicali	Mexico
188	Sokoll2	Mexicali	Mexico
189	Pauraq	Mexicali	Mexico
190	Chibia	Mexicali	Mexico
191	Danphe	Mexicali	Mexico
192	Wbll2	Mexicali	Mexico
193	Pfau5	Mexicali	Mexico
194	Sokoll1	Mexicali	Mexico
195	Sha7	Mexicali	Mexico
196	Pastor2	Mexicali	Mexico
197	PRL2	Mexicali	Mexico
198	Rolf07-2	Mexicali	Mexico
199	Rolf07-1	Mexicali	Mexico
200	PRL1	Mexicali	Mexico
201	PRL3	Mexicali	Mexico
202	Whear2	Mexicali	Mexico
203	PRL4	Mexicali	Mexico
204	Kachu9	Mexicali	Mexico
205	Kachu3	Mexicali	Mexico
206	Kachu1	Mexicali	Mexico
207	Attila2	Mexicali	Mexico
208	Babax4	Mexicali	Mexico
209	Kachu6	Mexicali	Mexico
210	Kachu11	Mexicali	Mexico
211	Kachu2	Mexicali	Mexico
212	Egypte3	Mexicali	Mexico
213	Chwink1	Mexicali	Mexico
214	Babax3	Mexicali	Mexico
215	Melon1	Mexicali	Mexico
216	Kachu10	Mexicali	Mexico
217	Hahn	Mexicali	Mexico
218	Kenya11	Mexicali	Mexico

219	Whear1	Mexicali	Mexico
220	Sokoll3	Mexicali	Mexico
221	Pbw343	Mexicali	Mexico
222	Wbll1	Mexicali	Mexico
223	Melon2	Mexicali	Mexico
224	Moka	Quebec	Canada
225	BS13-134	Quebec	Canada
226	Quebec29	Quebec	Canada
227	CRGB-13A.0202	Quebec	Canada
228	Migantic	Quebec	Canada
229	AACScota	Quebec	Canada
230	Topaze	Quebec	Canada
231	CFB1604	Quebec	Canada
232	Magog	Quebec	Canada
233	Dakosta	Quebec	Canada
234	Bangor	Quebec	Canada
235	12NQW-237	Quebec	Canada
236	CFB1605	Quebec	Canada
237	AW823	Quebec	Canada
238	Fuzion	Quebec	Canada
239	CFB1601	Quebec	Canada
240	CFB1511	Quebec	Canada
241	CFB1509	Quebec	Canada
242	BS12-218	Quebec	Canada
243	CFB1607	Quebec	Canada
244	CFB1603	Quebec	Canada
245	BS13-20,66	Quebec	Canada
246	Pasteur	Quebec	Canada
247	CFB1606	Quebec	Canada
248	CFB1602	Quebec	Canada
249	12NQW-1018	Quebec	Canada
250	Pokona	Quebec	Canada
251	SSBlomidon	Quebec	Canada
252	ECO448.1-38	Quebec	Canada
253	12NQW-488	Quebec	Canada
254	RGTPresidio	Quebec	Canada
255	Orlians	Quebec	Canada
256	CFB1510	Quebec	Canada
257	ERA-149R1	Quebec	Canada
258	Klios	Quebec	Canada
259	Furano	Quebec	Canada
260	Tugela-DN	/	South Africa
261	Chinesespring	/	South Africa
262	Betta	/	South Africa

263	Atlas66	/	South Africa
264	GAMTOUS	Capetown	South Africa
265	SST047	Capetown	South Africa
266	PALMIET	Capetown	South Africa
267	SST015	Capetown	South Africa
268	SST0127	Capetown	South Africa
269	SST056	Capetown	South Africa
270	SST88	Capetown	South Africa
271	SST087	Capetown	South Africa
272	SST027	Capetown	South Africa
273	SST347	Winter areas	South Africa
274	SST3149	Winter areas	South Africa
275	SST387	Winter areas	South Africa
276	SST316	Winter areas	South Africa
277	SST3156	Winter areas	South Africa
278	SST374	Winter areas	South Africa
279	SST356	Winter areas	South Africa
280	SST867	Spring areas	South Africa
281	SST866	Spring areas	South Africa
282	SST884	Spring areas	South Africa
283	SST806	Spring areas	South Africa
284	SST835	Spring areas	South Africa
285	SST877	Spring areas	South Africa
286	SST843	Spring areas	South Africa
287	SST875	Spring areas	South Africa
288	SST895	Spring areas	South Africa

II.2.2. Agro-morphological characterization

II.2.2.1. Description of the experimental site

Mbankolo is located at an altitude of 1069m above the sea, with an average temperature of 18-20 ° C, bimodal rainfall with an annual average of 1600 mm. On a rock, but very light and ferralitic. The climate is temperate sub-equatorial with four seasons. The vegetation is poor and mostly herbaceous, with some trees. In Mbankolo (Centre, Cameroon), soil analyses indicate appreciable properties (pH = 5.7, exchangeable acidity ($Al^{3+} + H^{+}$) = 0 cmol/kg) (Table III).
In addition, preliminary tests have shown that local varieties produce up to 0.5 t / ha in this zone (Ngo ngom, 2017). These favorable agro-climatic conditions for the cultivation of wheat seem to be important to justify the establishment of the trial of characterization of wheat varieties in this region of Cameroon (Mbankolo).

Table III: Physico-chemical composition of soils in the Mbankolo area, Cameroon

	Elements	Values
	Humidity %	**2.459**
organic material	**organic matter (g / kg)**	**35.524**
	Carbon (g / kg)	**20.606**
	Nitrogen (g / kg)	**1.980**
	C/N	**10.407**
exchangeable bases	**Ca^{2+} (cmol/kg)**	**1.207**
	Mg^{2+} (cmol/kg)	**1.371**
	K^{+} (cmol/kg)	**0.277**
	Na^{+} (cmol/kg)	**0.012**
	Som (cmol/kg)	**2.867**
	CEC (cmol/kg)	**9.953**
Acidity / alkalinity	**pH (H_2O) – 1 :5**	**5.70**
mechanical analysis	**Clay (in%)**	**23.25**
	Limon (in%)	**7**
	coarse silt (in%)	**3.96**
	Fine sand (in%)	**19.61**
	coarse sand (in%)	**45.30**
Phosphoric acid	**P total (mg/kg)**	**421.051**
exchangeable acidity	**Al^{3+} (cmol/kg)**	**0**
	Al^{3+} +H^{+} (cmol/kg)	**0**
Metals	**Zn (mg/kg)**	**114.379**
	Cu (mg/kg)	**37.285**
	Mn (mg/kg)	**89.976**
	Fe (mg/kg)	**23.663**

II.2.2.2. Experimental design

The experimental design is done in alpha-latice, with two repetitions (Fig. 19). In each block, (6 rows of 3 m for each variety), the wheat varieties were sown in completely randomized blocks (Rekha et al., 2014). The cultural practices recommended for characterization included

NPK fertilizer (15-15-15) as a base application at 200 kg / ha during soil preparation, and urea applied at 50 kg / ha, as a fraction greater than 14 days after sowing (at the first weeding) and at initiation of panicles (according to the physico-chemical composition of soils 1.980 g/kg of nitrogen (N); 10.407 g/kg of C/N, 421.0551 mg/kg of phosphoric acid (P total), 0.277 cmol/kg of potassium (K)). Cultivars were recorded for physio and agro-morphological traits during 2015/2016, and 30 traits or descriptors of wheat DUS (Distinctness, Uniformity and Stability) were collected on all cultivars.

Fig. 19. Experimental site of wheat breeding lines in Mbankolo (Yaounde, Cameroon)

II.2.2.3. Statistical analyses

The data collected were subject to multivariate analyses. The construction of the phylogenetic tree was done using software R. The Ward.D method was used for calculating Euclidean distance. The principal component analysis (PCA) was performed using the SPAD V5 software.

II.2.3. Genotyping-by-Sequencing

II.2.3.1. DNA extraction

Genomic DNA was extracted from 100 mg fresh young leaves using the DNeasy 96 Plant kit (Qiagen, cat. no. 69181) following the manufacturer‘s protocol. Steps are following: In each

tube containing a dry leaflet of wheat, two tungsten grinding balls made it possible to grind automatically about 8 times (frequency: 24 oscillations / s, duration: 30 seconds), making inversions with the plates between each milling: In order to isolate the genomic DNA, the lysis buffer left at 65 °C was stirred well, then 400 µl into each of the tubes. Invert 20X and incubate in the water bath at 65 °C for 30 min, inverting again every 10 min. Making a spin down [15 sec.] to bring down the liquid. Under the hood, add a volume of 400 µl of chloroform: isoamyl alcohol [24: 1]. Reverse 40X. Stay under the hood because the liquid may leak at this stage. Centrifuge 10 min at 6000 rpm in the plate centrifuge using the lab support (due to chloroform which may damage the plastic). Recover 200 µl of supernatant (warning: slope interface, must not be touched) and transfer to new well identified tubes. Add 0.8 X the volume of supernatant isopropanol [160 µl]. Invert 20X and centrifuge for 30 min at 6000 rpm. Remove the supernatant to keep the pellet and then drain on brown paper to remove the maximum amount of isopropanol. Making sure the caps are at the bottom of the tubes. Put 500 µl of 70% cold ethanol [-20 °C] and centrifuge for 30 minutes at 6000 rpm. Reverse quickly to get rid of the ethanol and then keep the pellet of DNA. Drain on brown paper to remove as much ethanol as possible. Centrifuge again briefly and then remove the rest of the liquid with a pipette. Put under the bell empty for 10 minutes (or more if necessary). Alternatively, leaving the oven at 40 °C for 10-15 min or until complete drying has also given good results. Ensure that all ethanol is evaporated as an ethanol residue may interfere with the subsequent activity of the enzymes. Suspend the DNA pellet in 50 µl of EB buffer. Let stand overnight in the refrigerator at 4 °C and re-suspend the genomic DNA pellet in the solution buffer.

II.2.3.2. Library preparation and sequencing

DNA was quantified using Quant-iT™ PicoGreen as well as Thermo Scientific Nanodrop 8000 spectrophotometer instrument (Fisher Scientific). DNA concentrations were normalized to 10 ng/µl and subsequently used for library preparation. Sequencing libraries were prepared according to the GBS protocol (Fig. 20) described by Elshire et al. (2011).

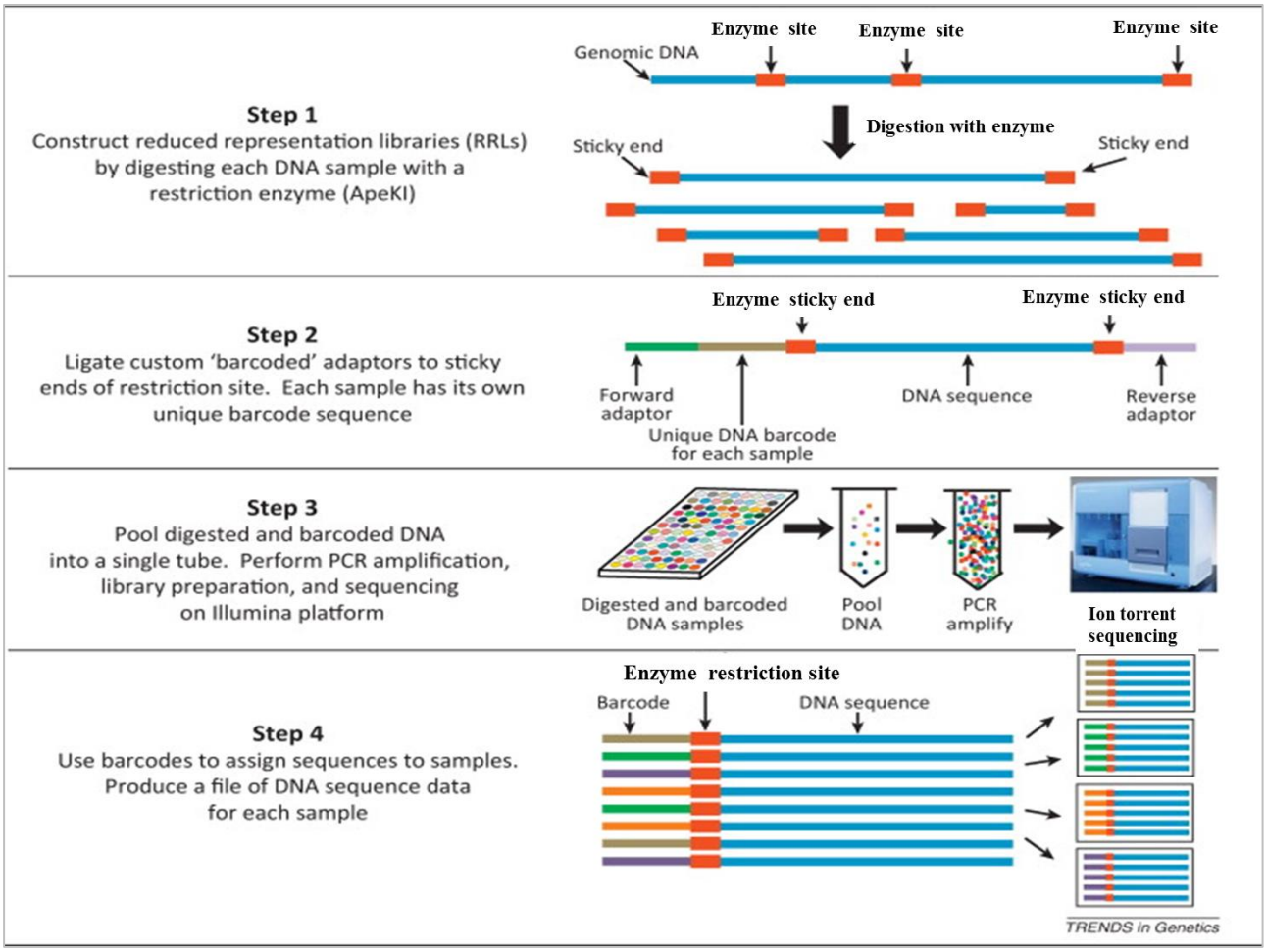


Fig. 20. Steps of libraries preparation and sequencing (The Genotyping-by-Sequencing libraries were prepared in 96-plex following co-digestion with two restriction enzymes *Pst*I (CTGCAG) and *Msp*I (CCGG), as per Elshire et al. (2011).

II.2.3.3. Genotyping-by-Sequencing analysis pipelines

The Genotyping by Sequencing **(**GBS) on Ion Torrent PGM platforms was performed on the Genomics and Bioinformatics Analysis Platform of the Institute of Integrative Biology and Systems (Laval University, Canada). Thus, the DNA sequences obtained (~ 2.4 million reads per wheat line) were analyzed on the FastGBS pipeline (Fig. 21), that align reads on the reference genome Chinese Spring v1.0 via BWA, relies on SAMtools to call SNPs as described by Sonah et al. (2013) and Sonah et al. (2014) and UNEAK pipeline (Universal Network Enabled Analysis kit pipeline, that call SNPs without a reference genome as describe by Lu et al., 2013). Both pipelines were run in the same conditions of depth of coverage (minDP=2), maximum mismatch for alignment (n = 3), Maximum Missing Taxa (MaxMT ≥80% and ≥ 30%), Minimum Sites Coverage (MnSCov=0.2 and 0.7) and Minimum Minor Allele Frequency (MinMAF=0.03).

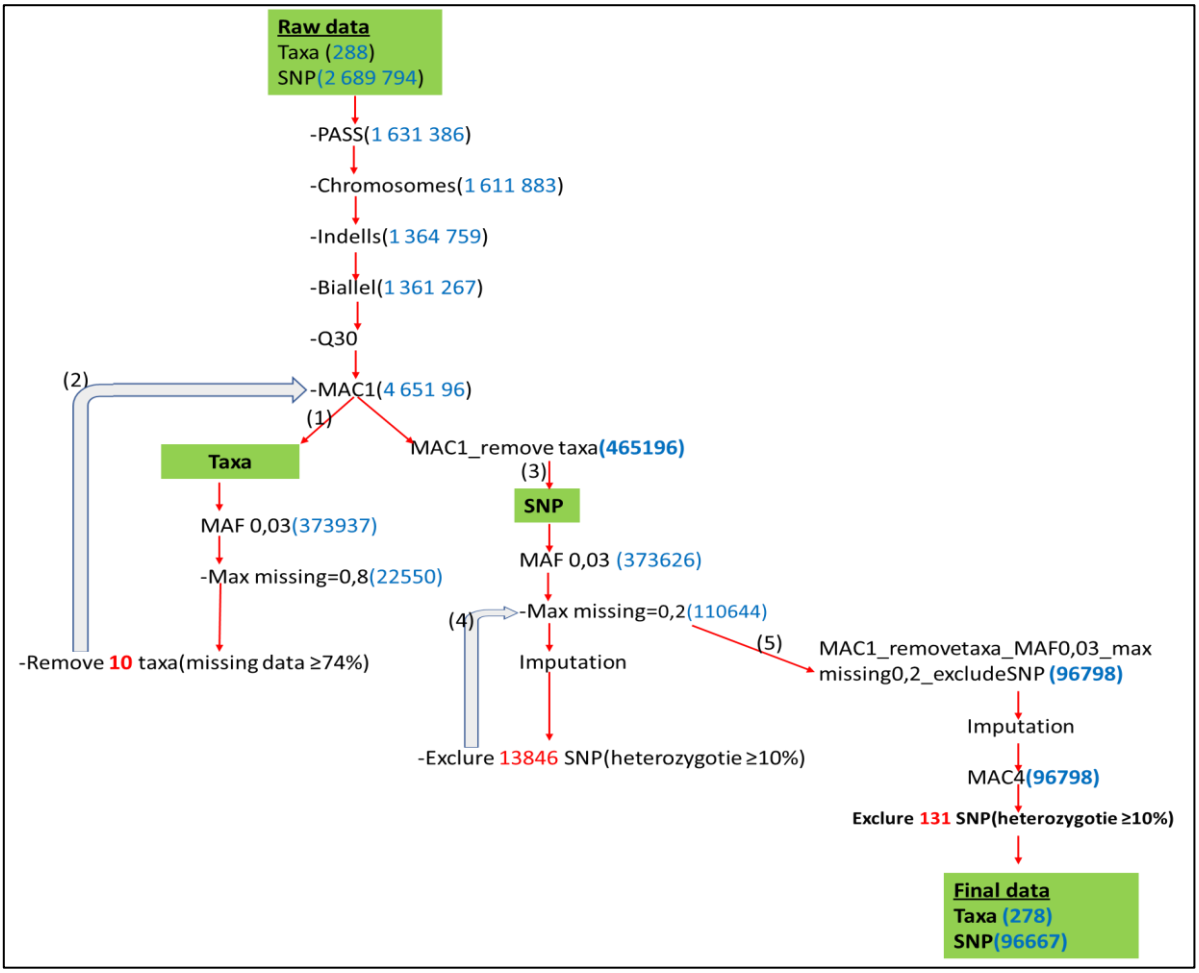


Fig. 21. Bioinformatics FastGBS' steps of filtering and imputation

The Fast-GBS analysis pipeline (Fig. 22) has been developed by integrating public packages with internally developed tools as describe by Torkamaneh et al. (2016) through the inclusion of the following core functions: (1) demultiplexing and cleaning of raw sequence reads; (2) read quality assessment and mapping; (3) filtering of mapped reads and estimation of library complexity; (4) re-alignment and local haplotype construction; (5) fit population frequencies and individual haplotypes; (5) raw variant calling; (6) variant and individual-level filtering; (7) identification of highly consistent variants. Since researchers may not always have immediate access to cluster resources, this pipeline allows either parallel processing of a large number of samples in a cluster or serial processing of multiple samples on a single machine. The "reads" were aligned to the reference genome Chinese spring v0.4 using the BWA software.

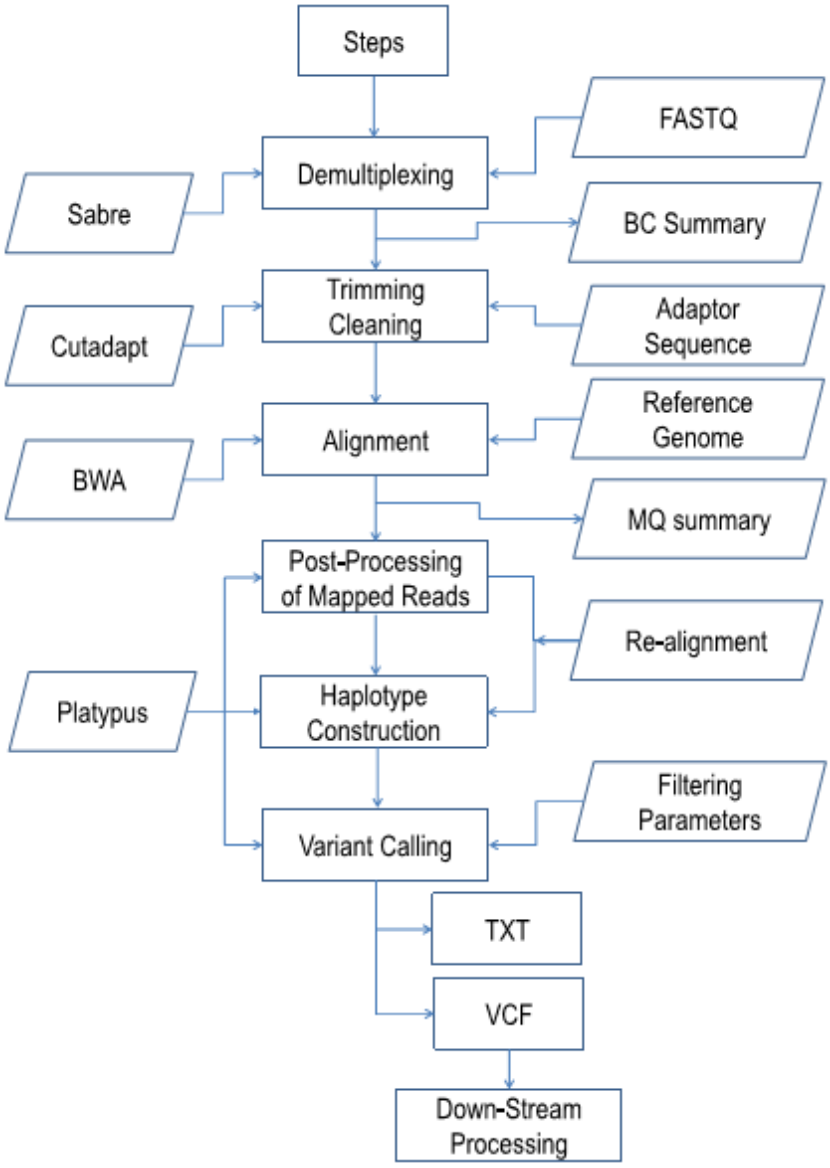


Fig. 22. Schematic representation of the analytical steps in the FastGBS pipeline (Torkamaneh et al., 2016).

The main steps in the analytical process are indicated in the central portion of the diagram, while the different software tools used are indicated to the left and inputs and outputs of each step to the right (Torkamaneh et al., 2016). The BEAGLE v4.0 (Browning and Browning, 2007) and IMPUTE2 (Howie et al., 2005) softwares tools were used to impute missing data. High proportions of missing taxa and loci, including the untyped loci were imputed.

In UNEAK, the general design is as follows: 1) reads are trimmed to 64 bp; 2) identical 64-bp reads are collapsed into tags; 3) pairwise alignment identifies tag pairs having a single base pair mismatch. These single base pair mismatches are candidate SNPs. A “network filter” is employed to discard repeats, paralogs and sequencing errors, resulting in a collection of reciprocal tag pairs, or SNPs (Fig. 23).

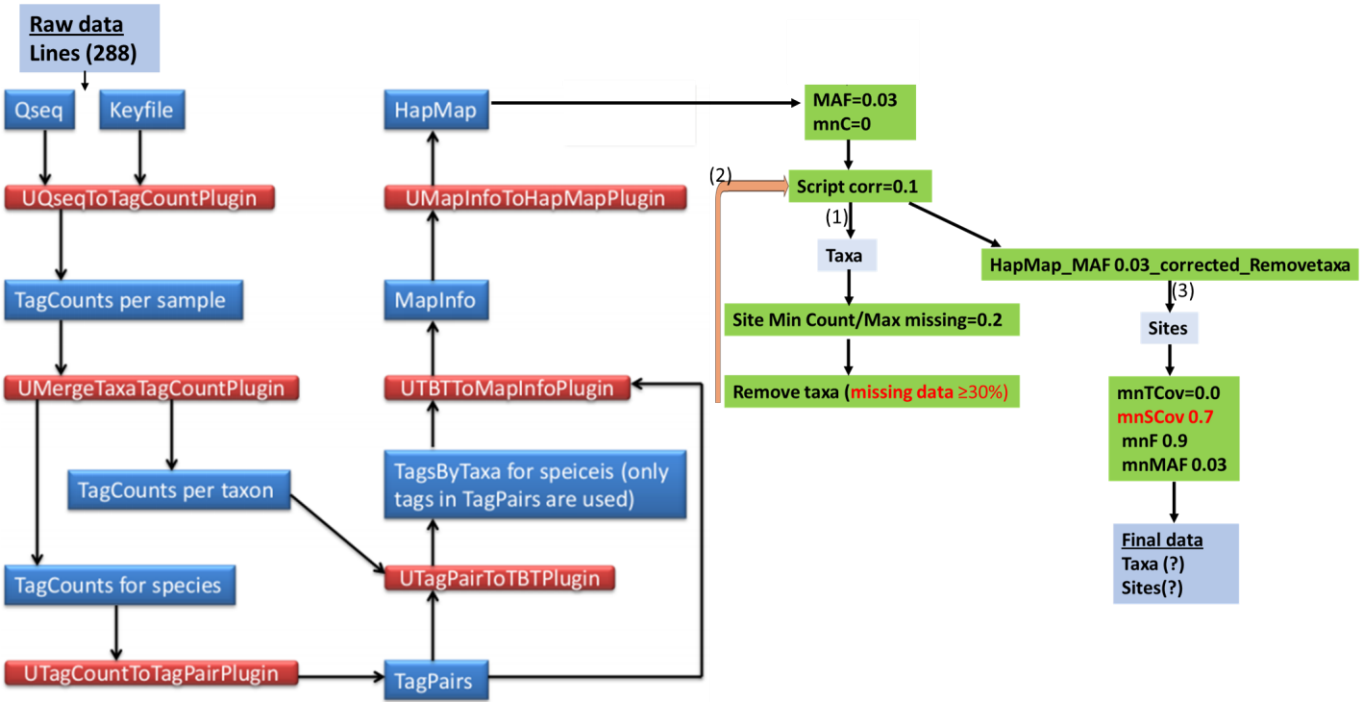


Fig. 23. Bioinformatics UNEAK analysis pipeline

II.2.3.4. Genotyping-by-Sequencing accuracy

The accuracy of genotyping by sequencing called SNP was determinate by comparing the SNP identifying using the SNPs array data, obtained from the characterization of the sub-collection of 71 wheat lines genomic diversity using a high-density 90k SNP Infinium array.

II.3. Identification of QTL for essential yield components in wheat

II.3.1. Plant material

The plant material, consisting of 170 winter and spring wheat varieties, representing the diversity present in Central Africa (Cameroon, Tchad), North Africa, South Africa, East Africa (Kenya, Ethiopia), France and Mexico (Mexicali and Baja California). Some accessions were originally obtained from the International Maize and Wheat Improvement Center (CIMMYT) and from the International Center for Agricultural Research in the Dry Areas (ICARDA).

II.3.2. Phenotyping

Field trials were conducted in the season 2015/2016 in Munt Mbankolo and during 2016/2017 in Nkolbisson. At each trial site, an incomplete alpha-lattice design with two replications was used. Each accession was planted in a single 3-m row with 25 cm spacing between the rows. The wheat spikes were harvested at physiological maturity. The grains were manually removed from the spikelet. Thousand seed weight (Wgr), grain length (Gle), grain width (Gwi) and grain

yield (Gyi) were recorded for each accession. The measurements for grain length and width were taken using a digital Vernier caliper on 20 grains for each accession.

II.3.3. Genotyping

Genomic DNA was extracted from 100 mg dry young leaves using the DNeasy 96 Plant kit (Qiagen, cat. no. 69181) following the manufacturer‘s protocol. DNA was quantified using Quant-iT™ PicoGreen as well as Thermo Scientific Nanodrop 8000 spectrophotometer instrument (Fisher Scientific). DNA concentrations were normalized to 10 ng/µl and subsequently used for library preparation.

GBS libraries were prepared in 96-plex following co-digestion with two restriction enzymes *Pst*I (CTGCAG) and *Msp*I (CCGG), as per Elshire et al. (2011), followed by ligation of barcoded adaptors. The samples were pooled per plate and PCR amplified (Poland et al., 2012). Each library was sequenced on an Ion Torrent PGM sequencer (Genomics Analysis Platform of the Institute of Integrative Biology and Systems, Laval University).

II.3.4. Statistical and population structure analyses

The analysis of variances for each trait were obtained using PROC MIXED in SAS 9.4. Each cultivar was considered as a fixed effect, whereas replications and environments were considered as random effects. Pearson correlation coefficients between pairs of phenotypic traits were also computed using PROC CORR in SAS 9.4. The broad-sense heritability for each trait was estimated by the formula $h^2 = \frac{\sigma_g^2}{\sigma_p^2} = \frac{(MS_{entry} - MS_{residual})/r}{MS_{entry}/r}$, where r: number of repetition; MS: mean square. σ_g^2: component of genotypic variance; σ^2: residual variance; σ_p^2: component of phenotypic variance.

The Genomics Analysis Platform of the Institute of Integrative Biology and Systems (Laval University, Canada) provided raw genotypic data with 2690424 SNPs. By applying several in-house filtering parameters, we selected only biallelic SNPs with a minor allele frequency (MAF) > 0.05. From this, 73 784 SNPs markers remained and were used for further analyses.

Population structure was evaluated using fastStructure version 1.0 (Anil et al., 2014) with a simple prior and 1000 iterations for K ranging from 1 to 12. To diminish the effect of high admixture within the population, structure analysis was performed on wheat accessions with MAF > 0.05 as recommended by Sobota et al. (2015). The optimal range of K was determined based on model complexity using the marginal likelihood method using the fastStructure script chooseK.py, as well as on visualization of the log marginal likelihood, and population visualization using Distruct version 1.1 (Rosenburg, 2004).

II.3.5. Genome wide association analysis (GWAS)

The GWAS were performed using the Genomic Association and Prediction Integrated Tool (GAPIT) (Lipka et al., 2012). A mixed linear model (GLM) was used with or without the covariate P from principal component analysis (PCA) and a kinship matrix was calculated either using the VanRaden method (K) or the EMMA method (K_) to determine relatedness among individuals (Lipka et al., 2012). A multi-locus mixed model (MLMM) incorporating a kinship matrix (K or K_) along with a P or Q matrix was used to test for marker-trait association (Segura et al., 2012). As mentioned by Wang et al. (2012) and Yang et al. (2013), the negative log(1/p) was used to establish a significance threshold. Best linear unbiased predictors (BLUP) were estimated for each line and each trait.

II.3.6. Identification of candidate genes

To identify a candidate gene that can control grain size in wheat, we investigated genes included in the interval of the significant QTL resulting to the GWAS analysis for the traits. Using the genome browser available for the wheat reference genome v1.0 for Chinese Spring on the *International Wheat Genome Sequencing Consortium* (IWGSC) website (https://urgi.versailles.inra.fr/jbrowseiwgsc/gmod_jbrowse), we positioned QTL flanking markers and manually investigated genes included in this interval. The function of these genes was inferred by BLAST their sequences to the UniProt the reference protein database (http://www.uniprot.org/blast/).

II.3.7. Examination of haplotype in and around the candidate gene

Using a catalog of 73,784 SNP markers, a total of 170 wheat breeding lines contrasting for the candidate gene were examined to define the SNP landscape around the gene using an in-house script (Tardivel et *al*., 2014; Tardivel et *al*., submitted).

This approach consists in identifying a restricted set of markers in the genomic area surrounding the gene of interest such that the markers in this set capture and summarize the genetic polymorphism observed in the vicinity of the gene. The underlying assumption is that every unique combination of markers defines a distinct haplotype that potentially correlates with phenotypic differences among different alleles of the gene of interest.

II.4. Ultra-Dense SNP Genotyping to discover new genomic sources for tolerance to acid soil with aluminum toxicity in wheat

II.4.1. Plant material

A subset of 45 lines of bread wheat (*Triticum aestivum* L.) were used for this study. Betta (acid soil-sensible), Atlas 66 (acid soil-tolerant) and Chinese spring (partially tolerant) were included as standards control varieties. These varieties were retained according to their level expressed under acidic soil conditions in South Africa, Kenya, Cameroon, and Ethiopia.

II.4.2. Greenhouse sand screening

Bread wheat varieties were evaluated under greenhouse conditions at the Welgevallen Experimental Farm, Stellenboch University, South Africa (Fig. 24).

Seeds were disinfected with 1% sodium hypochlorite (w/v) and then germinated in filter paper in Petri dishes for 4 days. Thereafter, 5 seedlings from each genotype were transferred into plastic pots containing sand. The treatments consist of T1 (non-acid soil at pH = 5.7 with Al =0 mg/L) and T0 (acid soil at pH 4.3 with 50 mg/L $AlCl_3$). All pots in the greenhouse were supplied daily with a complete nutritive solution consisted of 164 g Sol-u-fert-T3T; 2g Microplex GA (Fe-EDTA (60.84g/kg), Mn (32.24 g/kg), Zn(20.25 g/kg), Cu(3.75 g/kg), B (20.07 g/kg), Mo(2.16 g/kg), 77 ml potassium nitrate diluted in 100L H2O, 0.05% jik,Tap water; Calcinit (15.5 % N, 19% Ca); Fe-EDTA (60.84g/kg), Mn (32.24 g/kg), Zn(20.25 g/kg), Cu(3.75 g/kg), B (20.07 g/kg), Mo(2.16 g/kg), Sol-u-fert-T3T (164 g), N (15.5 %), Ca (19%), H2O. In T1, the Al treatments (50 mg/L $AlCl_3$) and the irrigation solution (pH 4.3 using 1M HCl) were supplied daily to plants after 4 days growing in pots.

The experiment was laid out in a factorial randomized block design with 3 replications in each treatment. Plants were harvested 20 days after planting of and the sand was washed off gently from the roots under tap water. The shoots were excised from the roots and both were rinsed in distilled water. The plant tops and roots were dried separately in a hot air oven at 35°C for 72 hours and the dry matter yields was determined.

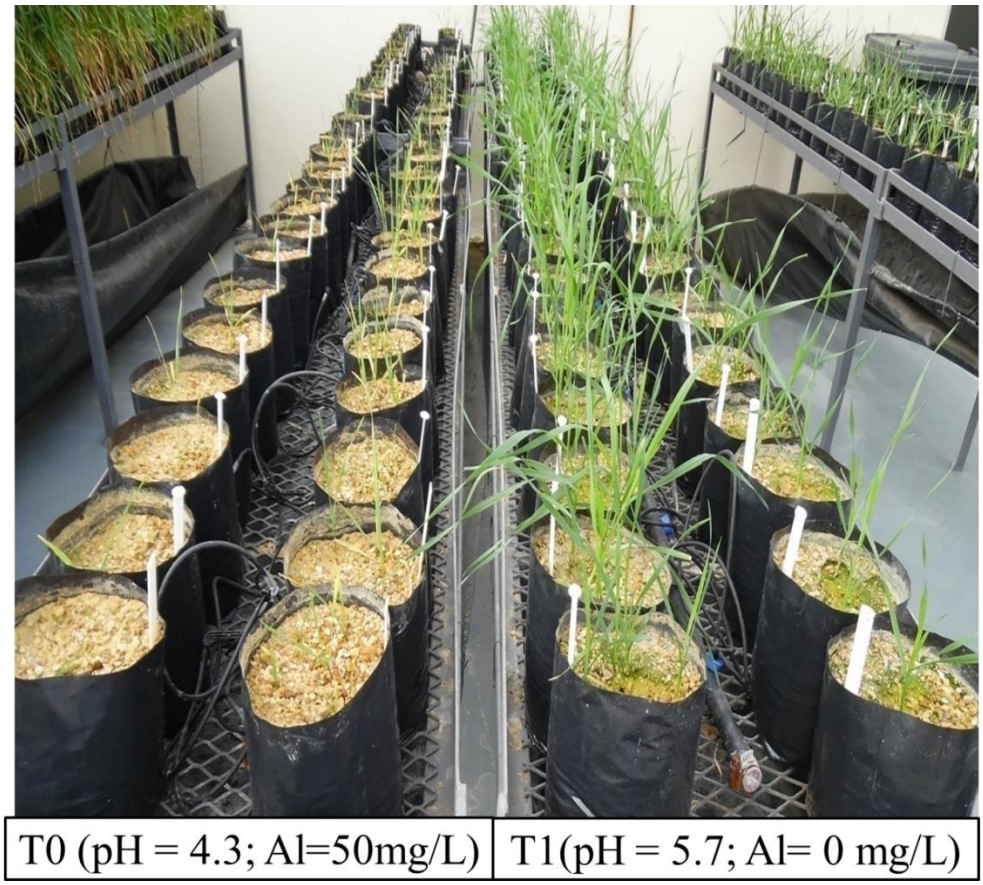


Fig. 24. Screening of wheat breeding lines for acid soil tolerance in greenhouse

At the Institute for Plant Sciences (Department Agriculture, Western Cape, South Africa) the pH of the representative 0 to 10 cm depth used like samples were tested. Sand samples were extracted with a probe 2.54 cm in diameter from more than 10 points across location in Welgevallen, Stellenbosch. All the samples from the 0 to 10 cm depth were compiled and mixed together to make a single location-specific sample; In addition to pH, the Department of soil analysis of Stellenbosch University measured extractable Al concentrations using 1.0 mol L–1 KCl extraction and inductively coupled plasma analysis (Table IV).

Table IV. The pH and aluminum (Al) concentration of sand plot soil samples taken at 0 to 10 cm depths in the Stellenbosch field used for acidity tolerance screening.

Soil Composition	**T1**	**TO**
Depth	0-10cm	0-10cm
pH (KCl)	5,7	4,3
Aluminium (mg/l)	0	50

II.4.3. Data collection and statistical analysis

- Data on root and shoot length (mm), and dry weight of root and shoot (g) were collected from each treatment in each replication 20 days after planting from all plants and the average of 3 plants were used for statistical analysis;
- Tolerance index (relative values) was compute as the ratio of the measured parameters under acid soil versus non-acid soil conditions.
- Operational heritability (h^2) was estimate from data analysis performed by Genstat software

$$h^2 = \frac{\sigma_g^2}{\sigma_p^2} = \frac{(MS_{entry} - MS_{residual})/r}{MS_{entry}/r}$$

, where r: number of repetition; MS: mean square. σ_g^2: component of genotypic variance; σ^2: residual variance; σ_p^2: component of phenotypic variance.

II.4. 4. Citrate efflux analysis

Organic anion efflux from excised root apices was measured on individual seedlings using as few as four apices by modifying the method described by Ryan et al. (1995a, 2009). Briefly, seeds were surface sterilized with bleach and thoroughly rinsed in sterile water. Twelve seeds were placed 8 in sterile conical flasks with 20 mL of 0.2 mM CaCl2, pH 4.3, and kept on a rotary shaker. After 6 days, approximately 4-mm root segments from each seedling was excised in petri dishes and washed in 1 mL of control solution (0.2 mM CaCl2, pH 4.3) for 1 h on a platform shaker (60 rpm). The solutions were replaced by 1 mL of treatments solution (200 µM CaCl2; 200 µM AlCl3) and returned to the shaker for 1.5 to 2.5 h. Aliquots were remove after 2.5 hours for malate and citrate analysis.

Then, the concentration of citrate was estimated with coupled enzyme assays that detect the production or consumption of NADH (Delhaize et al., 1993b; Wang et al., 2007).

Citric acids in nutrient solutions and roots were assayed using enzymic methods. For citric acid, 2.52 mL of sample was incubated with 0.24 mL of buffer (1 M Tris-C1, pH 7.8), 30 µL of 10 mM NADH, and 10 µL of a lactate dehydrogenase/malate dehydrogenase mixture (0.5 mg/mL for each). After a stable reading was obtained, 10 µL of citrate lyase (Boehringer Mannheim / R-Biopharm Citric acid Roche Cat. No. 10139076035 Manufacturer: R-Biopharm AG, Darmstadt, dissolved in water to 190 mg/mL) were added and the decline in A_{340} due to oxidation of NADH was monitored on a chart recorder. The amount of NADH oxidized in reactions is stoichiometric to the amount of citrate. NADH is determined by absorbance at 340nm.

Principle of the method:

Citric acid (citrate) --------CL-----> oxaloacetate + acetate

Oxaloacetate + NADH + H+ ----- L-MDH -------> L-malate + NAD+

Pyruvate + NADH + H+ ----------L-LDH---------> L-lactate + NAD+

The amount of NADH oxidized in reactions is stoichiometric to the amount of citrate. NADH is determined by absorbance at 340nm.

CL = Citrate lyase

L-MDH = L-Malate dehydrogenase

L-LDH = L-Lacate dehydrogenase

NADH = Reduced Nicotinamide-adenine dinucleotide

II.4.5. Genotyping and systematics analysis of haplotypes around the MATE1B gene

Genomic DNA was extracted from 100 mg dry young leaves using the DNeasy 96 Plant kit (Qiagen, cat. no. 69181) following the manufacturer‘s protocol. DNA was quantified using Quant-iT™ PicoGreen as well as Thermo Scientific Nanodrop 8000 spectrophotometer instrument (Fisher Scientific). DNA concentrations were normalized to 10 ng/µl and subsequently used for library preparation. The Genotyping-by-Sequencing libraries were prepared in 96-plex following co-digestion with two restriction enzymes *Pst*I (CTGCAG) and *Msp*I (CCGG), as per Elshire et al. (2011) then, followed by barcoded adaptors ligation of each samples. The samples were pooled per plate and PCR amplified (Poland et al., 2012). Each library was sequenced on the Ion Torrent PGM platforms.

The phenotypic variation observed among the 45 cultivars makes it possible to explore the genetic basis, in order to define the haplotypes and the cultivars carrying the common alleles / gene of interest. In order to know the allelic variations (around the MATE1B gene) encountered in the phenotyped wheat cultivars for Al tolerance, we defined haplotypes (using the catalog of 80 124 SNPs).

CHAPTER III: RESULTS AND DISCUSSION

III.1. RESULTS

III.1.1. Genetic diversity of Cameroonian bread wheat (*Triticum aestivum* L.) cultivars revealed by microsatellite markers

III.1.1.2. Characteristics of markers and genetic diversity

All pairs of primers specific for SSR locus used resulted in a positive amplification with allelic variations in size on all DNA of wheat accessions. A total of 77 microsatellite alleles were detected. The number of alleles per locus varied from 2 (Xgwm 125 and Xgwm 331) to 13 (Xwmc 177), with an average of 7 alleles per locus. Genetic diversity for microsatellite loci ranged from 0.46 (Xgdm 125) to 0.90 (Xgwm 177) with an average of 0.88. The polymorphism information Content (PIC) varied from 0.25 (Xwmc 331) to 0.89 (Xwmc 177), with an average of 0.69 (Table V).

The results indicated a significant correlation ($P < 0.01$) between gene diversity and number of alleles across wheat accessions in both Regions (Fig. 25). The correlation coefficient between these two variables over the 11 loci were 0.88 (Adamawa) and 0.76 (Northwest).

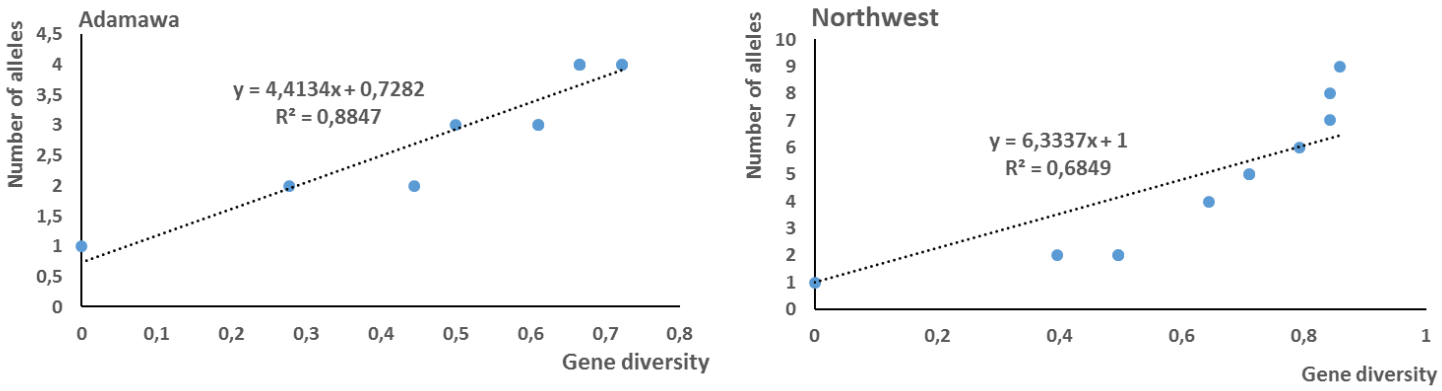


Fig. 25. Correlation between gene diversity and the number of alleles over 11 microsatellite loci in hexaploid wheat.

Table V. Description of SSR Markers

Locus	Chr position	Primers sequences	Repeat	Bases expected	Ann. temp.	Alleles frequency	Number of alleles	Gene Diversity	PIC
Xwmc 11	1A, 3A	5' TTGTGATCCTGGTTGTGTTGTGA 3' 5' CACCCAGCCGTTATATATGTTGA 3'	CT	177	61	0.29	8	0.83	0.81
Xwmc 59	1A. 6A	5' TCATTCGTTGCAGATACACCAC 3' 5' TCAATGCCCTTGTTTCTGACCT 3'	(CA)19	197	58	0.18	10	0.89	0.87
Xwmc 177	2A	5' AGGGCTCTCTTTAATTCTTGCT 3' 5' GGTCTATCGTAATCCACCTGTA 3'	(CA)21	184	52	0.18	13	0.90	0.89
Xgwm 190	5D	5' GTGCTTGCTGAGCTATGAGTC 3' 5' GTGCCACGTGGTACCTTTG 3'	(CT)22	201-253	55	0.18	9	0.87	0.86

Xgwm 437	7D	5' GATCAAGACTTTTGTATCTCTC 3' 5' GATGTCCAACAGTTAGCTTA 3'	(CT)24	109-111	47	0.18	10	0.88	0.87
Xgwm 539	2D	5' CTGCTCTAAGATTCATGCAACC 3' 5' GAGGCTTGTGCCCTCTGTAG 3'	(GA)27	143-157	60	0.24	8	0.83	0.81
Xdgm 125	4D	5' GCAGGCGTGTTACTCCAAGT 3' 5' CCGAGGTGGATAGGAGGAAA 3'			60	0.65	2	0.46	0.35
Xwmc 331	4D	5' CCTGTTGCATACTTGACCTTTTT 3' 5' GGAGTTCAATCTTTCATCACCAT 3'		128	61	0.82	2	0.29	0.25
Barc 133	3B	5' AGCGCTCGAAAAGTCAG 3' 5' GGCAGGTCCAACTCCAG 3'	(CT)24			0.65	4	0.52	0.47
Xgwm 133	6B	5' ATCTAAACAAGACGGCGGTG 3' 5' ATCTGTGACAACCGGTGAGA 3'	(CT)39			0.35	4	0.72	0.67
Xgwm 644	6B	5' GTGGGTCAAGGCCAAGG 3' 5' AGGAGTAGCGTGAGGGGC 3'	(GA)20			0.29	7	0.79	0.76
Mean						**0.36**	**7**	**0.72**	**0.69**

Xgwm and Xwmc markers were obtained respectively from Röder *et al*. (1998) and Somers and Isaac (2004; Grain Genes); PIC: Polymorphism Information Content.

III.1.1.3. Genetic relationship and diversity among different geographical regions

Genetic Distance value (GD) indicates that some accessions are closely related. The GD over accessions in all regions ranged from 0.18 (between Wassande2 and NGDEREM3) to 1 with a mean of 0.8 (80%). So, at 80% of genetic divergence, the 17 wheat cultivars studied were structured into 5 main groups (A, B, C, D and E) in the dendrogram based on the UPGMA analysis using SSR data (Fig. 26). Group A included 4 cultivars (Alexander wonder, Riba, Vrack and FUANB3). Very close to the 80% of genetic dissimilarity, the group B could be divided into two subgroups: subgroup B1 contained 6 accessions (BABANKIT, FUANB1, FUANB4, FUANB2, BBT2 and HGW) while subgroup B2 included only one cultivar (IRAT 10). It is noteworthy that the two varieties FUANB1 and FUANB4 are identical. Group C contained 4 cultivars (Ngderem1, Ngderem3, WASSANDE 2 and NGM2). Groups D and E each contained 1 cultivar, respectively Ngderem4 and SONALIKA.

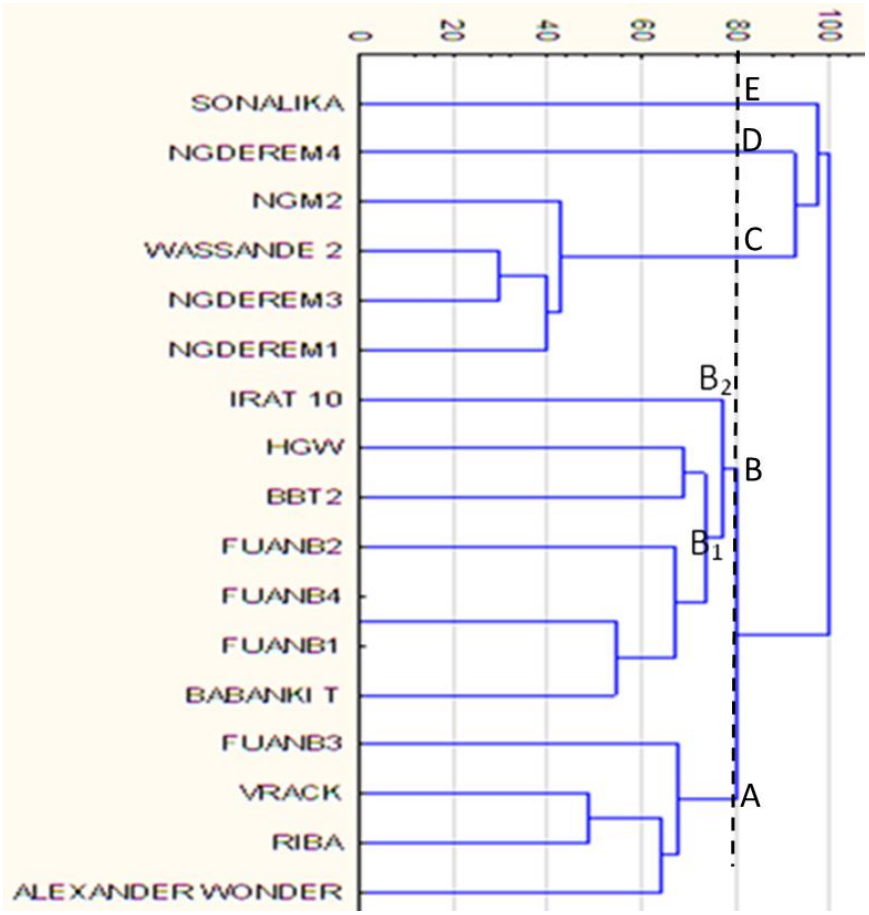


Fig. 26. Grouping according to the dissimilarities between 17 accessions of hexaploid wheat on the basis of the SSR profiles of 11 loci.

Furthermore, the Principal Component Analysis (PCA) for the six-collection village split the accessions into five clearly distinct groups. The first two principal components had Eigenvalues of 6.36 and 3.81. The PCA grouped the 17 wheat accessions into various components with the first two explaining 59.86 and 37.44% of the total variation. Accessions from each village were approximal clustered together (Fig. 27). So, 80% of the genetic material from the same geographical village could be clustered in specific groups.

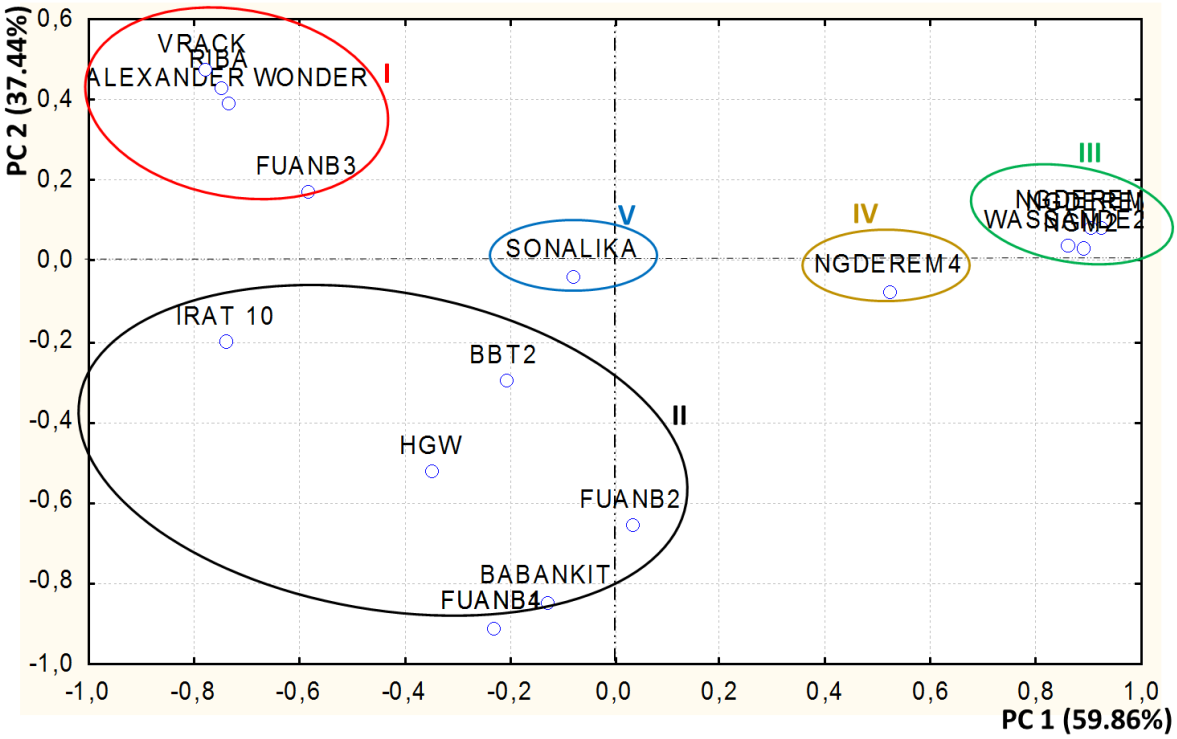


Fig. 27. Principal Component Analysis of 17 hexaploid wheat accessions from 6 villages in Cameroon. The grouping is based on Dice's similarity coefficients.

Accessions were then analyzed separately according to their region of origin (Adamawa and Northwest). A comparison of the genetic diversity of wheat accessions was done between two germplasm pools. The mean of gene diversity, number of alleles per locus, total number of alleles and the number of accessions carrying rare alleles were higher in Northwest, compared to those in Adamawa Region (Table VI). These results suggest that the Northwest area exhibited greater genetic diversity than Adamawa region, even after taking into account the effect of collection size.

Table VI. Analysis of geographical regions

Item	Northwest	Adamawa
Number of accessions	n = 11	n= 6
Total number of alleles	51	34
Average number of alleles per marker	4.64	3.09
Number of rare alleles	5	6
Mean of PIC-values	0.57	0.48
Mean of Gene diversity	0.62	0.54

III.1.2. Genetic population in a global accession of hexaploid wheat breeding lines

III.1.2.1. Agro-morphological variation

The phenotyped wheat breeding accessions were classified into 4 groups in a phylogenetic tree, to which each group showed clear specificity traits for which it differs from that of the others (Fig. 28). Also, most of wheat lines were grouped according to its adapted geographical region.

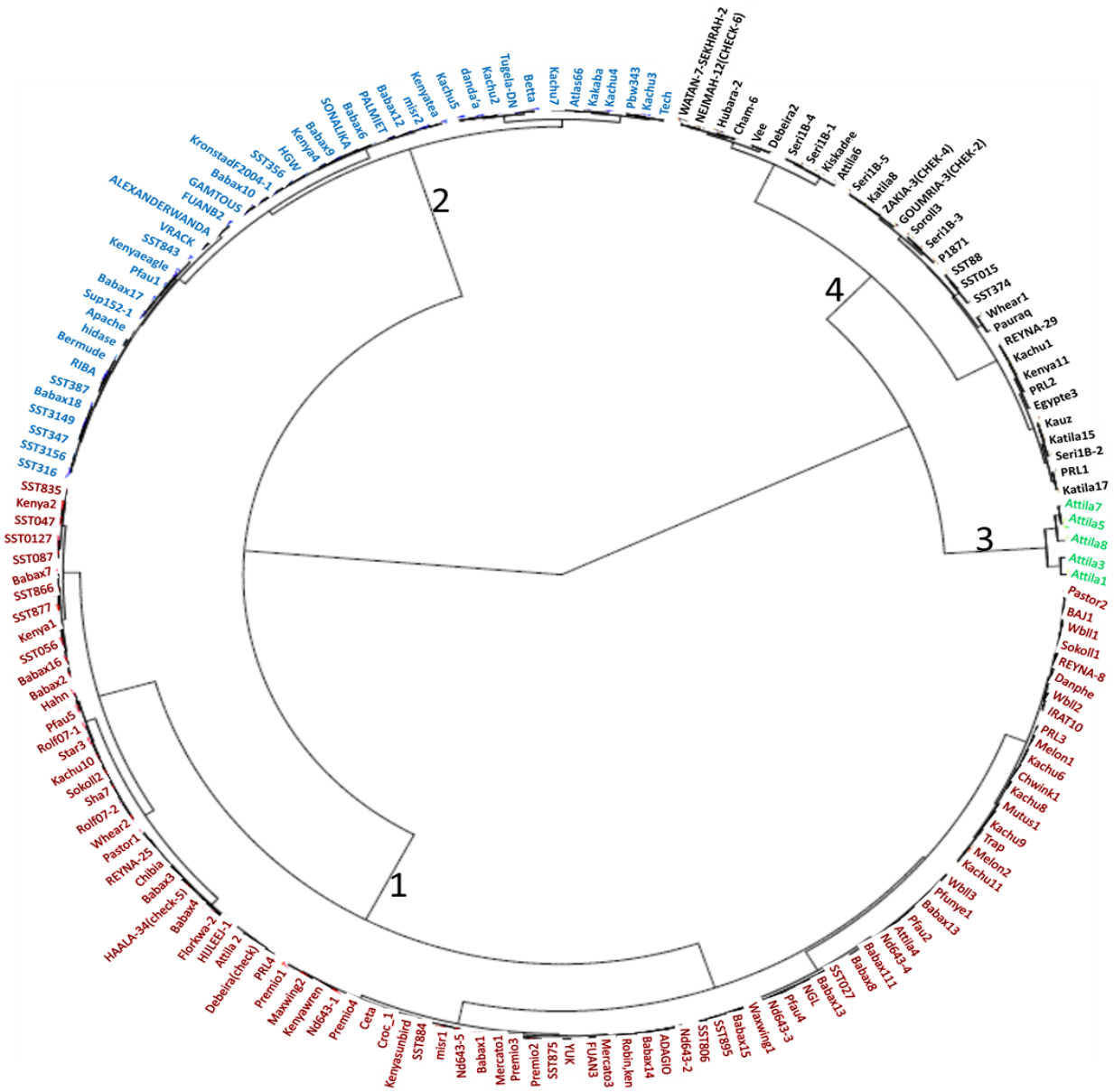


Fig.28. Dendrogram using 39 agro-morphological wheat descriptors. Colors indicated the memberships in group

The results of principal component analysis (PCA) indicated five principal components with Eigen-values greater than 1.0. The PCA grouped the 39 wheat descriptors into various

components with the first five explained 34.93 % of the variation, while the first principal component (PC1) explained alone 12.13 % of the variation (Table VII).

The PC1 had high loading for leaf blade angle, existence of awn, internode length, leaf length, leaf width, glume color, culm stiffness, leaf sheath pubescence, Ear prostrusion, Spike waxiness, ear length, ear width, weight of grains. The second component (PC2) explaining about 7.33 % of the total variation was correlated with plant habit, leaf color, leaf number, flowering date, culm length, maturity date, culm thickness, spike shape, spikelet density, humidity, grain color. The third component (PC3) explaining about 5.89 % of the total variation was correlated with existence of awn, culm angle, ear length, glossiness of grain and grain yield. The fourth component (PC4) explaining about 5.06 % of the total variation was associated with leaf flecking and awn length, while the fifth component (PC5) explaining about 4.53 % of the total variation was correlated with leaf angle, tiller number, culm thickness, glume pubescence and 1000-grain weight.

Table VII. Principal Component Analyzes of 39 wheat descriptors

		Eigen vectors				
Main Component Number		PC1	PC2	PC3	PC4	PC5
Eigen values		4.24	2.56	2.06	1.77	1.59
Cumulative (%)		12.13	19.45	25.34	30.40	34.93
Total variance (%)		12.13	7.33	5.89	5.06	4.53
		Eigen values				
plant habit	Pha	-0.02	0.46	-0.24	-0.05	-0.20
leaf color	Lco	0.28	0.37	0.19	-0.38	-0.30
leaf number	Lnu	0.24	0.42	0.23	0.20	0.07
leaf flecking	Lfl	-0.27	-0.08	0.03	0.38	0.16
leaf blade angle	Lba	0.53	0.10	-0.03	0.01	0.24
leaf angle	Lan	-0.29	-0.17	0.22	-0.35	0.38
tiller number	Tnu	0.25	0.08	0.07	-0.35	0.29
existence of awn	Eaw	0.20	0.06	0.20	0.19	0.16
flowering date	Fda	-0.21	0.34	-0.18	0.12	0.27
culm angle	Can	-0.78	-0.04	0.23	-0.02	-0.10
internode length	Ile	0.36	0.04	-0.20	-0.05	-0.29
leaf length	Lle	0.15	0.06	0.07	-0.39	-0.04
leaf width	Lwi	0.14	-0.08	-0.02	0.09	0.13
chlorophyll	Chl	-0.01	-0.16	-0.05	-0.19	-0.32
culm length	Cle	-0.01	0.25	-0.18	-0.05	0.00
glume color	Glco	0.58	-0.29	0.10	0.04	-0.20
maturity date	Mti	0.17	0.57	-0.32	0.23	0.18
culm thickness	Cth	0.08	0.63	-0.19	-0.20	0.22
culm stiffness	Cst	0.71	0.05	-0.17	0.05	-0.11
leaf sheath pubescence	Lsp	0.56	0.18	-0.12	-0.07	-0.12
spike shape	Ssh	-0.52	0.27	0.00	-0.50	-0.05
spikelet density	Sde	-0.26	0.11	-0.36	-0.45	0.05
Ear prostrusion	Epr	0.61	-0.16	0.07	-0.21	0.28
Spike waxiness	Swa	0.54	0.06	-0.25	-0.19	0.21
awn length	Ale	0.13	0.09	0.17	0.27	0.08
glume pubescence	Gpu	-0.07	-0.32	0.22	-0.38	0.25
ear length	Ele	0.39	-0.13	0.41	-0.05	-0.12
ear width	Ewi	0.34	-0.14	-0.11	0.04	0.16
humidity	Hum	-0.18	0.36	0.10	0.06	-0.16
grain color	Gco	-0.31	0.58	0.21	0.09	-0.01
grain quality	Gqu	-0.08	-0.31	-0.62	-0.09	-0.01
glossiness of grain	Ggr	0.16	0.33	0.58	-0.04	-0.12
1000-grain weight	1000gw	-0.02	0.02	0.04	0.19	0.45
weight of grains(g)	Wgr	0.24	0.04	0.19	-0.01	-0.34
grain yield(kg/ha)	Gyi	0.17	0.11	0.43	-0.16	0.29

III.1.2.3. Variant discovery in the wheat genome

Genotyping-by-Sequencing (GBS) approach was used to characterize the genetic diversity in a global collection of 288 hexaploid wheat breeding lines. The DNA sequences obtained were analyzed on the FastGBS and UNEAK pipelines.

The UNEAK pipeline having called more raw reads, has identified fewer SNPs that could be anchored to the reference assembly but more of these SNPs remained after filtering for minor allele frequency and proportion of missing data (Table VIII). So, the UNEAK pipeline is likely more robust to SNP discovery, particularly in duplicated sequences, as it uses a population-based filtering to identify properly segregating SNPs.

The proportion of missing data was zero because the tools of FastGBS pipeline allowed to impute the missing dataset. For each of the both pipelines and datasets, SNPs were identified at MAF > 0.03 with data present in 20% of wheat lines (Criteria of filtering proper to FastGBS pipeline, named FastGBS1) and with data present in 70% of wheat lines (Criteria of filtering proper to UNEAK pipeline, named UNEAK1).

In the FastGBS1 pipeline, the number of SNP calls increased to over 87092 when SNPs with up to 80% missing data with 20% of minimum sites coverage were tolerated, while in UNEAK2 pipeline, the number of SNP increased to over 45060 (Table VIII). So, it is quite evidence to understand why the different SNP calling pipelines were primarily responsible for differences in the SNPs identified. A scarcity of SNPs identified was noted in common between the both pipelines (Fig. 29), suggesting that the different alignment and filtering parameters between the pipelines played an important role in which SNPs were retained in the final data sets.

Indeed, using the same dataset, the UNEAK2 and FastGBS1 pipelines identified 10669 SNPs in common, but sets of 36293 and 76423 SNPs were unique to the UNEAK and FastGBS pipelines, respectively (Fig. 29c). Although differences were observed between the pipelines on which SNPs were discovered, there was high agreement on genotype calls for common SNPs, 97,22% and 96,53 % to the UNEAK2 and FastGBS1, respectively (Table VIII). But, the Fast-GBS1 pipeline recorded the high number of polymorphism SNPs. The estimation of it concordance (%) with SNP Infinium Array will be highly appreciated.

Table VIII. Comparison of GBS SNP data between UNEAK and FastGBS pipelines

Platform dataset (bioinformatics pipeline)	Fast-GBS1	Fast-GBS2	UNEAK1	UNEAK2
Number of reads per line	2.4 M	2.4 M	5.3 M	5.3 M
Number of SNP	87092	32045	7475	45060
Number of lines	278	255	269	280
Agreement in genotype calls(%)	96.53	88.54	93.40	97.22
Proportion Missing	0	0	0.14	0.57
Proportion Heterozygous	0.02	0.02	0.02	0.004
Average Minor Allele Frequency	0.23	0.21	0.22	0.20
Accuracy(%)	95.1			

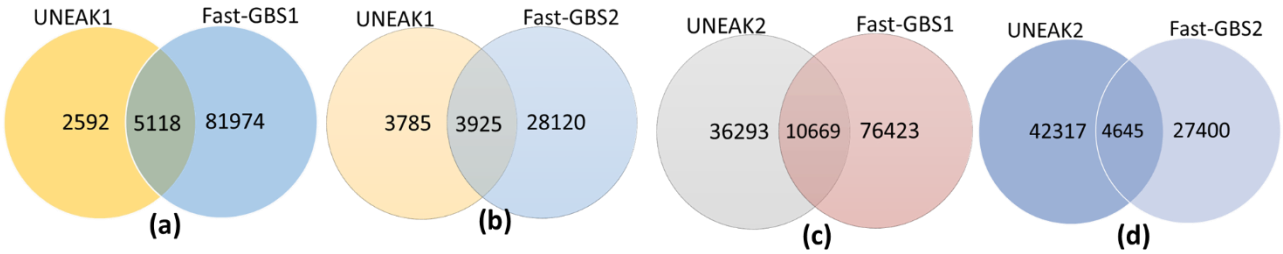


Fig. 29. Venn diagrams of the number of SNPs identified in each dataset and with the respective bioinformatics pipeline. (a) SNPs identified in the UNEAK1(Missing taxa ≥ 30%, mnSCov=0,7) with Fast-GBS1(Missing taxa ≥ 80%, max missing=0,2); (b) SNPs identified in the UNEAK1(Missing taxa ≥ 30%, mnSCov=0,7) with Fast-GBS2(Missing taxa ≥ 30%, max missing=0,7); (c) SNPs identified in the UNEAK2(Missing taxa ≥ 80%, mnSCov=0,2) with Fast-GBS1(Missing taxa ≥ 80%, max missing=0,2); (d) SNPs identified in the UNEAK2(Missing taxa ≥ 80%, mnSCov=0,2) with Fast-GBS2(Missing taxa ≥ 30%, max missing=0,7). mnSCov: minimum sites coverage.

III.1.2.4. Accuracy validation of SNP calling GBS

Using a subset of 71 lines for which SNP array (wheat 90K array) data were available, we found a high level of agreement (>95%) between SNP calls made at the same loci using both genotyping methods.

III.1.2.5. Distribution of polymorphic SNP markers on wheat genomes

The distribution of the polymorphic identified SNP markers mapped to individual chromosomes. All chromosomes were covered by unique SNPs but their distribution frequencies were different. So, using the FastGBS1 pipeline with high deep coverage of read, 96,667 polymorphic SNP were identified and positioned on the physical map (Fig. 30) and distributed over all 21 chromosomes of the hexaploid wheat (Table IX).

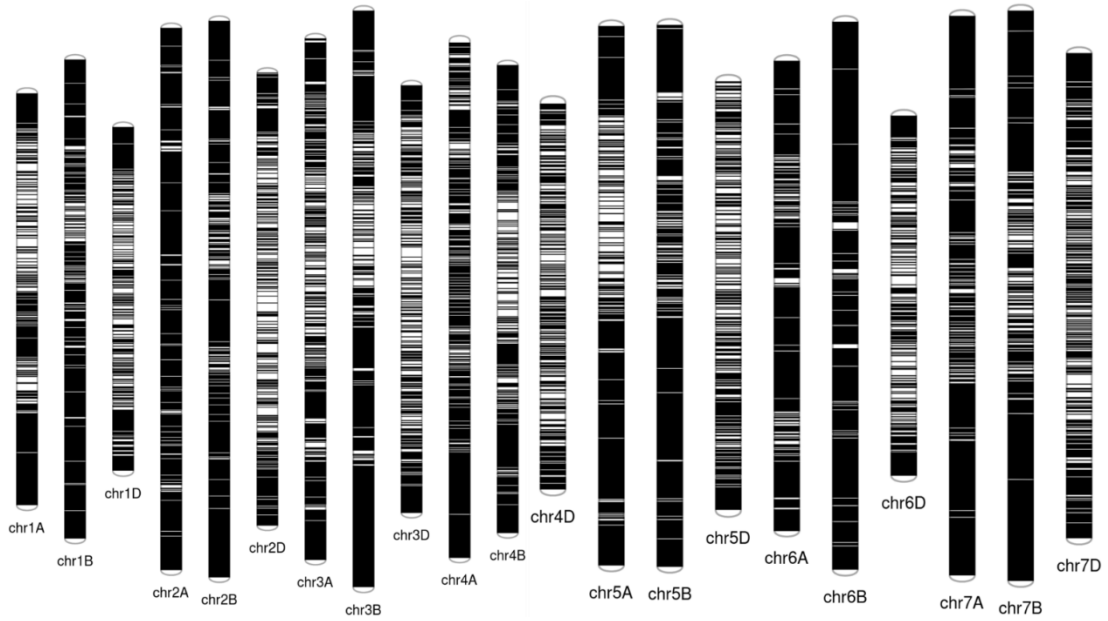


Fig.30. Genomic coverage of polymorphic SNP markers over physical map of 21 chromosomes in hexaploid wheat lines.

Consistent with previously observed levels of genetic diversity in the wheat genomes, the majority of mapped markers were located in the A (38481) and B (47215) genomes. Only 10971 of SNP markers mapped to the D genome (Table IX). In the whole genome, the lower number of SNP markers were located in 4D chromosome (581) while the greater where located in the 2B chromosome (8591). The distribution of SNP markers across homeologous chromosome ranged between 8897 (group 4) and 17538 (group 2). Althrough the 7 homeologous chromosome in hexaploid wheat, the identified 96667 SNPs were distributed following the order: group group 4(581), group 5(12104), group 1(12636), group 6(13603), group 3(14503), group 7(17386) and group 2(17538).

Table IX. Distribution of polymorphic SNP markers across the A, B and D genomes of hexaploid wheat.

Chromosomes	Wheat genome			Total
	A	B	D	
1	4629	6102	1905	12636
2	6467	8591	2480	17538
3	5048	7866	1589	14503
4	5233	3083	581	8897
5	4693	6346	1065	12104
6	4643	7606	1354	13603
7	7768	7621	1997	17386
Total	38481	47215	10971	96667

III.1.2.6. Relationship between global germplasm collections

Up to now, most genotyping methods have been designed to use on the diploid species. In our study, we have used a genotyping by sequencing approach to characterize the hexaploid wheat accessions from multiples origins in the world. The relationship between the accessions was determined by calculating a pairwise similarity matrix that was used to perform multidimensional scaling and build principal coordinate (PCO) plots. Six distinct groups were found, and each group showed clear traits of specificity for which it differs from those of others (Fig. 32).

The accessions from Canada, for which 89282 polymorphic SNP where identified (Table X), fell in the group 1, constituted of 2 sub-groups (Fig. 32). The first sub-group is composed of spring wheat lines materials from the Quebec's region and two winter lines from France, while the second sub-group englobed the spring wheat breeding lines from other regions in Canada. In this group, the pairwise similarity index between accessions is 0.29 (Table X). Furthermore, some spring wheat accessions from Western Cape town region in South Africa were found in this group.

The group 2 is mainly constituted of major diversity wheat breeding lines from South Africa and the group 3 contained in accessions mostly from North Africa. These materials of group 3 are from ICARD's bread wheat breeding program. The crosses were made in Syria at Telhadya, and yield evaluations were made across North African locations including at Egypt. These materials are high yielding with resistance/tolerance to heat stress. It is quite evidence to understand why these materials are far from Canadian material with a pairwise genetic differentiation Fst index estimated at 0.13 (Table XI). The spring genetic materials of group 4 were adapted under Central African growing areas (Cameroon and Tchad). However, some of the North African material were cluster with lines from Mexico (group 5 and 6). Also, the Eastern African accessions (Kenya and Ethiopia) were grouped with those materials resistant to heat stress. Indeed, the accessions from the group 5 were collected in Mexicali (Capital of the Mexican state of Baja California. This region is located in the northwest of the country, not far from the Californian city of San Diego) and those of group 6 were cultivated in Baja California. We noticed that the GBS approach was able to separate the hexaploid wheat accessions according to their geographical origins.

To confirm that the GBS approach was able to dissect the substructure of the adaptation type of accessions in their growing areas, we examined the phylogenetic analysis (Fig. 31).

Many clusters were evidenced; winter wheats and spring wheats/facultative accessions associated with their geographical origins were observed (Fig. 31). Eight winter (South Africa), two spring (west-Canada), twelve spring (Mexico) and two spring (North Africa) were separated from their main geographical groups; these accessions should carry the vernalization VRN-1 gene expression.

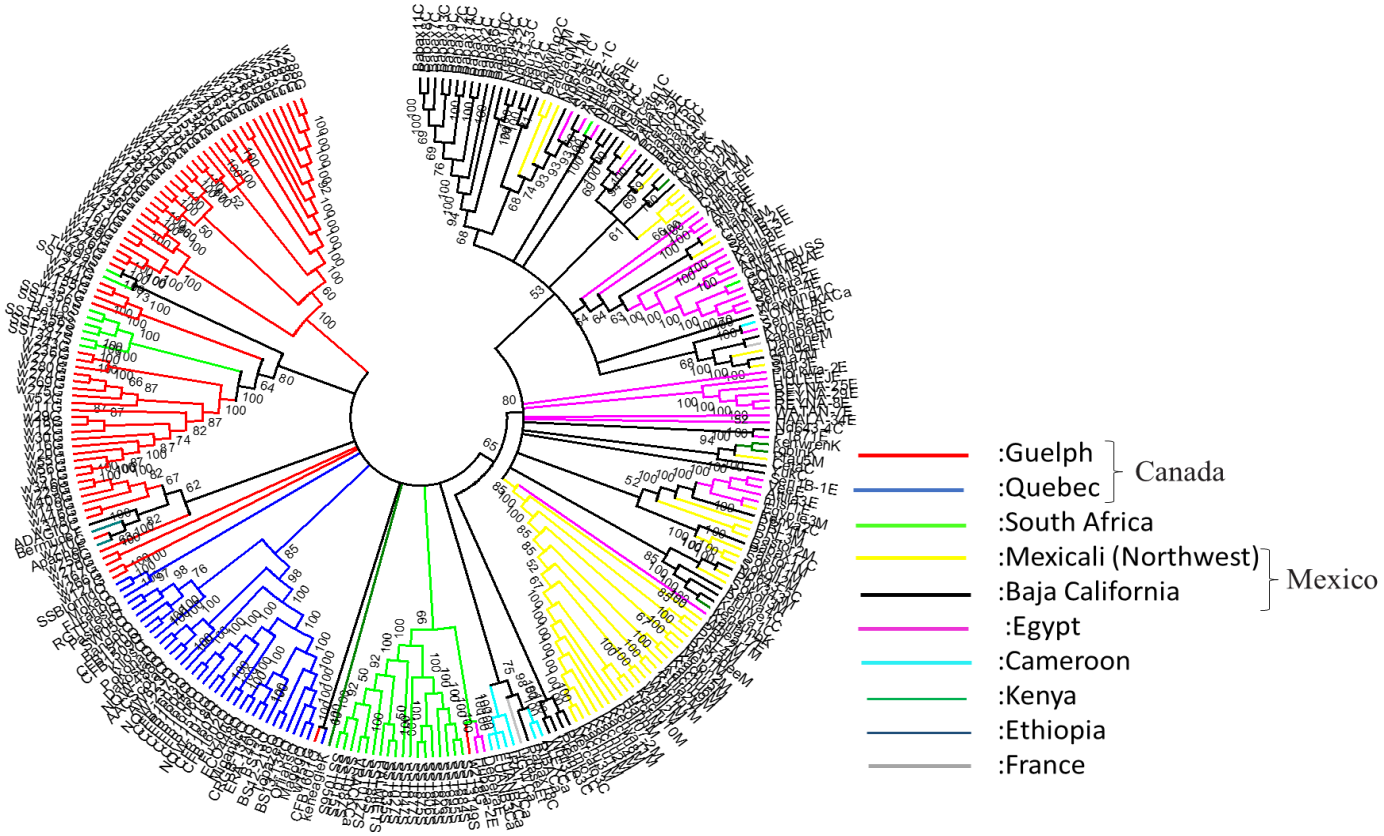


Fig. 31. Phylogenetic relationship of the 278 wheat lines accessions based on the analysis of SNPs generated by GBS. The rooted genetic distance tree was constructed using the neighbor-joining method. Numbers above branches indicate bootstrap values (>50%). Colors of branches represent the classification of accessions into their geographical localization.

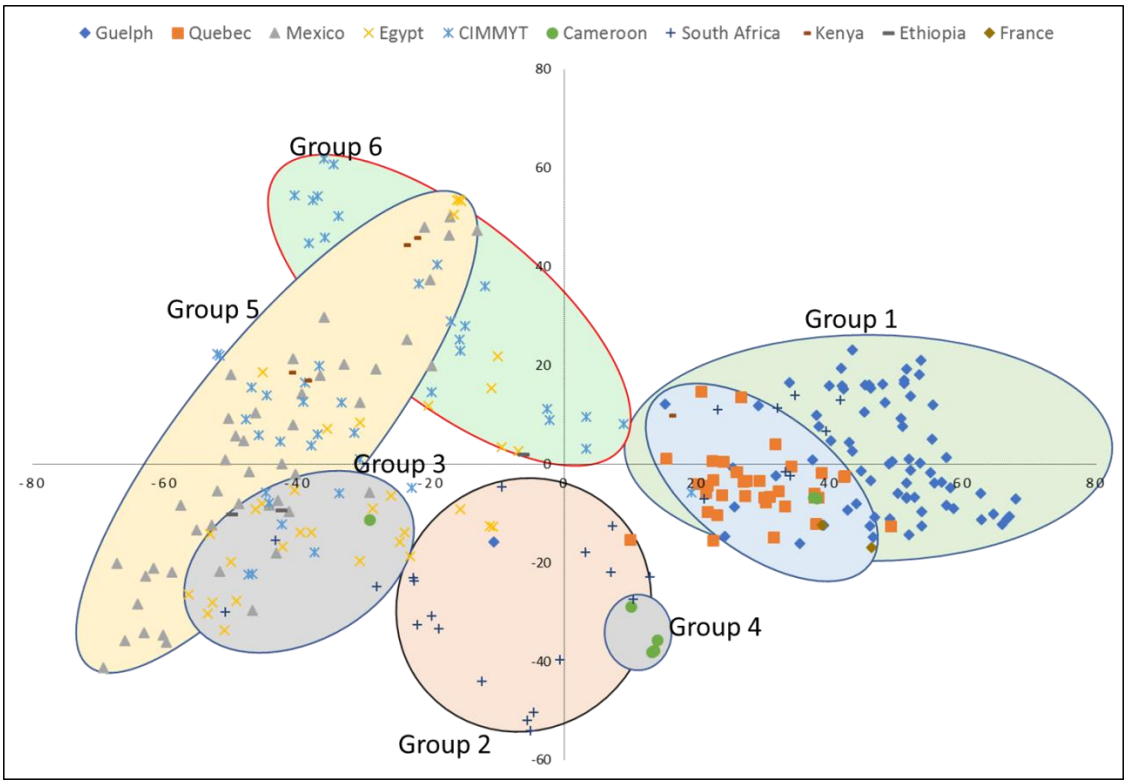


Fig.32. Principal coordinate plot (multidimensional scaling) of all 278 lines (across the A, B and D genomes of hexaploid wheat) against 96667 SNP-markers.

The GBS approach was tested by surveying SNP variation in a set of samples of 288 hexaploid wheat accessions including cultivars of different geographic origin from Canada, Mexico, France, South Africa, Central Africa, East Africa and North Africa. Due to the non-representative number of samples from France, there were not considered in the populational analysis. In the considered global collection, the nucleotide divergency on a per-SNP basis varied from 0.29 (Mexico and Canada) to 0.47 (Central and East Africa). Our analysis showed that the proportion of SNP heterozygous was higher in the Canadian materials. Regarding the heterozygosity on a per-individual basis, specfically the inbreeding coefficient estimated for each individual varied from 0.89 (Canadian collection) to 0.98 (East African material) indicating the degree of fixation of wheat lines in the collections. Furthermore, the number of biallelic polymorphic loci per population varied from 37416 in Central Africa material to 89282 in Canadian material. The larger Canadian cultivars, Mexico and South African collections had a higher number of polymorphic SNPs, with a cumulative prevalence of intermediate to high Minor Alleles Frequency SNP loci (Table IX). The lower numbers of shared polymorphisms between these and other collections are representative of the narrow genetic base compromising the collections of deletion and mutation lines which are developed in a single genetic background.

The effect of collection size on the number of polymorphic SNPs within a collection was also perceptible. The minor allele frequencies (MAF) of SNPs and pairwise similarity index within different germplasm collections were calculated as a measure of allelic diversity (Table X). So, the level of genetic diversity in the accessions was either comparable or higher than that of the population of winter and spring material, probably due to ascertainment bias in the SNP discovery panel, which included mainly of cultivars.

Table X. SNP diversity summary assessed in cultivars collection

	South Africa	**Central Africa**	**East Africa**	**North Africa**	**Mexico**	**Canada**
Samples number	27	8	8	34	91	108
Nucleotide divergency	0.37	0.47	0.47	0.34	0.29	0.29
Heterozygous proportion (%)	0.01	0.02	0.01	0.01	0.01	0.03
Average Minor Allele Frequency	0.27	0.35	0.35	0.25	0.21	0.21
Relatedness	0.07	0.22	0.22	0.06	0.02	0.02
Average pairwise similarity distance	0.37	0.45	0.45	0.34	0.29	0.29
Inbreeding coefficient	0.97	0.96	0.98	0.96	0.96	0.89
SNP markers number	67592	37416	38842	65251	79305	89282

To check the transferability of SNP markers across populations, we evaluated the number of shared alleles and the level of genetic differentiation (F_{ST}) between the lines populations (Table XI). The larger number of polymorphic SNPs were shared among Canadian and Mexico population (74577), following by South African and Canadian populations (66154) being transferrable between and useful within different collections. In contrary, the lower number were shared between Central and East African population (19609). The sharing of the majority of polymorphic SNPs suggest that the targeting of SNPs with both alleles found in at least two individuals in the discovery panel enriched the catalogue for common SNP variants.

So, this remark is coherent with the prevalence of SNPs of intermediate to high average of Minor Alleles Frequency in the populations (Table X).

However, the F_{ST} variation between the populations of different geographical origin is possibly caused by the use of different parents or by allele frequency divergence during the development of locally adapted populations. In our study, the F_{ST} varied from 0.02 (between Mexico and East African material) to 0.17 (between North and Central African material) (Table XI). This high proportion of shared alleles between wheat populations suggest that the majority of alleles for wheat improvement were contributed by local well adapted cultivars.

Table XI. Number of SNP markers shared between populations (above diagonal) and the estimates of pairwise FST (below diagonal) *

	South Africa	Central Africa	East Africa	Norht Africa	Mexico	Canada
South Africa		31494	34206	54301	61794	66154
Central Africa	0.11		19609	29873	33375	36769
East Africa	0.08	0.14		35130	38466	37871
North Africa	0.10	0.17	0.05		61728	62715
Mexico	0.11	0.16	0.02	0.04		74577
Canada	0.09	0.10	0.11	0.13	0.13	

*Weir and Cockerham's unbiased pairwise F_{ST}.

III.1.3. Detection of QTLs for important yield components

III.1.3.1. Phenotypic variation

The descriptive statistics of grain length (Gle), width (Gwi), 1000-grain weight (Wgr) and grain yield (Gyi) revealed a large variation in this collection of hexaploid wheat lines (Table XII). For grain length, the coefficient of variation was 2.54 % with a mean of 3.15 ± 0.08 mm, 2.55 % (1.57 ± 0.04) for grain width, 3.12 % (25.65 ± 0.8 mg) for 1000-grain weight, and 4.31 % (2.55 ± 0.11 t/ha) for grain yield. Broad sense heritability estimates were 90.64 % for grain length, 97.94 % for grain width, 61.6 % for 1000-grain weight and 56.01 % for grain yield. The analysis of variance revealed significant differences among genotypes for all these traits. The effect of environment was no significant for both grain length and width (Table XII).

Table XII: Descriptive statistics, broad sense heritability (h^2) and F-value from analysis of variance for the grain size descriptors

Traits	Unit	Mean±SD	CV%	h^2	F-values		
					Genotype (G)	Environment (E)	GxE
Gle	(mm)	3.2±0.08	2.5	90.6	10.7***	36.9	1.1
Gwi	(mm)	1.6±0.04	2.6	97.9	48.6***	11.5	1.3
Wgr	(g)	25.7±0.8	3.1	61.6	30.9***	15.7**	2.6*
Gyi	(t/ha)	2.6±0.11	4.3	56.0	66.3***	174.9***	2.2*

SD: standard deviation, CV: coefficient of variation, h^2: broad sense heritability, Gle: grain length, Gwi: grain width, Wgr: 1000-grain weight and Gyi: grain yield.

The results of correlation analysis showed that there is a strong positive and highly significant correlation between grain length and width (r = 0.88). In addition, positive and highly significant correlations were identified between grain yield and length (r = 0.509), width (r = 0.535) and weight (r = 0.957), indicating that these traits are the components of grain yield in wheat (Table XIII).

Table XIII: Correlation between seed traits and yield

	Grain length	Grain width	1000-grain weight	Grain yield
Grain length	1			
Grain width	0.880**	1		
1000-grain weight	0.567**	0.574**	1	
Grain yield	0.509**	0.535**	0.957**	1

Furthermore, the distribution of grain length and width shows a bimodal distribution, indicating that the genetic basis of these traits is not controlled by many genes (Fig. 33). In this collection, we found that wider grains were completely separated from shorter, just as the longer distinguished themselves from the narrower ones (Fig. 34).

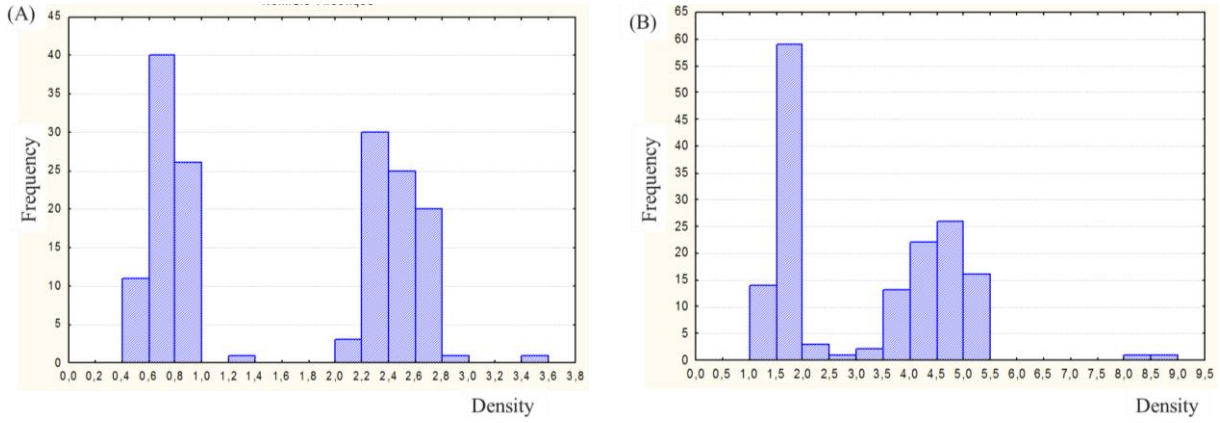


Fig 33. Bimodal distribution of grain width (A) and length (B)

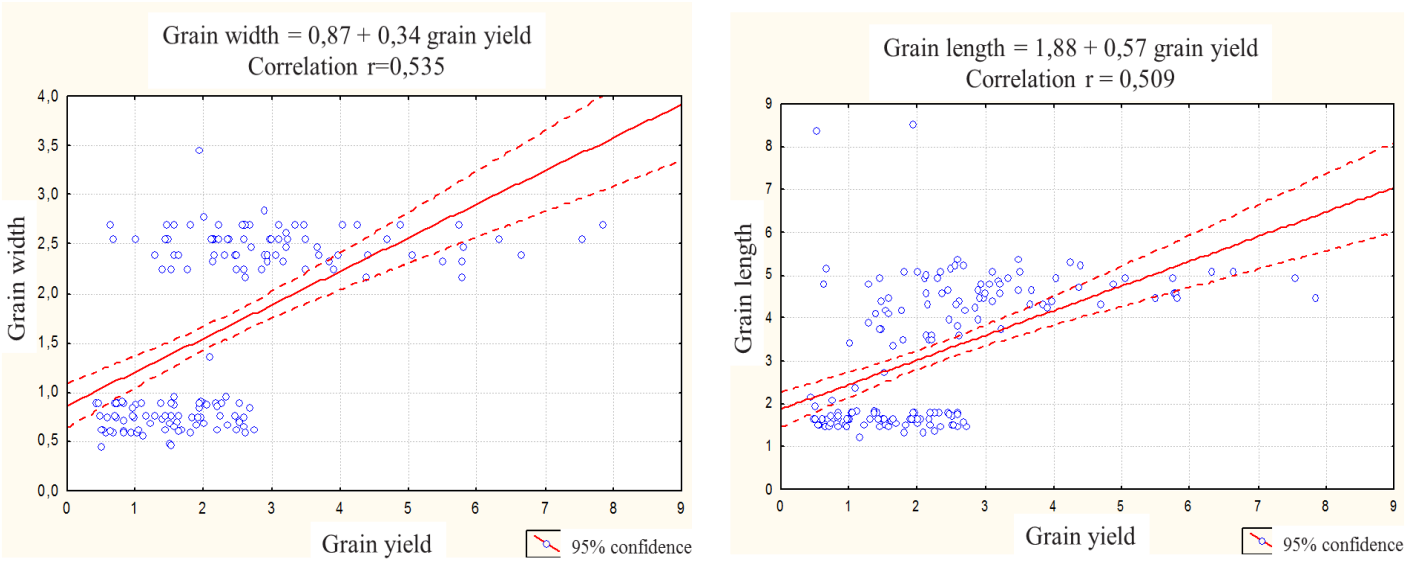


Fig 34. Linear regression between grain size traits

III.1.3.2. Molecular diversity and analysis of population structure

Genetic diversity and population structure analyses were performed with 73,784 SNPs markers distributed over the entire hexaploid wheat genome. To investigate population structure, we used the fastSTRUCTURE software on the entire population of 170 individuals based on 73784 SNPs, with increasing numbers of assumed subpopulations (K) from 1 to 12. The complexity model that best explained the structure for this population was model K = 6 (Fig. 35) and these clusters were found to correspond to the country of origin of the lines. The number of wheat accessions in each of the six subpopulations ranged from 8 to 46. Maximum numbers of accessions were found in Mexico 1 (46) and minimum were observed in East and Central Africa (8).

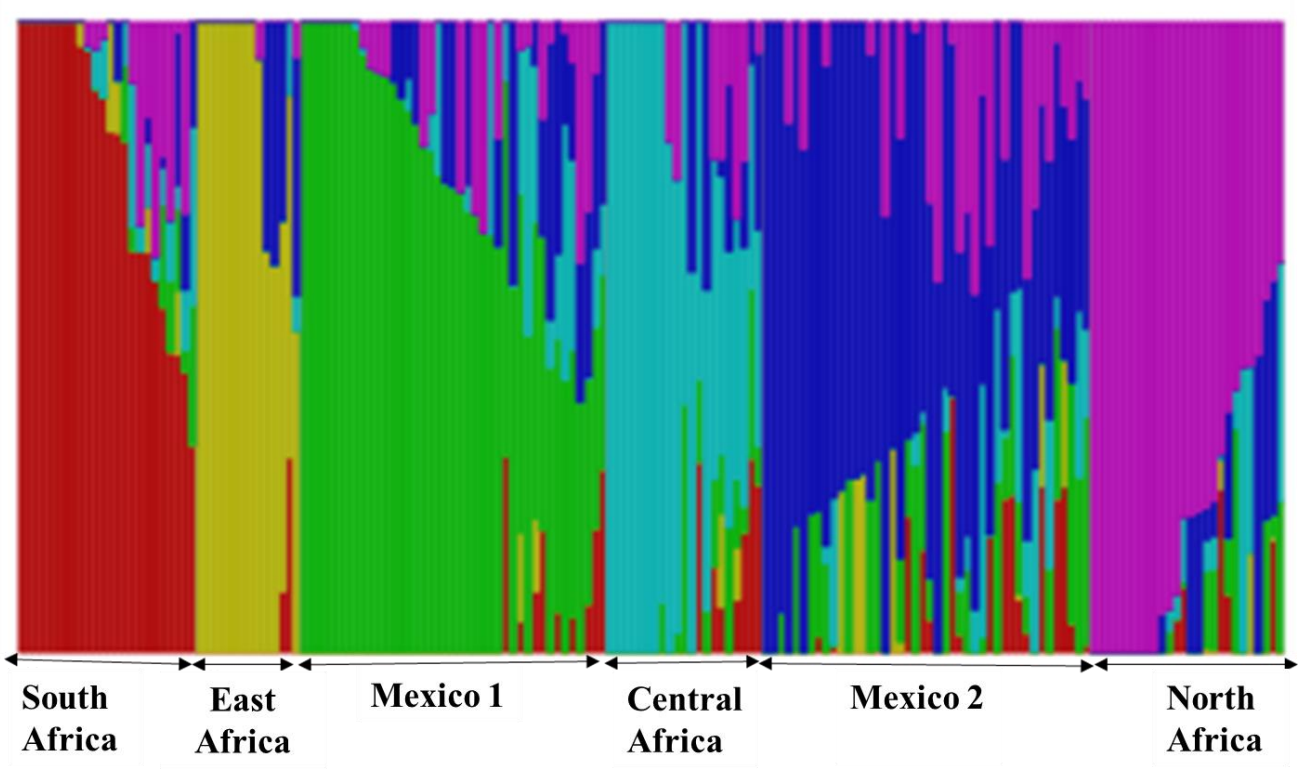


Fig 35. Population structure of 170 hexaploid wheat cultivars, where each vertical line represents a cultivar and each color, a separate subpopulation.

III.1.3.3. Marker–trait associations for grain size in bread wheat

A genome-wide association conducted using 170 hexaploid wheat accessions and 73,784 SNPs markers detected a total of seven SNPs on chromosomes 2D, 1D and 4A were found to be significantly associated with grain length and/or width (Table XIV). The Q–Q plots illustrating observed associations between SNPs and traits (grain length and width) contrasted to expected associations after considering for population structure and line relatedness are represented (Fig. 36C).

For grain length, five (5) significant associations on chromosome 2D and one (1) on 1D were mapped through genome wide association analysis with the most significant marker-trait associations on 2DS (Fig. 36A). For grain width, the most significant association was mapped on 2DS with a total of five (5) associations on 2D, one (1) on 1D and one (1) on 4A detected over the threshold value (Fig. 36B).

Table XIV. Details of loci associated with grain size traits identified via a genome-wide association study in hexaploid wheat

Traits	Loci	Chr	Position	P,value	MAF	R^2	Allelic effect	FDR
Grain length	chr2D:452812899	2D	452812899	6.42E-07	0.304	0.406	-0.27	0.038
	chr2D:403935865	2D	403935865	1.31E-06	0.288	0.398	0.265	0.038
	chr2D:444560418	2D	444560418	2.08E-06	0.275	0.394	-0.269	0.038
	chr2D:452644656	2D	452644656	2.08E-06	0.275	0.394	-0.269	0.038
	chr2D:442798939	2D	442798939	3.25E-06	0.282	0.389	-0.258	0.041
	chr1D:166874041	1D	166874041	3.34E-06	0.294	0.389	0.253	0.041
Grain width	chr2D:452812899	2D	452812899	7.03E-06	0.304	0.551	-0.118	0.332
	chr2D:403935865	2D	403935865	1.29E-05	0.288	0.546	0.116	0.332
	chr2D:444560418	2D	444560418	3.44E-05	0.275	0.539	-0.115	0.423
	chr2D:452644656	2D	452644656	3.44E-05	0.275	0.539	-0.115	0.423
	chr1D:166874041	1D	166874041	3.29E-05	0.294	0.539	0.111	0.423
	chr4A:713365388	4A	713365388	1.35E-05	0.142	0.546	0.131	0.332

Chr: Chromosome, MAF: Minor Allele Frequency, R square of model with SNP, FDR: False Discovery Rate

In the interval of «442798939 – 452812899 pb » region of chromosome 2D, harbored significant QTLs for both grain length and grain width. Similarly, position 166,874,041 bp on chromosome 1D also revealed a QTL for both traits while at position 713365388 bp on chromosome 4A, we found a QTL only for grain width. All of the six (6) SNPs marker associations on chromosomes 2D and 1D appeared to overlap with grain length and width.

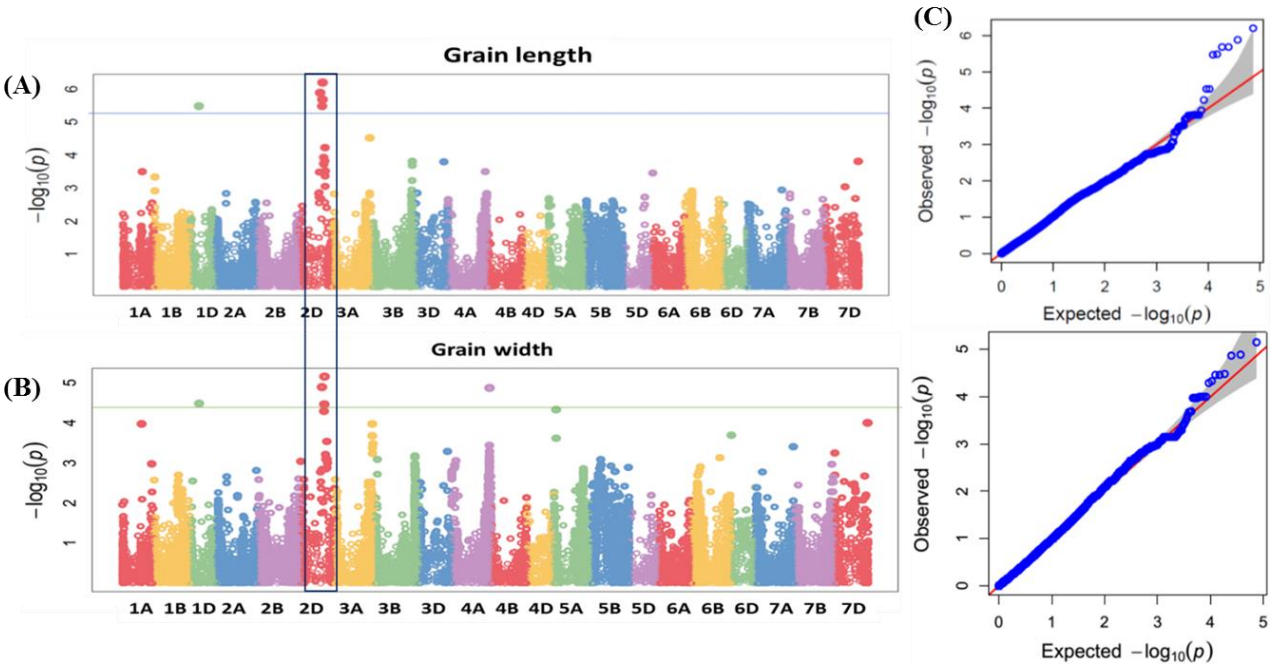


Fig.36. Genome-wide association studies of grain traits in hexaploid wheat.

Manhattan plots for grain length (A) and grain width (B). Negative log10 (P) values from a genome-wide scan are plotted against position on each of 21 chromosomes. The Q–Q plots illustrating the observed associations between SNPs and traits (C).

III.1.3.4. Candidate gene detection for grain size

To identify a candidate gene that can control the grain size (grain length and width) phenotype in the studied wheat collection, we investigated genes included in the QTL interval identified by the GWAS analysis located on chromosome 2D (Fig. 37). In this interval (chr2D: 423364963:425452211), a total of 33 genes (including 15 and 18 high- and low-confidence genes respectively) were observed. Of these, the most promising candidate gene appears as the TraesCS2D01G331100 gene that is highly conserved to the rice CYP724B1 protein gene commonly known as the *D11* gene. The *D11* gene was previously reported as involved in the regulation of grain size in rice due to his role in brassinosteroid biosynthesis. Brassinosteroids are a type of plant hormone and are key regulators of plant growth including seeds by promoting the expansion and elongation of cells in conjunction with auxin hormone.

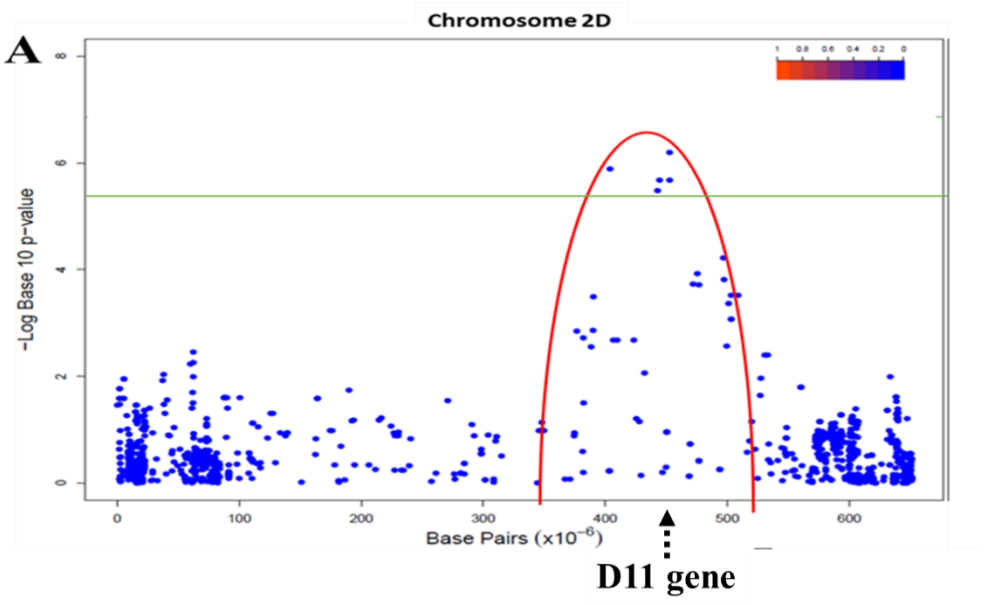


Fig. 37. Candidate gene (D11) in associated region on chromosome 2D

III.1.3.5. Characterization of haplotype around the D11 gene

Our automated approach for haplotype identification narrowed the polymorphism surrounding the D11 locus to only two SNP markers that defined 3 different haplotypes where individuals of haplotype A carry alleles AT, those of haplotype B appears with alleles CT while those of the haplotype C carry alleles CC (Fig. 39).

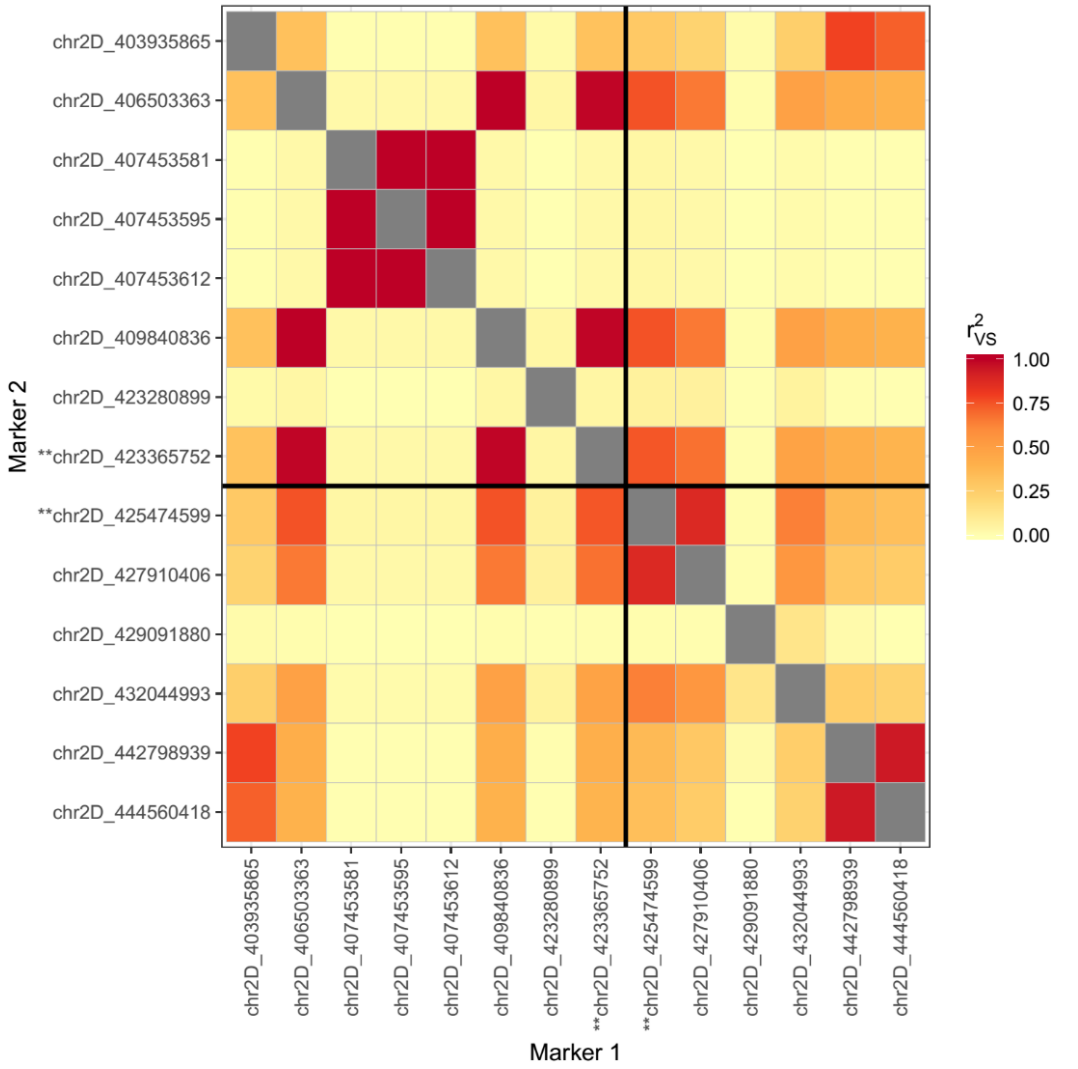


Fig. 38. Linkage disequilibrium plots across wheat lines showing SNPs markers around the D11 gene

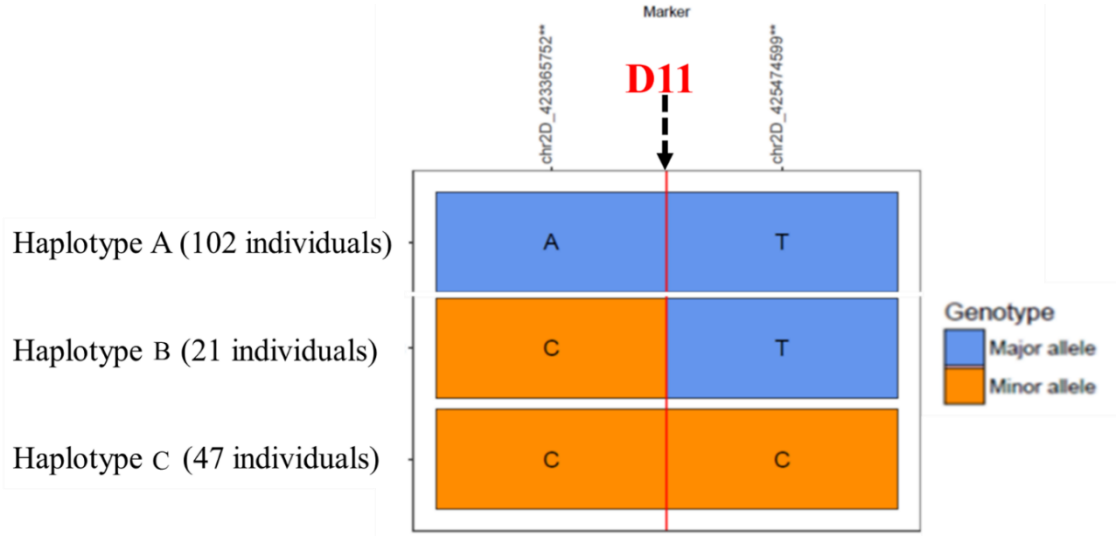


Fig. 39. Three haplotypes around the D11 gene. Individuals of haplotype A carry alleles (A, T); haplotype B with alleles (C, T) and the haplotype C with alleles (C, C).

III.1.3.6. Grouping of haplotypes using phenotypic traits

By comparing the phenotypes of members of the D11 haplotypes A (102 individuals), B (21 individuals) and C (47 individuals), relatively to grain yield and its components, we highlighted highly significant differences (Fig. 40). For grain length, significant differences were observed between all haplotypes, and haplotype A had the individuals with the longest grains averaging 3.76 mm. Furthermore, as haplotype A is also statistically different from haplotypes B and C for grain width and yield, we concluded that individuals of haplotype A are the highest yielding (2.61 t / ha) and produce the longest (3.76 mm) and widest (2.03 mm) grains. This information will be very useful in breeding programs to increase wheat productivity.

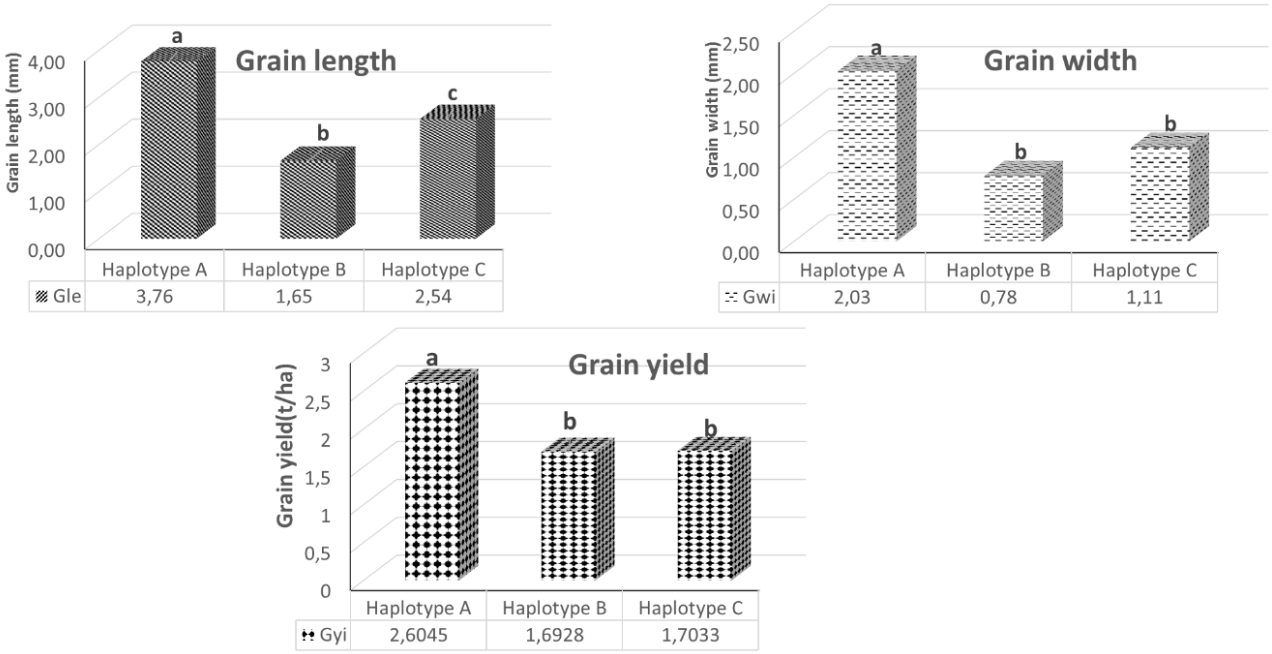


Fig. 40. Histograms showing the impact of haplotypes on the grain yield and its components.
Gle: grain length, Gwi: width, Wgr: 1000-grain weight and Gyi: grain yield

III.1.4. Ultra-Dense SNP Genotyping to discover new genomic sources for tolerance to acid soil with aluminum toxicity in wheat.

III.1.4.1. Performance of wheat lines for acid soil tolerance in greenhouse

To evaluate the performance of 45 wheat genotypes on acid soil with aluminum toxicity (pH = 4.3, Al = 50mg / L) and non-acidic (pH = 5.7, Al = 0mg / L), 6 variables were measured: Plant height (PH), Fresh biomass shot (FBS), Fresh biomass root (FBR), Dry biomass shoot (DBS), and Dry biomass roots (DBS).

The results of analysis of the variances showed highly significant differences between genotypes across acidic and non-acidic treatments for variables PH, RL and FBS (Table XV). The treatment effect was highly significant in all variables. The repetition effect was not

significant for all variables. The effect of the Genotypes-treatment interaction was significant for FBS.

Broad sense heritability estimate varied from 31.58% (DBR) to 70.31% (PH). Plant height had the high heritability (70.31%) and the root length 'heritability was 53.77% (Table XV).

Table XV. Analysis of variances of the phenotypic traits for acid soil tolerance

SV	df	PH	RL	FBS	FBR	DBR	DBS
Genotype (G)	44	3.26**	155.9**	5.18**	2.26	0.19	1.97
Treatments (T)	1	1715.35**	31042.3**	1747.53**	431.91**	28.5**	284.9**
G x T	44	1.53	0.94	5.53**	2.53	0.21	2.01
Repetition (R)	2	2.26	128.98	3.76	15.31	3.65	0.74
G x R	44	1.34	69.52	2.16	1.01	0.16	1
Model GML		21.64**	449.88**	25.15**	7.27**	0.52**	5.21**
Error		0.968	72.07	2.01	1.34	0.25	1.04
h^2		70.31	53.77	61.20	40.71	31.58	47.21

** significant at P<0.001; SV: Sources of variation; df: degree of freedom; h^2; Heritability, PH: Plant height; RL: Root length; FBS: Fresh Biomass Shoot; FBR:Fresh Biomass Root; DBR:Dry Biomass Root; DBS:Dry Biomass Shoot.

The root growth rate of all genotypes was reduced with the addition of Al to the pots and the Al-sensitive and Al-tolerant wheat genotypes were clearly identified. Genotypes with intermediate Al-tolerance levels showed variable root lengths in response to Al stress.

A highly significant variability in root length (RL) was observed in the wheat genotypes evaluated (Table XVII). Indeed, in an acid treatment, the RL varied from 2 cm (Tugela-DN) to 26.40 cm (misr2) with an average of 11.25 cm, and a highly significant difference between genotypes. On the other hand, in non-acid treatment, the RL varied from 15.5 cm (GAMTOUS) to 46,7cm (misr2) with an average of 31.84 cm (Table XVII).

The tolerance index was computed as the ratio of the measured traits under acid versus non-acid soil conditions, so that to identify the tolerant genotypes. So, the standard checks showed the tolerance index of 70.1% (Atlas66), 47.9% (Chinese spring) and 23.6% (Betta). These differences were clearly visible in these control varieties (Fig. 42).

The tolerance index varied from 6% (Tugela-DN) to 72.6% (SST867) with an average of 34.9%. Of the 45 genotypes evaluated, 18 were selected for their high tolerance to acidic soils, relative to the three standard checks (Fig. 41).

Table XVI. Pearson's correlation coefficient between phenotypic traits

	PH	RL	FBS	DBR	FBR	DBS
RL	0.674**					
FBS	0.475**	0.195				
DBR	0.486**	0.269	0.903**			
FBR	0.413**	0.244	0.917**	0.765**		
DBS	0.515**	0.347*	0.814**	0.826**	0.903**	
CIT	0.112*	0.36*	-0.095	0.034	-0.095	-0.006

**. The correlation is significant at the 0.01 level; *. The correlation is significant at the 0.05 level; PH: Plant height; RL: Root length; FBS: Fresh Biomass Shoot; FBR: Fresh Biomass Root; DBR:Dry Biomass Root; DBS:Dry Biomass Shoot; CIT: Citrate

The significant positives correlations were found between plant height and all others traits (Root length, Fresh Biomass Shoot, Fresh Biomass Root, Dry Biomass Root, Dry Biomass Shoot and Citrate effux) (Table XVI). The highest significant correlation was find between Root length and plant height (0.674, P<0.01).

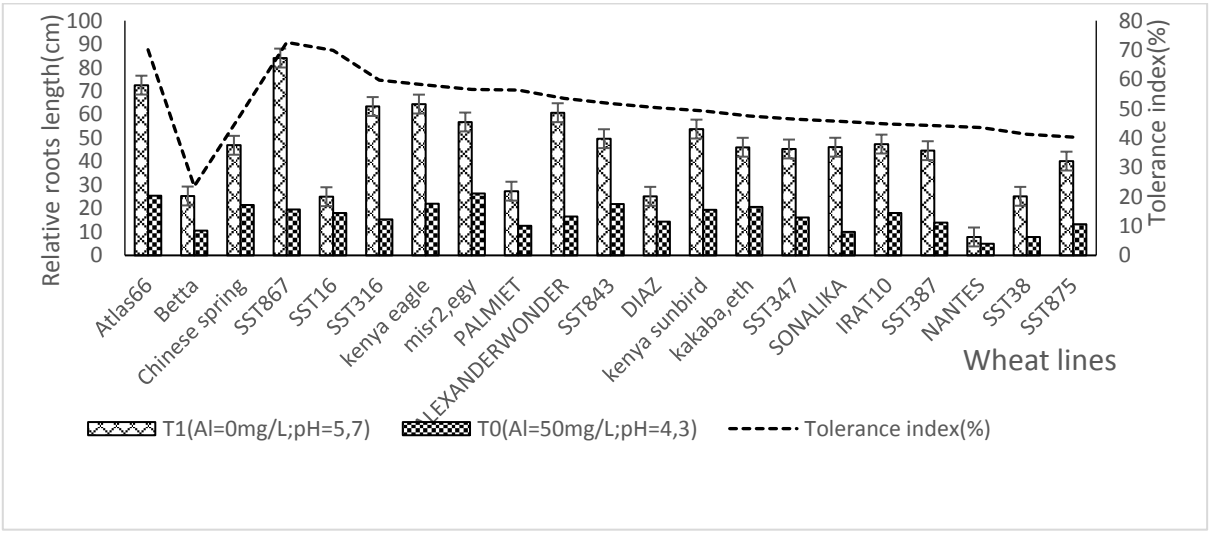


Fig.41. Relative roots length of best wheat lines for Al tolerance

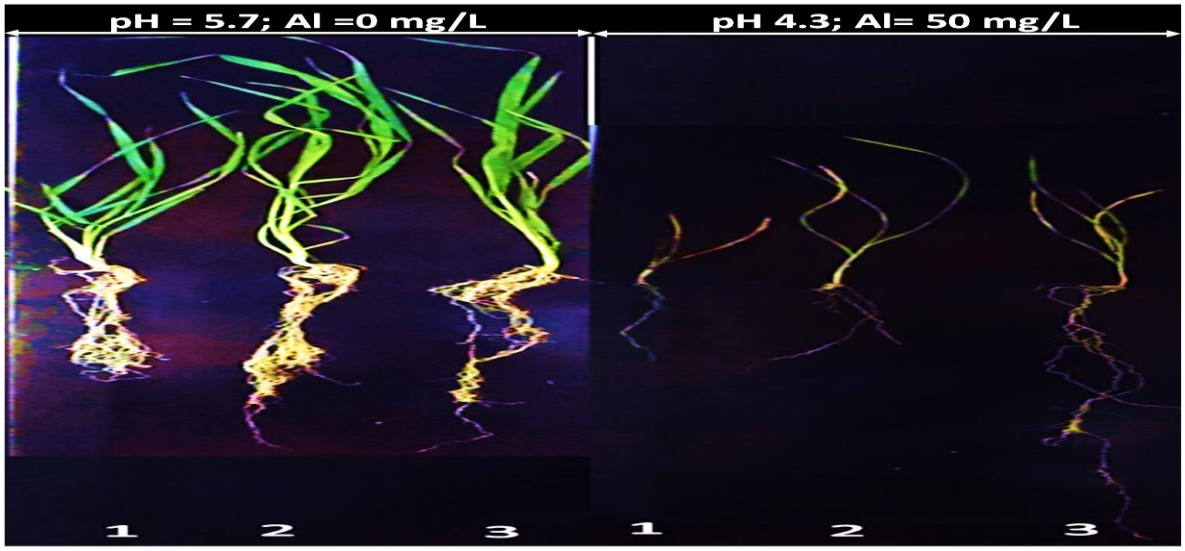


Fig.42. Tolerants and sensitives standard checks of wheat lines grown on acid soil with Al toxicity. 1-Betta; 2-Chinese spring; 3-Atlas 66

III.1.4.2. Physiological response of genotypes for acid soil tolerance

Of the two genes mainly controlling the Al tolerance, TaMATE1 reportedly responds to Al stress based on citrate efflux. So, the organic anion efflux from excised root apices was measured on individual seedlings using as few as four apices. An important variability of citrate efflux was observed in the 45 genotypes evaluated under acidic condition with aluminum toxicity. Indeed, citrate concentration values in genotypes varied from 1.9umol/apex/2h (SST3149) to 9.4umol/apex/2h (SST3156) with a general average of 6umol/apex/2h. The standard checks showed the citrate efflux concentration of 8.6 µmol/apex/2 h (Atlas66), 7.66 µmol/apex/2 h (Chinese spring) and 6.1 µmol/apex/2 h (Betta) (Table XVI). Among the evaluated genotypes, the 18 best lines retained for their high tolerance index also showed high concentrations of efflux citrate (Fig. 43), confirming once more time that the mechanism of Al resistance acting in Al-resistant wheat varieties is exclusion of Al from the root apex, the site of Al phytotoxicity**.**

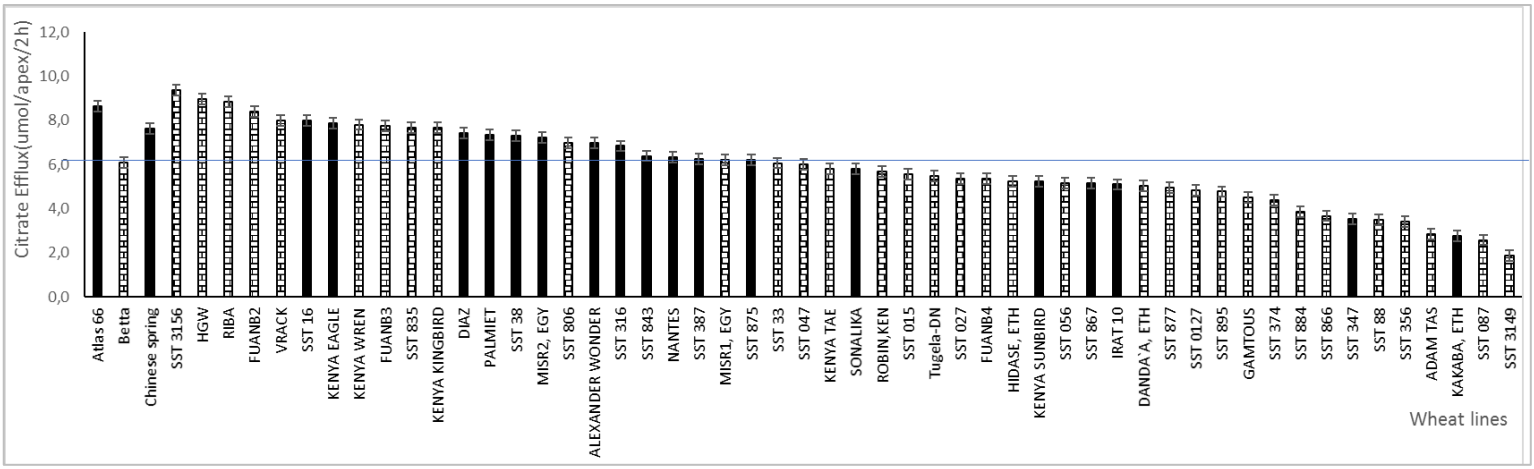


Fig.43. Citrate efflux from excised root apices of evaluated wheat lines for Aluminium tolerance. Top Al-tolerant lines released high amount of citrate efflux (above the black line) comparatively to the Al-sensitives.

Table XVII. Mean performance of wheat lines evaluated under acid soil with aluminum toxicity conditions.

Genotypes	PH(cm)		RL(cm)		Relative root growth (%)	Citrate Efflux (umol/apex/2h)
	T1	T0	T1	T0		
SST867	6.6	1.5	26.9	19.5	72.6a	5.2
Atlas66	7.3	2.4	36.5	25.6	70.1ab	8.6
SST16	4.8	3.3	24.4	18	69.9abcd	8
SST316	7.3	1.4	25.7	15.3	59.7abcd	6.8
kenyaeagle	6.7	2.4	37.8	22	58.1abc	7.9
misr2,egy	6.9	2.8	46.7	26.4	56.6abcd	7.2
PALMIET	4	2.2	19.9	12.6	56.3abcd	7.4
alexander wonder	7.3	2.6	30.9	16.5	53.5abcd	7
SST843	8.4	3	42.3	21.9	51.7abcd	6.4
DIAZ	5.7	2	26.3	14.4	50.4abcd	7.4
kenyasunbird	6.8	1	39.4	19.4	49.3abcd	5.2
Chinesespring	8.3	2.3	44.8	21.5	47.9abcd	7.6
kakaba,eth	8	1.3	43.6	20.7	47.5abcd	2.8
SST347	7.7	2.1	34.8	16.2	46.4abcd	3.5
SONALIKA	4.7	2.1	21.8	10	45.6abcd	5.8
IRAT10	8.9	3	40.5	18.1	44.7abcd	5.1
SST387	5.3	1.1	31.6	14	44.2abcd	6.3
NANTES	4.9	1.2	20.3	7.9	43.5abcd	6.3
SST38	4.1	1.2	18.9	7.8	41.3abcd	7.3
SST875	7.7	0.8	32.9	13.3	40.3abcd	6.2
FUANB2	7.9	2.6	33.4	13.2	39.4abcd	8.4
danda`a,eth	6.8	2	41.2	16.1	39.1abcd	5
kenyawren	6.4	1.5	31.4	11.8	37.7abcd	7.8
SST056	6	1.4	26.6	9.9	37.4abcd	5.2
RIBA	9	1.4	36.4	13.4	36.9abcd	8.8
FUANB3	7.7	2.3	32.2	11.5	35.8abcd	7.7

Toto1	7.3	2.2	33.5	11.4	34abcd	8
SST895	6.5	1.3	36.7	12.3	33.6abcd	4.8
kenyatae	5.7	1.3	30.5	9.8	32.2abcd	5.8
SST866	6	1.1	29.1	8.4	28.8abcd	3.7
SST884	6.3	1.5	32.1	9.2	28.6abcd	3.9
kenyakingbird	6.6	0.8	33.3	9.4	28.3abcd	7.7
SST015	5.4	1.7	29.1	8	27.5abcd	5.6
SST3156	8	1.5	38.1	10.3	27abcd	9.4
SST374	7.8	1.8	28	7.4	26.4abcd	4.4
SST835	5.8	1.3	36.9	9.1	24.7abcd	7.7
SST047	4.7	1	22.6	5.6	24.6abcd	6
SST087	5.6	1.1	29.1	7.1	24.3abcd	2.6
SST877	6.6	1.2	29.4	7.1	24.1abcd	5
Betta	7	0.6	44.6	10.5	23.6abcd	6.7
HGW	7.8	2.5	30.8	7.2	23.4abcd	9
robin,ken	6.5	1.6	40.1	8.4	21.1bcd	5.7
SST3149	7.2	0.8	27.8	5.7	20.6bcd	1.9
SST806	6	1.3	35.5	7.3	20.6bcd	7
GAMTOUS	4.3	1.3	15.5	3	19.4bcd	4.5
SST356	6.6	0.5	32.9	6.3	19bcd	3.4
FUANB4	6.3	2	17	3	17.4bcd	5.4
SST88	5.9	1.2	31.3	5.5	17.4bcd	3.5
hidase,eth	6.6	1	31.3	5.2	16.6bcd	5.2
SST027	4.7	0.7	29.2	4.8	16.5bcd	5.4
ADAMTAS	3.9	1.2	24.6	4	16.1bcd	2.8
SST33	5.5	1.6	27.3	4.3	15.7bcd	6.1
misr1,egy	6.9	0.9	39.6	5.5	13.9bcd	6.2
SST0127	6.1	1	34.9	4.6	13.2bcd	4.8
Tugela-DN	6.2	0.4	33.5	2	6d	5.5
Mean	6.5	1.6	31.8	11.4	34.9	6.0
Minimun	3.9	0.5	15.5	3.0	6.0	1.9
Maximum	9.0	3.3	46.7	26.4	72.6	9.4
SED						

RL: root length; PH: Plant height; T0: non-acid soil (pH = 5.7 with Al =0 mg/L); T1: acid soil (pH 4.3 with 50 mg/L $AlCl_3$).

III.1.4.3. Identification of haplotypes around the MATE1B gene

In order to identify the haplotypes around the MATE1B gene, population structure was evaluated in 45 wheat breeding lines using the software fastSTRUCTURE. The estimated log probability of the data for each k between 1 and 12 increased continuously. To identify the genetic clusters (subpopulation), we ran the software on the entire population of 45 individuals based on 80 124 SNPs markers. The best complexity model component used to explain the structure was K = 9 populations (Fig. 44).

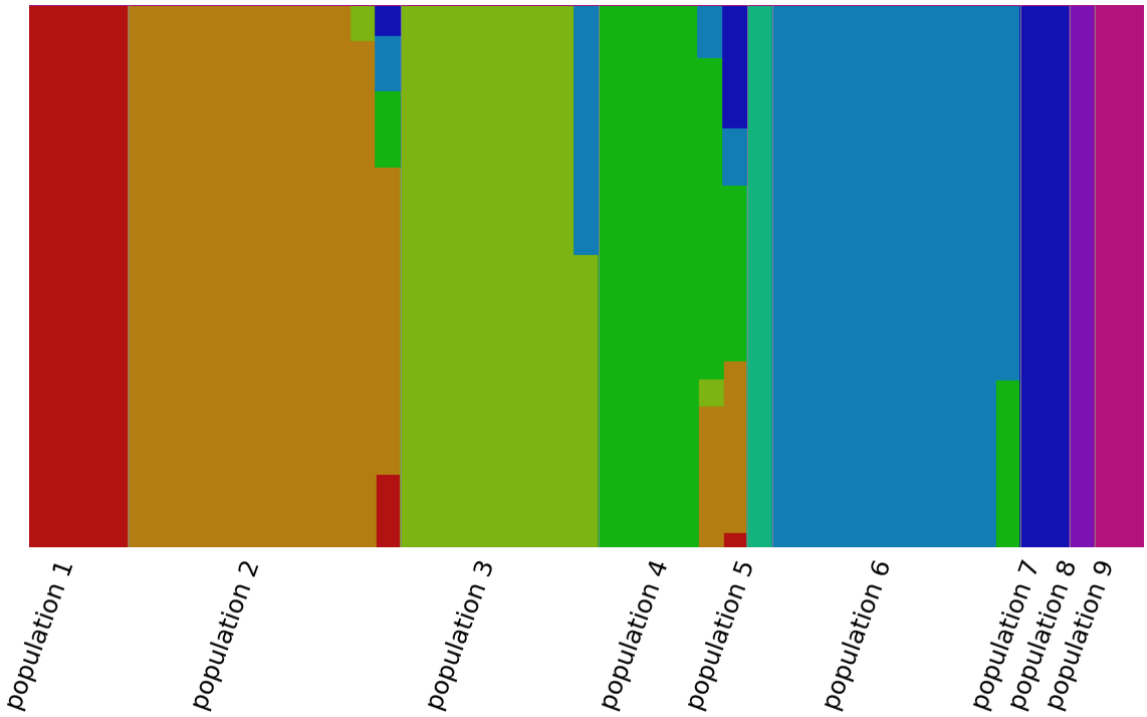


Fig. 44. Population structure analysis of 45 bread wheat lines for aluminum tolerance

To define the genetic distance represented by the haplotype window at the MATE1B locus, we identified four closest neighboring SNPs markers, in linkage disequilibrium (LD) which could be used in the markers assisted selection program. Thus, those four significant SNPs markers were revealed to be in LD and used to define haplotypes potentially having a functional interest (Fig. 45).

Furthermore, those five haplotypes around the MATE1B locus were identified by systematics approaches to classify he individuals according to their level of tolerance to the aluminum toxicity (Fig. 46). The impact of wheat haplotypes around the MATE1B gene on the phenotypic traits confirms the tolerant, moderate and sensitive accessions (Fig. 47).

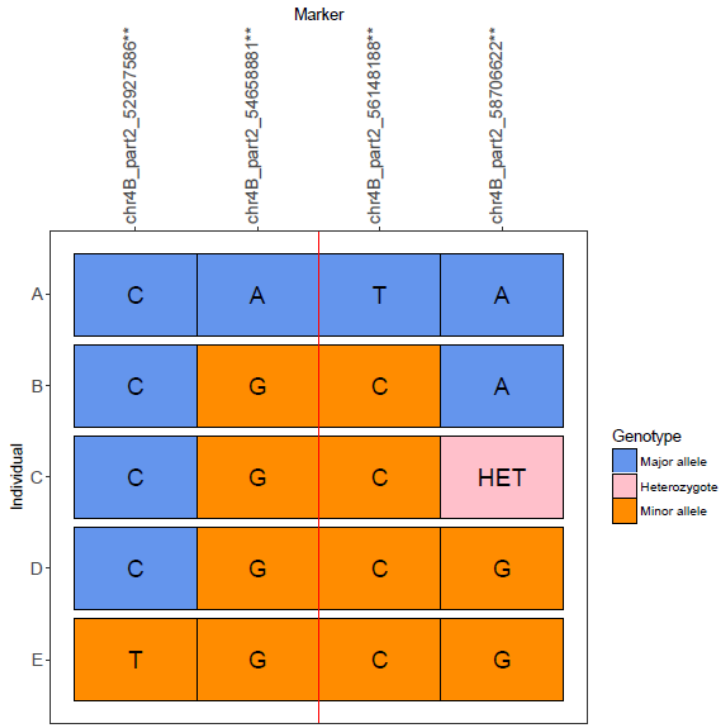


Fig. 45. Alleles of wheat cultivars haplotypes around the MATE1B gene

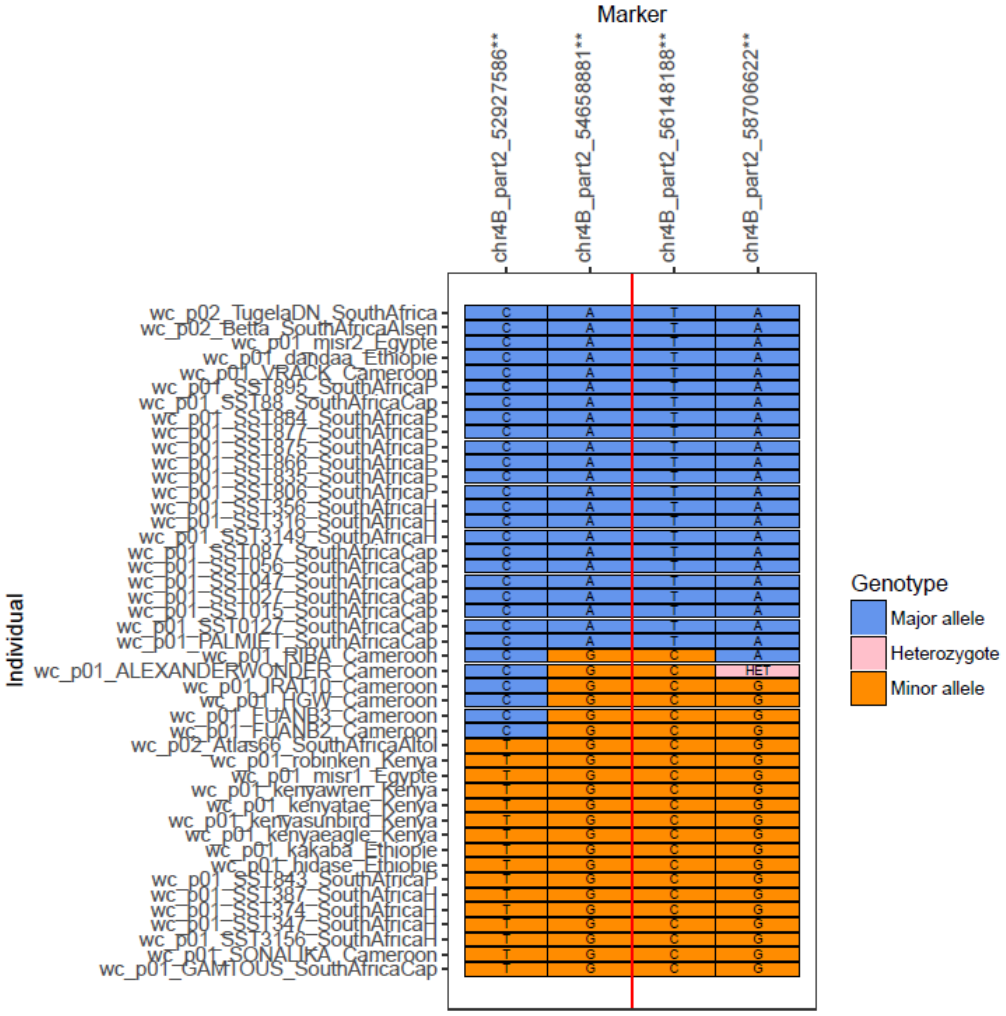
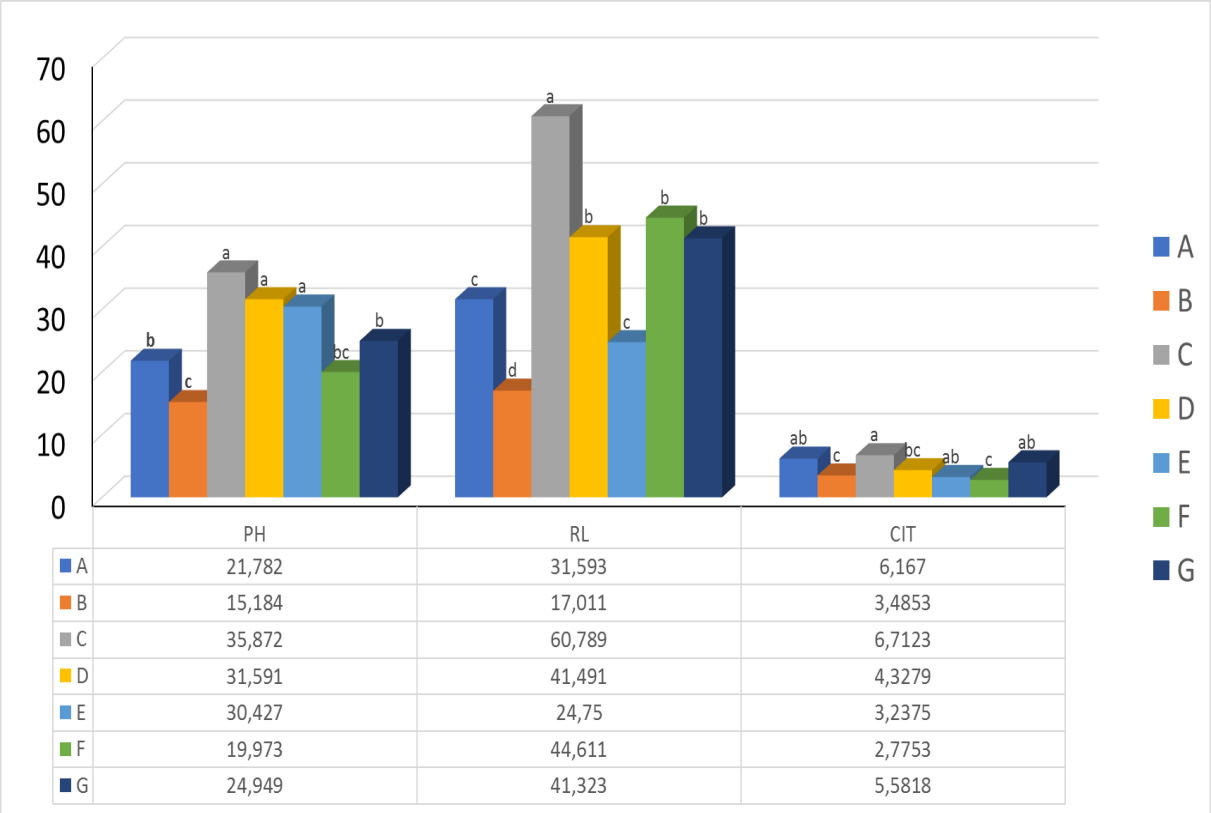


Fig. 46. Groups of wheat lines haplotypes around the MATE1B gene



	PH	RL	CIT
A	21,782	31,593	6,167
B	15,184	17,011	3,4853
C	35,872	60,789	6,7123
D	31,591	41,491	4,3279
E	30,427	24,75	3,2375
F	19,973	44,611	2,7753
G	24,949	41,323	5,5818

Fig. 47. Impact of wheat haplotypes for the MATE1B gene on the phenotypic traits

III.2. Discussion

III.2.1. Genetic diversity in Cameroonian bread wheat cultivars revealed by SSR markers

In our study, 11 microsatellite markers revealing 77 alleles allowed to discriminate 17 cultivars of hexaploid wheat collected in Cameroon. The number of alleles per locus ranged from 2 to 13 with an average of 7. Röder et al. (2002) detected an average of 10.5 alleles per locus from 502 recent European wheat varieties, using 19 microsatellite markers. Khaled et al. (2015) used 17 SSR markers to assess genetic diversity of 33 genotypes of hexaploid wheat from Egypt and detected an average of 5.59 alleles per locus. The average number of alleles per locus in our study is thus comparable to those observed in previous studies. In addition, the microsatellite markers we used had an average PIC value of 0.69, which means that these markers are highly informative in our study. Indeed, Botstein et al. (1980) reported that a PIC value higher than 0.5 is considered to be a sign of a very informative marker, while 0.5> PIC> 0.25 corresponding to an informative marker. In previous studies, Röder et al. (2002) found an average PIC value of 0.67 in 500 genotypes. The choice of these SSR loci is therefore relevant for our study.

Cluster analysis discriminated all cultivars of Cameroonian hexaploid wheat into 5 main groups. Most cultivars clustered according to their geographical location. Indeed, all accessions from groups B and C are cultivated in the Northwest and Adamawa Regions respectively. Similarly, the 4 varieties of group A were collected in the Northwest Region. Moreover, the variety NGDEREM4 of group D comes from Adamawa and the SONALIKA variety of group E was introduced in Cameroon since 1975 through the SODEBLE Company, from Mexico. Huang et al. (2002) reported that the genetic diversity of hexaploid wheat was not completely related to geographic distribution. Then, these authors reported that these results might be explained by the fact that similar genetic variation occurred independently in the different geographic regions or that artificial transfer of accessions from one region to others resulted in a false determination of the geographic origin. Similar results were obtained by Khaled et al. (2015) in hexaploid wheat genotypes grown in Egypt. On the other hand, Al-Khanjari et al. (2007) found that all local varieties of wheat from the same geographical area clustered in the same group. In our case, we can hypothesize that the genetic proximity of the cultivars based on their geographical origin results from a local selection and diversification, coupled with weak or non-existent exchanges of seeds between regions, inducing a geographical structuration and a form of isolation by distance.

`

The overall gene diversity increased with the number of alleles at a given locus. We found significant correlation between gene diversity and the number of alleles in Adamawa ($r = 0.88$, $P < 0.01$) and Northwest ($r = 0.76$, $P < 0.01$). Therefore, the number of alleles could be used for the assessment of genetic diversity in hexaploid wheat. Similar results were found by Huang et al. (2002) in a set of 24 microsatellite markers used to characterize 998 accessions of hexaploid wheat germplasm. So, those authors reported that the characterization of a reliable correlation coefficient needs a large sample size. The results in the present study disagrees with those reported by Prasad et al. (2000) who indicated that the Polymorphism Information Content value was not correlated with the number of alleles in 55 wheat accessions. According to Huang et al. (2002), the number of alleles was also correlated with the repeat number of microsatellite DNA and its relative distance from the centromere. It has been suggested that the three mechanisms for creating a new allele at SSR loci are slippage replication (Tachida and Iizuka 1992), unequal crossing-over and genetic recombination (Harding et al. 1992).

The value of Genetic distance (GD) indicated that some accessions were closely related. Averages of GD over accessions in all Regions were ranged from 0.18 (between Wassande2 and NGDEREM3) to 1 with a mean of 0.8 (80%). The high GD coefficient values indicate the presence of high gene diversity in the accessions.

The mean of gene diversity was relatively higher in Northwest (0.62) compared to Adamawa (0.54). These results suggested that the Northwest exhibited greater genetic diversity than Adamawa region. The Northwest was the presumed center of origin of hexaploid wheat in Cameroon and Adamawa was the site were the SODEBLE was established.

The results obtained in our study provided new information on the relationships between the Cameroonian bread wheat cultivars. The set of the used microsatellite markers showed a high level of polymorphism and sufficient information to discriminate the cultivars of hexaploid wheat grown in Cameroon. Overall our study provides a first description about the molecular genetic diversity of Cameroonian wheat varieties. The results are consistent with expectations and provide a first base for further investigations. The important level of the genetic diversity reported in the present study should be taken into account in developing wheat breeding programs in agro-ecological zones of Cameroon. Morphological and phenotypic studies will also be required to couple our results of molecular analyzes.

III.2.2. Characterization of a global wheat collection using Genotyping-by-Sequencing

Using the Genotyping-by-Sequencing (GBS) protocol to characterize six wheat breeding lines populations, we were able to generate 96 667 polymorphic SNP markers mapped across wheat

chromosomes. These results are compared with SNP number generated previously by others genotyping technologies such as the Illumina iSelect 90k wheat array (Wang et al., 2014; 46 977 mapped markers) and the Affymetrix Axiom_ HD Wheat Genotyping Array (Winfield et al., 2015; 56 505 mapped markers). Furthermore, we found a high level of agreement (>95%) between SNP calls made at the same loci using GBS and SNP array (wheat 90K array) genotyping methods, indicating the high quality of the wheat GBS data.

The levels of genetic diversity in the wheat genomes showed that the majority of mapped markers were located in the B (47215) and A (38481) genomes. Only 10971 of SNP markers mapped to the D genome. This lack of genetic diversity in the D genome of hexaploid wheat cultivars is a well-documented phenomenon attributed to the genetic bottleneck experienced during the initial hybridization to create the hexaploid and the subsequent limited gene flow into bread wheat from *A. tauschii* compared to that from tetraploids (*T. turgidum*) into the A and B genomes (Dvorak, 2006, Halloran et *al*., 2008). Allen et al. (2017) reported the same results, reflecting in the analysis of the average MAF of A, B and D genome markers, where the synthetic hexaploid lines (bred to specifically increase D genome diversity) screened showed higher MAF in the D genome compared to the A and B genomes, a trend opposite to that observed in conventional cultivars and landraces.

Screening the GBS with a range of hexaploid wheat lines demonstrated its utility on a wide range of germplasm from different geographical areas in the world. In general, a high number of polymorphisms SNP were shared between collections.
A relationship between polymorphism level and collection size was observed. The genotyping data were further investigated to examine the relationships between diverse collections of this global wheat breeding lines.

Mostly, a high number of shared polymorphisms SNP-markers and low F_{ST} was observed between populations of different geographical origin, showing that there has been an overlap of germplasm used within these breeding programmes. The principal coordinate analysis plots reflect low F_{ST} measures with overlaps in particular between Western and Eastern Canada, Southern Africa, Central Africa; Mexican and North African accessions. Group 1 (Western and Eastern Canada) is unsurprising given the overlap between the geographical origins of these collections.

The group 2 reflect the signature of Southern Africa breeding material. The relationships between the populations overlapping in group 3 are less clear, although climatic conditions

within these countries are similar, making the exchange of adapted germplasm conceivable thanks to the efforts of the International Center for Agricultural Research in the Dry Areas (ICARDA). Group 5 and 6 reflects the significant impact the International Maize and Wheat Improvement Center (CIMMYT) developed lines have had on breeding programmes. It has been reported that during the 20th century, the global community of wheat breeders freely shared genetic materials (Kronstad, 1997), particularly in efforts led by CIMMYT and ICARDA.

III.2.3. Identification of QTLs for essential yield component in wheat

In the present study, a genome-wide association study was conducted on an international collection of hexaploid wheat to uncover genomic regions controlling variation for grain size. Accessions showed a significant level of variation for grain length, width, thousand-seed weight, and yield with high heritability. Significant positive correlation of grain length, width and weight with grain yield indicated that increase in grain length and width both contribute to enhanced grain yield. Our results showed that grain width comparatively had more positive impact on grain weight and yield than grain length. More recently, Arora et al. (2017) reported also a more positive correlation impact on grain weight than grain length in a collection of *Aegilops tauschii*. They pointed out that correlations between these characters points to a causal relationship between grain size and weight as longer and broader grains are able to accumulate more starch and hence have more grain weight. Previously, Rasheed et al., (2014) mentioned that a moderate to strong correlations between grain weight and size have been reported in a collection of synthetic hexaploid wheats. In fact, Simmonds et al. (2016) noticed that both grain width and length, induce the increase in thousand grain weight due to a mutation in TaGW2-A1 in tetraploid and hexaploid wheat through wider and longer grains.

In the present study, seven SNPs were found to be associated with both traits, grain length and width identifying three quantitative trait loci (QTLs) located on chromosomes 1D, 2D and 4A. Furthermore, Yan et al. (2017) indicated that the grain size and shape represents a classical example of traits with variations that happen after polyploidization and domestication in wheat. A long, thin primitive grain was transformed into a wider, shorter modern grain during wheat domestication, showing that grain shape became rounder during wheat domestication (Gegas et al., 2010). A noteworthy aspect of polyploid wheat evolution is genomic asymmetry in the control of grain shape, and that is the predominant control of grain shape by the A genome (Feldman et al., 2012). However, the large variation in grain size and shape observed among Ae. tauschii (ancestor, donor of D genome) cultivars is retained in the synthetic allohexaploid

wheat using natural tetraploid species as the AB genome donor, suggesting that the D genome partially affects grain size and shape of allohexaploid wheat (Röder et al., 2008; Okamoto et al., 2013; Rasheed et al., 2014). Many studies have identified QTLs for grain size and shape in bread wheat accessions and QTLs have been assigned to several chromosomes (Breseghello and Sorrells 2006, 2007, Gegas et al 2010, Sun et al 2009, 2010, Tsilo et al., 2010, Williams et al., 2013). Okamoto et al. (2013) performed QTL analyzes using four synthetic wheat F2 populations to identify the genetic loci responsible for grain size and shape variation in the hexaploid background. On chromosome 2DS, QTL for variables related to grain length were found in the population Ldn / PI476874 // Ldn / KU-2069. A QTL for grain length (qGL3), which encodes a putative protein phosphatase with a Kelch-like repeat domain (OsPPLK1), was found in the centromeric region of rice chromosome 3 (Qi et al., 2012). Furthermore, genes or QTLs associated with grain shape and size are of interest for domestication and enhancement (Simons et al., 2006, Williams et al., 2013). Many studies have revealed that in the genome-D of common wheat donor, *Aegilops Tauschii* provides a great source of genetic variability and improves agronomic traits (Reynolds et al., 2007, Rana et al., 2013). As a result, many synthetic allohexaploid wheat lines were produced by hybridization of tetraploid wheat and *Aegilops* tauschii, and successfully used in wheat cultivation (Trethowan and Mujeeb-Kazi, 2008, Takumi et al., 2009). At the genomic level, several QTL analyzes were used to identify D-genomic regions of synthetic allohexaploid wheat lines associated with grain size and shape (Okamoto et al., 2013, Yu et al., 2014). Particularly, Tg-D1 on the chromosome 2D is one of the well-known loci that have been recruited for the domestication of wheat grain size and shape. At the allohexaploid wheat speciation, a dramatic change in grain shape occurred due to the mutation in the Tg gluteal toughness gene (Nalam et *al*., 2007, Dvorak et al., 2012). To reveal genetic factors on the D genome that control variation in grain size and shape, Yan et al. (2017) analyzed QTL using F2 and F2: 3 populations derived from a common allohexaploid wheat strain TAA10 and synthetic allohexaploid wheat XX329, which have almost identical AABB genomes and different DD genomes. Based on genotyping using 660K wheat single nucleotide polymorphism (SNP), a total of nine (9) stable QTLs associated with grain size and shape were mapped to chromosomes 2D and 7D, and verified using near isogenic lines (NILs) with allohexaploid synthetic wheat XX329 contributing favorable alleles. These authors identified the candidate genes for grain size (GIF1 and PFPβ), orthologs in rice. However, these authors provide further information to support the genetic factors that shaped grain morphology during the evolution and domestication of wheat, using Microarray data approach in wheat. In addition, recent advances in genomic technologies have allowed for a better understanding of

the genetic basis of variation in large sets of genetic material using Genome-wide Association Study (GWAS). GWAS is one of those approaches that can be used for the identification and high-resolution mapping of genetic variability useful from sets of genetic material that have resulted from many historical recombination cycles (Yu and Buckler, 2006). Arora et al. (2017) conducted GWAS in a collection of *Aegilops Tauschii* accessions for grain length, width, and weight using SNPs markers. Based on Genotyping-by-Sequencing approach, genetic similarity was calculated between accessions and GWAS was performed using 114 non-redundant accessions and 5249 SNPs markers. A total of 17 SNPs associated with granulometric characteristics distributed over all seven chromosomes, with particularly the chromosomes 2D, 5D and 6D, harboring the most important marker-trait associations. Furthermore, some useful genes/QTL controlling desirable traits have been identified on the D genome of synthetic allohexaploid wheat, including disease resistance, abiotic stress tolerance, suitable quality and anti-sprouting (Tadesse et al., 2007; Imtiaz et al., 2008; Li et al., 2012; Ilyas et al., 2015).

Grain length QTL on chromosomes 5D and 6D in *Aegilops tauschii* were observed by Zhao et al. (2015) and on chromosomes 2D, 4D, and 7D by Okamoto et al. (2013) in synthetic wheat-based populations. Interestingly, one pleiotropic locus on chromosome 2DS significantly contributed to the determination of wheat grain shape, which corresponded to that of Tg-D1 (Dvorak et al., 2012). Consistent with the result, Yan et al. (2017) found that a major QTL with an LOD score of higher than 9.8 was located on the short arm of chromosome 2D and the allele from synthetic wheat at the QTL produced longer grains.

Limited understanding of grain size and shape is available in polyploid wheat where QTL for grain size and shape have been identified (Zhang et *al*., 2010; Williams and Sorrells, 2014; Simmonds et *al*., 2016), but no gene affecting grain weight independent of length and width has yet been cloned. This is in contrast to rice where genes with large effects on grain size have been identified revealing an independent genetic control of grain length and width (Weng et *al*., 2008).

In the present study, we found a promising candidate gene (TraesCS2D01G331100) in the vicinity of the peak SNP on chromosome 2D, whose was homologous to the rice *D11* (CYP724B1) gene previously reported as involved in the regulation of grain size in rice (Tanabe et al. 2005). The *D11* gene has been reported to encode a novel cytochrome P450 with homology to enzymes involved in brassinosteroid biosynthesis. These markers will be useful in breeding for enhanced wheat productivity (Tanabe et al. 2005). With orthologous genes controlling similar phenotypes across many grass species, including wheat, rice, barley, and

sorghum, comparative genomics has shown collinearities and provided a powerful tool for gene discovery in wheat (Valluru et *al*., 2014). Furthermore, the peak region of QTgw.cau-2D was syntenic to rice chromosome 4 at 18.49–23.76Mb, and one cloned gene GIF1 (Os04g0413500) encoding a cell-wall invertase required for carbon partitioning during early grain filling, which was located at 20.44Mb on rice chromosome 4 (Wang et *al*., 2008). In wheat though TaGW2 has been shown to be a negative regulator of grain weight as mutant alleles of TaGW2-A1 have been found to increase grain weight by contributing both to length and width (Simmonds et *al*., 2016). Genes controlling seed size have been extensively studied in rice. The characterized genes have been revealed to function in G-protein signaling (Huang et *al*., 2009), or in the ubiquitin–proteasome pathway (Song et *al*., 2007). TaGS5 an ortholog of rice GS5 has been shown to be a positive regulator for grain size in wheat (Ma et *al*., 2016). In rice, a gene GL3.1, regulating grain length and yield was cloned and found to belong to serine/threonine phosphatase of the PPKL family. GL3.1 regulates grain length by mediating cell cycle progression through affecting the phosphorylation status of cell cycle proteins, such as cyclin-T1;3, thereby controlling grain yield (Qi et *al.*, 2012).

III.2.4. Ultra-Dense SNP Genotyping for genomic selection of wheat lines for acid soil under aluminum tolerance

In the present study, the root growth rate of all genotypes was reduced with the addition of Al to the pots and the Al-sensitive and Al-tolerant wheat genotypes were clearly identified. Genotypes with intermediate Al-tolerance levels showed variable root lengths in response to Al stress. Jayasundara et al. (1998) reported that acid soils are phytotoxic as a result of nutritional disorders: deficiency or unavailability of essential nutrients such as Ca, Mg, Mo and P, and toxicity of Al and Mn. The solubility of soil compounds, and hence the availability of nutrients for plants, is related to soil pH. The common responses of plants to Al toxicity are mainly structural and ultra-cellular changes in the leaves, increased rates of diffusion resistance, reduced stomatal opening, reduced photosynthetic activity leading to leaf and size leaf chlorosis and necrosis; and decreased biomass (Thornton et al., 1986; Tekeu et al. 2015). Furthermore, acid soils are generally characterized by by Al^{3+}, Mn^{2+} and H^{+}, with deficiencies of Ca^{2+}, Mg^{2+} and $PO4^{3+}$ (Krstic et al., 2012), low pH (pH <5.5), low levels of organic matter, reduced cation exchange rate and excessive leaching of elements in areas with high rainfall. They may be due to the toxicity of Al, Fe and Mn, and / or they may be caused by the absence or absence of N, P, Ca, Mo, Si and Mg, inhibiting root development and reduction of water absorption (Mossor-Pietraszewska, 2001, Velasquez et al., 2008).

In our present study, we found that in acid soil treatment, the root length varied from 2 cm to 26.40 cm with an average of 11.25 cm, while in the non-acid treatment, the root length varied from 15.5 cm to 46,7cm with an average of 31.84 cm. Reynolds et al. (2001) mentioned that the soil acidity complex creates a chemical barrier that reduces root growth and development; these roots become ineffective at absorbing nutrients and water. Low availability of P and highly exchangeable Al inhibit root growth with concomitant losses in grain yield and quality (Von Baer, 2007). In a neutral or basic environment, Al is insoluble and is in the form of monovalent and divalent oxides or aluminum hydroxide Al(OH)3, and is a normal part of any soil, useful for plants (Harter, 2007). In an acid environment, it is soluble and is in trivalent form with the $Al3+$ ion becoming the dominant species (Krill et al., 2010, Kochian et al., 1995). The $Al3+$ ions disrupt many physiological processes in plants by both apoplastic and symplastic interactions. Al delays the growth of the primary root and inhibits lateral root formation (Krstic et al., 2012). This inhibition of root growth is the main symptom of toxicity in Al (Delhaize and Ryan, 1995, Tabuchi and Matsumoto, 200, Krill et al., 2010). The reduction in root growth and function leads to increased sensitivity to other constraints, mainly mineral deficiencies and drought, due to the limited capacity of Al's poisoned roots to acquire sufficient water and nutrients (Krill et al., 2010). The roots of the plants subjected to Al toxicity have a spatula appearance, the truncated ends turning brown and the root system being entirely coronoid and having several truncated and filiform lateral roots, without shoots. Al toxicity predisposes the plant root to fungal attack (Ota, 1968). In the roots, Al interferes with cell division, reduces DNA duplication by increasing the rigidity of the double helix (Rout et al., 2000). Al binds phosphorus in assimilated form, interferes with certain enzymes governing the deposition of polysaccharides in the cell membrane and increases their rigidity. Young plants are generally more affected by the toxicity of aluminum than older plants (Rout et al., 2000). Al toxicity is the most important factor limiting agricultural productivity on 67% of the total area of acid soils (Eswaran et al., 1997).

In the present study, we observed that among the evaluated genotypes, the best wheat lines retained for their high tolerance index of root length also showed high concentrations of efflux citrate, confirming once more time that the mechanism of Al resistance acting in Al-resistant wheat varieties is exclusion of Al from the root apex by citrate efflux, the site of Al phytotoxicity (Kochian et al., 2015). Of the two genes mainly controlling the Al tolerance, TaMATE1 reportedly responds to Al stress based on citrate efflux. Indeed, the Al tolerance genes belong mainly to the MATE (multidrug extrusion and extrusion of toxic compounds) and

ALMT (aluminum-activated malate transporter) families. MATE genes code for transporters excreting a wide range of metabolites and xenobiotics in eukaryotes and prokaryotes (Weston et al., 2012), and members of the ALMT family code for vacuolar malued channels (Kovermann et al., 2007). In wheat, Al tolerance is mainly controlled by two genes: TaALMT1, which encodes a malate transporter on chromosome 4D, is constitutively expressed on the root apex (Sasaki et al., 2004). TaMATE1 would respond to Al stress based on citrate efflux (Ryan et al., 2009, Kochian et al., 2015).

CHAPTER IV: CONCLUSION AND PERSPECTIVES

IV.1. Conclusion

The present study was investigated to characterize the genetic background and select high-yielding and tolerant wheat varieties to aluminum toxicity.

- Our study revealed that in 17 bread wheat cultivars collected in Cameroon, 11 microsatellite markers showed 77 alleles allowing discriminating those cultivars. The number of alleles per locus ranged from 2 to 13 with an average of 7. The results obtained in our study provided new information on the relationships between the Cameroonian bread wheat cultivars. The set of the used microsatellite markers showed a high level of polymorphism and sufficient information to discriminate the cultivars of hexaploid wheat grown in Cameroon. Overall, our study provides a first description about the molecular genetic diversity of Cameroonian wheat varieties.
- Using the Genotyping-by-Sequencing (GBS) protocol to characterize an international collection of six wheat breeding lines populations, we were able to generate 96 667 polymorphic SNP markers mapped across 21 wheat chromosomes. Furthermore, we found a high level of agreement (>95%) between SNP calls made at the same loci using GBS and SNP Microarray (wheat 90K array) genotyping methods, indicating the high quality of the wheat GBS data. The levels of genetic diversity in the wheat genomes showed that the majority of mapped markers were located in the B (47215) and A (38481) genomes. Only 10971 of SNP markers mapped to the D genome. In general, a high number of polymorphisms SNP were shared between collections. Mostly, a high number of shared polymorphisms SNP-markers and low F_{ST} was observed between populations of different geographical origin, showing that there has been an overlap of germplasm used within these breeding programmes. The principal coordinate analysis plots reflect low F_{ST} measures with overlaps in particular between Western and Eastern Canada, Southern Africa, Central Africa; Mexican and North African accessions. We concluded that the Genotyping-by-Sequencing (GBS) approach was able to rapid and accurate dissect the substructure of the adaptation type of hexaploid wheat accessions in their growing areas;
- To uncover the genomic regions controlling variation for grain size in wheat, a genome-wide association study was conducted on wheat accessions which showed significant level of variation for grain size, 1000-weight, and yield with high heritability. Significant positive correlation of grain length, width and weight with grain yield indicated that increase in grain length and width both contribute to enhanced grain yield. So, the GBS-GWAS approach was able to rapidly and accurately identify SNP markers significantly associated with QTLs (located on chromosomes 1D and 2D) governing grain length and width in wheat, and many of these

were in close proximity to a strong candidate gene (D11), which is known to control grain size in rice. In total, 102 high yielding wheat lines varieties were selected;

- Using an ultra-dense SNP genotyping with phenotypic traits for acid soil tolerance with aluminum toxicity in wheat lines, the root growth rate of all genotypes was reduced with the addition of Al to the pots and the Al-sensitive and Al-tolerant wheat genotypes were clearly identified. We found that in acid soil treatment, the root length varied from 2 cm to 26.40 cm with an average of 11.25 cm, while in the non-acid treatment, the root length varied from 15.5 cm to 46,7cm with an average of 31.84 cm. Furthermore, we observed that among the evaluated genotypes, the best wheat lines retained for their high tolerance index of root length also showed high concentrations of efflux citrate, confirming once more time that the mechanism of Al resistance acting in Al-resistant wheat varieties is exclusion of Al from the root apex, the site of Al phytotoxicity. Using the systematic approach with 80124 SNPs markers, 16 wheat genotypes around the MATE1B gene were selected for acid soil tolerance under aluminum toxicity conditions. Those genotypes could be useful in the wheat breeding program;
- Definitely, we found seven (7) high yielding wheat varieties, which are tolerant to aluminum toxicity (having in the same time the MATE1B and D11 functional genes).

IV.2. Perspectives

The present project allows to point out the level of genetic diversity available in Cameroonian and in a collection of international wheat breeding lines, select high-yielding and tolerant wheat varieties for aluminum toxicity. For an eventual vulgarization, the below perspectives should be considerate:

- The important level of the genetic diversity reported in the present study should be taken into account in developing wheat breeding programs in agro-ecological zones of Cameroon.
-Develop wheat breeding programs in the agro-ecological zones of Cameroon,
-Use the D11 markers and gene identified in genetic improvement programs; Molecular techniques such as CRISPR_cas and qPCR should be used for further investigation on D11 gene.
-Evaluate the seven (7) highly productive wheat varieties tolerant to aluminum toxicity in multi-local and peasant trials for dissemination;
-Develop and maintain genebanks for the conservation of wheat varieties in Cameroon;

REFERENCES

Allen A.M., Barker G.L., Berry S.T., Coghill J.A., Gwilliam R., Kirby S., Robinson P., 2011. Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). Plant Biotechnol. J. 9, 1086–1099.

Allen A.M., Mark O., Winfield A.J. Burridge B., 2017. Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum* L.). Plant Biotechnology Journal 15, pp. 390–401 doi: 10.1111/pbi.12635.

Al-Maskri A.H., Sajjad M., and Khan S. H., 2012. Association mapping: a step forward to discovering new alleles for crop improvement. *Int. J. Agric. Biol*, *14*, 153-160.

Anil R., Matthew S., and Jonathan K.P., 2014. *fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets* , (Genetics), 197:573-589 [Genetics, Biorxiv].

Aniol A., 1990. Genetics of tolerance to aluminum in wheat *Triticum aestivum* L. Thell, Plant Soil 123, 223–228.

Aniol A., Gustafson J.P., 1984. Chromosome location of genes controlling aluminum tolerance in wheat, rye and triticale, Can. J. Genet. Cytol. 26, 701–705.

Arora S., Singh N., Kaur S., Bains N.S., Uauy C., Poland J. and Chhuneja P., 2017. Genome-Wide Association Study of Grain Architecture in Wild Wheat Aegilops tauschii. Front. Plant Sci. 8:886. doi: 10.3389/fpls.2017.00886.

Ayuk-Takem J., 1984. Performance de 12 variétés de blé (Triticum aestivum) en croissance à UPPER FARM (1982 m). Rapport d'activités annuelles de IRA (Institut de la Recherche Agronomique) Bambui. 1982-1986.

Basten C.J., Weir B.S., and Zeng Z.B, 2004. QTL cartographer, version 1.17. *Department of Statistics, North Carolina State University, Raleigh, NC*.

Bastien M., 2013. Étude de la résistance à la sclérotiniose chez le soya. Thèse de Doctorat (Ph.D.) en biologie végétale, *Universite Laval*, 196 p.

Bentley D.R., Balasubramanian S., Swerdlow H.P., Smith G.P., Milton J., Brown C.G., et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53–59. doi: 10.1038/nature07517.

Botstein D., White R.L., Skolnick M., and Davis R.W, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.

Bradbury P.J, Zhang Z., Kroon D.E, Casstevens T.M, Ramdoss Y., Buckler E.S, 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23 (19):2633-2635.

Breseghello F., and Sorrells M. E., 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177. doi: 10.1534/genetics.105.044586.

Breseghello F., and Sorrells M.E., 2007. QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. *Field Crops Res*. 101: 172–179.

Browning S., Browning B., 2007. Rapid and accurate haplotype phasing and missing-data inference for wholegenome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet*.; 81, pp. 1084–97 PMID: 17924348.

Buckler IV E.S., Thornsberry J.M., 2002. Plant molecular diversity and applications to genomics, Curr. Opin. Plant Biol. 5, 107–111.

Catchen J., Hohenlohe P.A, Bassham S., Amores A., Cresko W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol Ecol*; 22(11):3124–40. doi: 10.1111/mec.12354.

Chen Y., 2013. Dissection of Agronomic Traits in Crops by Association Mapping. In *Diagnostics in Plant Breeding* (pp. 119-142). Springer Netherlands.

Cheverud J.M., 2001. A simple correction for multiple comparisons in interval mapping genome scans. Heredity 87 (1):52-58.

Choulet F., Alberti A., Theil S., Glover N., Barbe V., Daron J., et al., 2014. Structural and functional partitioning of bread wheat chromosome 3B. *Science*;345:1249721.

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B., and Pang E.C.K., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement : The basic concepts. *Euphytica*, *142*, 169–196. doi:10.1007/s10681-005-1681-5

Collard B.C.Y., Mackill D.J., Collard B.C.Y., and Mackill D.J., 2008. Marker-assisted selection : an approach for precision plant breeding in the twenty-first century Marker-assisted selection : an approach for precision plant breeding in the twenty-first century. *Phil. Trans. R. Soc. B*, *363*(February), 557–572. doi:10.1098/rstb.2007.2170.

Colorado Wheat, 2013. Why is the wheat genome so complicated? http://coloradowheat.org/2013/11/why-is-the-wheat-genome-so-complicated/

Cuthbert J.L., Somers D.J., Brule-Babel A.L., Brown P.D., and Crow G.H., 2008. Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (Triticum aestivum L.). *Theor. Appl. Genet*. 117, 595–608. doi: 10.1007/s00122-008-0804-5.

Cuthbert P.A, Somers D.J, Brule-Babel A., 2007. Mapping of *Fhb2* on chromosome 6BS: a gene controlling Fusarium head blight field resistance in bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 114 (3):429-437. doi:10.1007/s00122-006-0439-3.

Dai H., Cao F., Chen X., Zhang M., Ahmed I.M., et al., 2013. Comparative proteomic analysis of aluminum tolerance in Tibetan wild and cultivated barleys. *PLOS ONE* 8:e63428.

Danial D., Parlevliet J., Almekinders C. and Thiele G., 2007. Farmers' participation and breeding for durable disease resistance in the andean region, *Euphytica,* 153(3), pp. 385- 396.

Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M., and Blaxter M.L., 2011. Genome wide genetic marker discovery and genotyping using next generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012.

Delaneau O., Marchini J., 2014. The 1000 Genomes Project Consortium. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nature Communications; 5 3934 doi: 10.1038/ncomms4934 PMID: 25653097.

Delhaize E., and Ryan P.R., 1995. Aluminum toxicity and tolerance in plants. *Plant Physiology*, 107**:** 315-321.

Delhaize E., Craig S., Beaton C.D., Bennet R.J., Jagadish V.C., Randall P.J., 1993. Aluminum tolerance in wheat (Triticum aestivum L.): I. Uptake and distribution of aluminum in root apices, Plant Physiol. 103, 685–693.

Dholakia B., Ammiraju J., Singh H., Lagu M., Röder M., Rao V.S., et al., 2003. Molecular marker analysis of kernel size and shape in bread wheat. Plant Breed. 122, 392–395. doi: 10.1046/j.1439-0523.2003.00896.x

Dvorak J., Deal K.R., Luo M.C., You F.M., Von Borstel, K., and Dehghani H., 2012. The origin of spelt and free-threshing hexaploid wheat. J. Hered. 103, 426–441. doi: 10.1093/jhered/esr152.

Dvorak. J., Akhunov E.D., Akhunov A.R., Deal K.R. and Luo M.C., 2006. Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. Mol Biol Evol. 23, 1386–1396.

Eeuwijk F.A. Van Bink M.C.A.M., Chenu K., and Chapman S. C., 2010. Detection and use of QTL for complex traits in multiple environments. *Current Opinion in Plant Biology*, *13*(2), 193–205. doi:10.1016/j.pbi.2010.01.001.

Elmer I., 2014. Caractérisation de ressources génétiques conférant une résistance partielle à la sclérotiniose chez le soja. Thèse de Doctorat (Ph.D.) en biologie végétale, Universite Laval, 201 p.

Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E. S., et al., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi: 10.1371/journal.pone.0019379.

Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., et al., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi: 10.1371/journal.pone.0019379.

Eswaran H., Reich P., and Beigroth F., 1997. Global distribution of soils with acidity. In Plant soil interactions at low pH. Ed. A C Monitz. pp. 159-164. *Brazilian soil science society*, campines Brazil.

Falush D., Stephens M., and Pritchard J. K., 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, *7*(4), 574-578.

FAO, 2013. How to feed the world in 2050. 2009. [Online] Available: http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf Accessed: 3 March 2013.

Feldman M., Levy A.A., Fahima T., and Korol A., 2012. Genomic asymmetry in allopolyploid plants: wheat as a model. J. Exp. Bot. 63, 5045–5059. doi: 10.1093/jxb/ers192.

Feldman, 2001. Origin of cultivated wheat. In: Alain P. Bonjean WJA (ed) The World wheat Book: A history of wheat breeding. Lavoister Publications, Paris.

Flint-Garcia S.A., Thornsberry J.M., and IV B., 2003. Structure of Linkage Disequilibrium in Plants*. *Annual Review of Plant Biology*, *54*(1), 357-374.

Froese P.S., and Carter H.A., 2016. Single Nucleotide Polymorphisms in the Wheat Genome Associated with Tolerance of Acidic Soils and Aluminum Toxicity. Crop Sci. 56:1662–1677 (2016). doi: 10.2135/cropsci2015.10.0629.

Fu Y.B., Peterson G.W., Richards K.W., Somers D., DePauw R.M., Clarke J.M., 2005. Allelic reduction and genetic shift in the Canadian hard red spring wheat germplasm released from 1845 to 2004. Theoretical and Applied Genetics 110, 1505–1516.

Gan Y., and Stobbe E.H., 1996. Seedling vigor and grain yield of 'Roblin' wheat affected by seed size. Agron. J. 88, 456–460. doi: 10.2134/agronj1996.00021962008800030016x.

Gegas V.C., Nazari A., Griffiths S., Simmonds J., Fish L., Orford S., et al., 2010. A genetic framework for grain size and shape variation in wheat. Plant Cell. 22, 1046–1056. doi: 10.1105/tpc.110.074153.

Goedert W.J., Lobato E., Lourenco S., 2016. Nutrient use efficiency in Brazilian acid soils: nutrient management and plant efficiency, in: A.C. Moniz, et al., (Eds.), Plant Soil Interaction. www.wheatgenome.org, 2016. IWGSC poster updated_V2.

Gore M., Bradbury P., Hogers R., Kirst M., Verstege E., van Oeveren J., et al., 2007. Evaluation of target preparation methods for single-feature polymorphism detection in large complex plant genomes. *Crop Sci.* 47, S135–S148. doi: 10.2135/cropsci2007.02.0085tpg.

Gorji A.H., and Zolnoori M., 2011. Genetic diversity in Hexaploid Wheat Genotypes using Microsatellite Markers. Asian Journal of Biotechnology, 2011. Knowledgia Review, Malaysia. ISSN 1996-0700/DOI, 10.3923/ajbkr.

Groos C., Robert N., Bervas E., and Charmet G., 2003. Genetic analysis of grain protein-content, grain yield and thou- sand-kernel weight in bread wheat. Theor. Appl. Genet. 106, 1032–1040. doi: 10.1007/s00122-002-1111-1.

Gupta N., Gaurav S.S., and Kumar A., 2013. Molecular basis of aluminium toxicity in plants: A review *American Journal of Plant Sciences*, 4, pp. 21-37.

Halloran G.M., Ogbonnaya F.C., and Lagudah E.S., 2008. Triticum (Aegilops) tauschii in the natural and artificial synthesis of hexaploid wheat. Aust. J. Agric. Res. 59, 475–490.

Harter R.D., 2007. Les sols acides des tropiques. *Echo Note Technique*. 12 p.

Hong-Xiang M., Gui-Hua B. & Wei-Zhong L., 2006. Quantitative trait loci for Aluminum resistance in wheat cultivar Chinese Spring. Plant and Soil (2006) 283:239–249, DOI 10.1007/s11104-006-0008-1.

Howie B.N., Donnelly P., Marchini J.,2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genetics; 5(6): e1000529 doi: 10.1371/ journal.pgen.1000529 PMID: 19543373.

Hua C., 2016. Quantitative trait loci and genome wide association mapping in western Canadian spring wheat (*Triticum aestivum* L.), a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Plant Science, *Department of Agricultural, Food and Nutritional Science University of Alberta*, 180p.

Huang X., Qian Q., Liu Z., Sun H., He S., Luo D., et al., 2009. Natural variation at the DEP1 locus enhances grain yield in rice.Nat. Genet. 41, 494–497. doi: 10.1038/ng.352.

Huang X.Q., Coster H., Ganal M.W., and Roder M.S., 2003. Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (Triticum aestivum L.). Theor. Appl. Genet. 106, 1379–1389. doi: 10.1007/s00122-002-1179-7.

Ilyas M., Mahmood T., Ali A., Babar M., Rasheed A., and Mujeeb-Kazi A., 2015. Characterization of D-genome diversity for tolerance to boron toxicity in synthetic hexaploid wheat and in silico analysis of candidate genes. Acta Physiol. Plant. 37, 1–13. doi: 10.1007/s11738-014-1765-8.

Imtiaz M., Ogbonnaya F.C., Oman J., and Van Ginkel M., 2008. Characterization of quantitative trait loci controlling genetic variation for preharvest sprouting in synthetic backcross-derived wheat lines. Genetics 178, 1725–1736. doi: 10.1534/genetics.107.084939.

IWGSC, 2014. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science*, Vol. 345, Issue 6194, DOI: 10.1126/science.1251788.

Jayasundara H.P.S., Thomson B.D. and Tang C., 1998. Responses of cool season grain legumes to soil abiotic stresses. Advances in Agronomy 63:77-151.

Jiangfeng H., Xiaoqing Z., Laroche A., Zhen-Xiang L., HongKui L. and Ziqin L.**,** *2014. Genotyping-bysequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front. Plant Sci.* ***5****:484. doi: 10.3389/fpls.2014.00484.*

Kochian L.V., 1995. Cellular Mechanisms of Aluminum Toxicity and Resistance in Plants. *Annual Reviews of Plant Physiology and Plant Molecular Biology,* 46**,** 237-260.

Kochian L.V., Pineros M.A., Liu J., and Magalhaes J.V., 2015. Plant Adaptation to Acid Soils: The Molecular Basis for Crop Aluminum Resistance. Annu. Rev. Plant Biol. 2015. 66:23.1–23.28. doi: 10.1146/annurev-arplant-043014-114822.

Kovermann P., Meyer S., Hoertensteiner S., Picco C., Scholz-Starke J., Ravera S., Lee Y., Martinoia E., 2007. The Arabidopsis vacuolar malate channel is a member of the ALMT family, Plant J. 52, 1169–1180.

Kraakman A.T., Niks R.E., Van den Berg P.M., Stam P., and Van Eeuwijk F.A., 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, *168*(1), 435-446.

Kraakman A.T.W., Martinez F., Mussiraliev B., Van Eeuwijk F.A., and Niks R.E., 2006. Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Molecular Breeding*, *17*(1), 41-58.

Krill A.M., Kirst M., Kochian L.V., Buckler E.S. and Hoekenga O.A., 2010. Association and Linkage Analysis of Aluminum Tolerance Genes in Maize. *Plos one*, 5.

Kronstad, W.E., 1997. Agricultural development and wheat breeding in the 20th Century. In: Wheat: Prospects for Global Improvement. Proceedings of the 5th International Wheat Conference, 10–14 June, 1996, Ankara, Turkey.

Krstic D., Ivica D., Dragoslav N. and Dragana B., 2012. Aluminium in Acid Soils: Chemistry, Toxicity and Impact on Maize Plants.*Food Production.*

Kumar S., Banks T.W., and Cloutier S., 2012. SNP Discovery through Next-Generation Sequencing and Its Applications. *International Journal of Plant Genomics, 2012*. doi:10.1155/2012/831460.

Lander E.S., 1996. The new genomics: global views of biology. *Science* 274, 536– 539. doi: 10.1126/science.274.5287.536.

Li H., Durbin R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25: 1754-1760. doi: 10.1093/bioinformatics/btp324. PubMed: 19451168.

Li H., Ye G., and Wang J., 2007. A Modified Algorithm for the Improvement of Composite Interval Mapping. *Genetics*, *175*, 361–374. doi:10.1534/genetics.106.066811.

Li Y., Zhou R., Wang J., Liao X., Branlard G., and Jia J., 2012. Novel and favorable qtl allele clusters for end-use quality revealed by introgression lines derived from synthetic wheat. Mol. Breed. 29, 627–643. doi: 10.1007/s11032-011-9578-6.

Lincoln S.E., and Lander S.L., 1993. Mapmaker/exp 3.0 and Mapmaker/QTL 1.1. *Whitehead Institute, Cambridge, MA*.

Lipka A.E., Tian F., Wang Q., Peiffer J., Li M., Bradbury P.J., Gore M., Buckler E. and Zhang Z., 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397-2399.

Lozet J. et Mathieu C., 1986. Dictionnaire de science du sol. Technique et Documentation LAVOISIER.

Lu F., Lipka A.E., Glaubitz J., Elshire R., Cherney J.H., Casler M.D., et al., 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genetics* 9:e1003215. doi: 10.1371/journal.pgen.1003215.

Ma H.X., Bai G.H., Carver B.F., Zhou L.L., 2005. Molecular mapping of a quantitative trait locus for aluminum tolerance in wheat cultivar Atlas 66, Theor. Appl. Genet. 112, 51–57.

Ma L., Li T., Hao C., Wang Y., Chen X., and Zhang X., 2016. TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. Plant Biotechnol. J. 14: 1269–1280. doi: 10.1111/pbi.12492.

Mackay I., and Powell W., 2007. Methods for linkage disequilibrium mapping in crops. *Trends in Plant Science*, *12*(2), 57-63.

Mammadov J., Aggarwal R., Buyyarapu R., and Kumpatla S., 2012. SNP Markers and Their Impact on Plant Breeding. *International Journal of Plant Genomics, 2012*. doi:10.1155/2012/728398.

Manly K.F., Cudmore Jr R.H., and Meer J.M., 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome*, *12*(12), 930-932.

Mantovani P., Maccaferri M., Sanguineti M.C., Tuberosa R., Catizone I., Wenzl P., Thomson B., Carling J., Huttner E., DeAmbrogio E., Kilian A., 2008. An integrated DArT-SSR linkage map of durum wheat, Mol. Breed. 22, 629–648.

Marathi B., Guleria S., Mohapatra T., Parsad R., Mariappan N., Kurungara V.K., Atwal S.S., Prabhu K.V., Singh N.K. and Singh A.K., 2012. QTL analysis of novel genomic regions associated with yield and yield related traits in new plant type based recombinant inbred lines of rice (Oryza sativa L.). BMC Plant Biol. 12: 137.

Marshall D.R., Ellison F.W., and Mares D.J., 1984. Effects of grain shape and size on milling yields in wheat. I. Theoretical analysis based on simple geometric models. Aust. J. Agric. Res. 35, 619–630. doi: 10.1071/AR9840619.

Matsuoka Y., 2011. Evolution of polyploidy triticum wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol. 52, 750–764.

Meyer R.S., Purugganan M.D., 2013. Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14:840–52. Nat Rev Genet. 2013 Dec;14(12):840-52. doi: 10.1038/nrg3605.

Miftahudin G., Scoles J., Gustafson J.P., 2002. AFLP markers tightly linked to the aluminum-tolerance gene Alt3 in rye (Secale cereale L.), Theor. Appl. Genet. 104 (2002) 626–631.

Mir R.R., Kumar N., Jaiswal V., Girdharwal N., Prasad M., Balyan H.S., et al., 2012. Genetic dissection of grain weight in bread wheat through quantitative trait locus interval and association mapping. Mol. Breed. 29, 963–972. doi: 10.1007/s11032-011-9693-4.

Monthe E., Birie habas J., 1980. Premiers résultats sur la mise au point de la culture du ble sur les plateaux de l'adamaoua au cameroun. O.R.S.T.O.M. Fonds documentaire No 17751. 40p.

Mossor-pietraszewska T., 2001. Effect of aluminium on plant growth and metabolism. *Acta Biochimica Polonica,* 48**,** 673-686.

Narum S.R., Buerkle C.A., Davey J.W., Miller M.R., and Hohenlohe P.A., 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* 22, 2841–2847. doi: 10.1111/mec.12350.

Navakode S., Neumann K., Kobiljski B., Lohwasser U., Börner A., 2014. Genome wide association mapping to identify aluminium tolerance loci in bread wheat. Euphytica, 198: 401. doi:10.1007/s10681-014-1114-4.

Navakode S., Weidner A., Varshney R.K., Lohwasser U., Scholz U., Roeder M.S., Boerner A., 2010. A genetic analysis of aluminium tolerance in cereals, Agric. Conspec. Sci. 75, 191–196.

Nelson J.C., 1997. QGENE: software for marker-based genomic analysis and breeding. *Molecular breeding*, *3*(3), 239-245.

Ngo Ngom K.A., 2017. Evaluation du potentiel genetique des accessions de ble (Triticum aestivum L.) en zone forestiere humide a pluviometrie bimodale du Cameroun.Master. UYI.

Ngonkeu M. E. L., 2009. Tolérance de certaines variétés de maïs aux sols à toxicité aluminique et manganique du Cameroun et diversités moléculaire et fonctionnelle des mycorhizes à arbuscules. Thèse de Doctorat Ph/ D. université de Yaoundé I. 255 p.

Nielsen R., Paul J.S., Albrechtsen A., & Song Y. S., 2011. Genotype and SNP calling from next-generation sequencing data. Nature Rev. Genet. 12, 443–45. doi: 10.1038/nrg2986.

Njomgang R., Yemefack M., Nyobe T., Tchienkoua M., Ambassa-Kiki R., Agoume V. and Ndaka B.S., 2010. Profil du pays sur les besoins en éléments nutritifs des plantes cultivées dans les principales zones agro-écologiques du cameroun: Description sommaire des zones agro-ecologiques (zae) du cameroun, *Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) and Institut de Recherche Agricole pour le Développement (Cameroun)*, p. 6.

Okamoto Y., Kajimura T., Ikeda T.M., and Takumi S., 2012. Evidence from principal component analysis for improvement of grain shape- and spikelet morphology-related traits after hexaploid wheat speciation. Genes Genet. Syst. 87, 299–310. doi: 10.1266/ggs.87.299.

Okamoto Y., Nguyen A.T., Yoshioka M., Iehisa J.C., and Takumi S., 2013. Identification of quantitative trait loci controlling grain size and shape in the D genome of synthetic hexaploid wheat lines. Breed. Sci. 63, 423–429. doi: 10.1270/jsbbs.63.423.

Ota, 1968. Studies on the occurrence of the physiological disease of rice called bronzing." *bull. Nat Inst. Agr. Sci.* Nishigahara, Tokyo, japon ser . D (Plant hysiology, genetics and crops in genera), (18) 108 p.

Pasam R.K., Sharma R., Malosetti M., van Eeuwijk F.A., Haseneyer G., Kilian B., and Graner A., 2012. Genome-wide association studies for agronomical traits in a world wide spring barley collection. *BMC Plant Biology*, *12*(1), 16.

Paux E., Sourdille P., Mackay I., and Feuillet C., 2012. Sequence-based marker development in wheat : Advances and applications to breeding. *Biotechnology Advances*, *30*(5), 1071–1088. doi:10.1016/j.biotechadv.2011.09.015.

Poczai P., Varga I., Laos M., Cseh A., Bell N., Valkonen J.P.T., Hyvonen J., 2013. Advances in plant gene-targeted and functional markers: a review, Plant Methods 9, 6.

Poland J., Endelman J., Dawson J., Rutkoski J., Wu S. et al., 2012b. Genomic selection in wheat breeding using genotyping-by-sequencing. The Plant Genome 5: 103-113. doi:10.3835/plantgenome2012.06.0006.

Poland J.A., and Rife T.W., 2012. Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5, 92–102. doi: 10.3835/plantgenome2012.05.0005.

Poland J.A., Brown P.J., Sorrells M.E., and Jannink J.L., 2012a. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7:e32253. doi: 10.1371/journal.pone.0032253.

Qi P., Lin Y.S., Song X.J., Shen J.B., Huang W., Shan J.X., Zhu M.Z., Jiang L., Gao J.P. and Lin H.X., 2012. The novel quantitative trait locus GL3.1 controls rice grain size and yield by regulating Cyclin- T1;3. Cell Res. 22: 1666–1680.

Rajaram S., and Hettel G.P., 1994. Wheat Breeding at CIMMYT: Commemorating 50 Years of Research in Mexico for Global Wheat Improvement. Wheat special report N° 29. Ciudad Obregon, Sonora, Mexico 21-25 March, 1994.

Raman H., Raman R., Wood R., Martin P., 2006. Repetitive inDel markers within the ALMT1 gene conditioning aluminium tolerance in wheat (Triticum aestivum L.), Mol. Breed. 18, 171–183.

Raman H., Ryan P.R., Raman R., Stodart B.J., Zhang K., Martin P., Wood R., Sasaki T., Yamamoto Y., Mackay M., Hebb D.M., Delhaize E., 2008. Analysis of TaALMT1 traces the transmission of aluminum resistance in cultivated common wheat (Triticum aestivum L.), Theor. Appl. Genet. 116, 343–354.

Raman H., Stodart B., Emebiri L., Raman R., Coombes N., and Milgate A., 2010. Genome-wide association analyses of common wheat (*Triticum aestivum* L.) germplasm identifies multiple loci for aluminium resistance. Genome **53**: 957–966, doi:10.1139/G10-058.

Rana R.M., Bilal M., Rehman S.U., Iqbal F., and Shah M.K.N., 2013. Synthetic wheat; a new hope for the hungry world. Asian J. Agric. Biol. 1, 91–94..

Rasheed A., Xia X.C., Ogbonnaya F., Mahmood T., Zhang Z., Mujeeb-Kazi A, et al., 2014. Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. BMC Plant Biol. 14:128. doi: 10.1186/1471-2229-14-128.

Reeves T.G. and Cassaday K., 2002. History and past achievements of plant breeding, *Australian Journal of Agricultural Research*, 53, pp. 851-863.

Reynolds M., Dreccer F., and Trethowan R., 2007. Drought-adaptive traits derived from wheat wild relatives and landraces. J. Exp. Bot. 58, 177–186. doi: 10.1093/jxb/erl250.

Reynolds M.P., Ortiz Monasterio J.I., McNab A., 2001. In: Reynolds MP, Ortiz Monasterio JI, McNab A (eds) Application of Physiology in Wheat Breeding, CIMMYT, Mexico, D.F.

Reza M.D., Goodarz N.E. and Aram S., 2015. Investigation of genetic diversity of some durum and bread wheat genotypes using SSR markers. *Journal of Biodiversity and Environmental Sciences*. 2222-3045 (Online) Vol. 6, No. 3, p. 24-32.

Riede C.R., Anderson J.A., 1996. Linkage of RFLP markers to an aluminum tolerance gene in wheat, Crop. Sci. 36, 905–909.

Rife T., Wu S., Bowden R. and Poland J., 2015. Spiked GBS: a unified, open platform for single marker genotyping and whole-genome profiling. BMC Genom. 16, 248–000.

Robain H., 1993. Caractérisation et classification des sols de la zone forestière humide .In. R.

Röder M.S., Huang X.Q., and Börner A., 2008. Fine mapping of the region on wheat chromosome 7D controlling grain weight. Funct. Integr. Genomics 8, 79–86. doi: 10.1007/s10142-007-0053-8.

Rosenburg NA., 2004. DISTRUCT: a program for the graphical display of population structure. Mol Ecol;4:137–8.

Rout G.R., Samantary S. and Das P., 2000. Aluminum toxicity in plants. a *review agronomic*,21:3-21.

Ryan P.R., Raman H., Gupta S., Horst W.J., Delhaize E., 2009. A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots, Plant Physiol. 149, 340–351.

Sakamoto T., Matsuoka M., 2008. Identifying and exploiting grain yield genes in rice. *Curr. Opin. Plant Biol.* 11:209–14. Volume 11, Issue 2, April 2008, Pages 209-214. doi.org/10.1016/j.pbi.2008.01.009.

Sansaloni C.P., Petroli C.D., Carling J., Hudson C.J., Steane D.A., Myburg A.A., Grattapaglia D., Vaillancourt R.E., Kilian A., 2010. A high-density diversity arrays technology (DArT) microarray for genome-wide genotyping in Eucalyptus, Plant Methods 6 16.

Sasaki T., Yamamoto Y., Ezaki B., Katsuhara M., Ahn S.J., Ryan P.R., Delhaize E., Matsumoto H., 2004. A wheat gene encoding an aluminum-activated malate transporter, Plant J. 37, 645–653.

Scheet P., Stephens M., 2006. A fast and flexible statistical model for large scale population genotype data: application to inferring missing genotypes and haplotypic phase. Am. J. Hum. Genet.; 78, pp. 629–44 PMID: 16532393.

Segura V., Vilhjálmsson B.J, Platt A., Korte A., Seren U., Long Q., et al., 2012. An efficient multi-locus mixedmodel approach for genome-wide association studies in structured populations. Nat. Genet.; doi:10.1038/ng.2314.

Semagn K., Bjornstad A., and Xu Y., 2010. The genetic dissection of quantitative traits in crops. *Electronic journal of Biotechnology*, *13,* 1–45. doi:10.2225/vol13-issue5-fulltext-21.

Semagn K., Bjornstad A., Ndjiondjop M.N., 2006. Principles, requirements and prospects of genetic mapping in plants, Afr. J. Biotechnol. 5, 2569–2587.

Semagn K., Bjornstad A., Skinnes H., Maroy A.G., Tarkegne Y., William M., 2006. Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population, Genome 49, 545–555.

Shendure J., and Ji H., 2008. Next-generation DNA sequencing. *Nat. Biotechnol.* 26, 1135–1145. doi: 10.1038/nbt1486.

Sibiya J., Tongoona P., Derera J. and Makanda I., 2013. Farmers' desired traits and selection criteria for maize varieties and their implications for maize breeding: A case study from kwazulu-natal province, south Africa, *Journal of Agriculture and Rural Development in the Tropics and Subtropics*, 114(1), pp. 39-49.

Simmonds J., Scott P., Brinton J., Mestre T.C., Bush M., del Blanco, A., et al., 2016. A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. Theor. Appl. Genet. 129, 1099–1112. doi: 10.1007/s00122-016-2686-2.

Simons K.J., Fellers J.P., Trick H.N., Zhang Z.C., Tai Y.S., Gill B.S., et al., 2006. Molecular characterization of the major wheat domestication gene Q. Genetics 172, 547–555. doi: 10.1534/genetics.105.044727.

Slatkin M., 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, *9*(6), 477-485.

Sobota RS, Shriner D, Kodaman N, Goodloe R, Zheng W, Gao YT, Edwards TL, Amos CI, Williams SM., 2015. Addressing population-specific multiple testing burdens in genetic association studies. *Ann Hum Genet*. 2015;79:136–47.

Somers D.J., Gustafson J.P., 1995. The expression of aluminum stress induced polypeptides in a population segregating for aluminum tolerance in wheat (*Triticum aestivum* L), Genome 38,1213–1220.

Sonah H., Bastien M., Iquira E., Tardivel A., Legare G., Boyle B., et al., 2013. An Improved Genotyping by Sequencing (GBS) Approach Offering Increased Versatility and Efficiency of SNP Discovery and Genotyping. *PLoS ONE*; 8(1): e54603. doi: 10.1371/journal.pone.0054603 PMID: 23372741.

Sonah H., Bastien M., Iquira E., Tardivel A., Legare G., Boyle B., et al., 2013. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE* 8:e54603. doi: 10.1371/journal.pone.0054603.

Sonah H., O'Donoughue L., Cober E., Rajcan I., Belzile F., 2014. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnol. J.*; doi: 10.1111/pbi.12249.

Song X.-J., Huang W., Shi M., Zhu M.-Z., and Lin H.-X., 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat. Genet. 39, 623–630. doi: 10.1038/ng2014.

Soto-Cerda, B.J., and Cloutier S., 2012. Association mapping in plant genomes. *Genetic diversity in plants. InTech, Rijeka*, 29-54.

Stodart B.J., Raman H., Coombes N., Mackay M., 2007. Evaluating landraces of bread wheat Triticum aestivum L. for tolerance to aluminium under low pH conditions, Genet. Resour. Crop Evol. 54, 759–766.

Sun X., Marza F., Ma H., Carver B.F. and Bai G., 2010. Mapping quantitative trait loci for quality factors in an inter-class cross of US and Chinese wheat. Theor. Appl. Genet. 120: 1041–1051.

Sun X., Wu K., Zhao Y., Kong F., Han G., Jiang H., et al., 2009. QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165, 615–624. doi: 10.1007/s10681-008-9794-2.

Tabuchi H., and Matsumoto H., 2001. Changes in cell wall properties on wheat (*Triticum aestivum*) roots during aluminium induced growth inhibition. *Physiol. Plant*, 112, p. 353-358.

Tadesse W., Schmolke M., Hsam S.L., Mohler V., Wenzel G., and Zeller F.J., 2007. Molecular mapping of resistance genes to tan spot [Pyrenophora triticirepentis race 1] in synthetic wheat lines. Theor. Appl. Genet. 114, 855–862. doi: 10.1007/s00122-006-0484-y.

Takumi S., Naka Y., Morihiro H., and Matsuoka Y., 2009. Expression of morphological and flowering time variation through allopolyploidization: an empirical study with 27 wheat synthetics and their parental Aegilops tauschii accessions. Plant Breed. 128, 585–590. doi: 10.1111/j.1439-0523.2009. 01630.x.

Tang Y., Garvin D.F., Kochian L.V., Sorrells M.E., Carver B.F., 2002. Physiological genetics of aluminum tolerance in the wheat cultivar Atlas 66, Crop. Sci. 42, 1541–1546.

Tanksley S.D. and McCouch S.R., 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Science, 277, 1063–1066.

Tardivel A., Sonah H., Belzile F., and O'Donoughue L.S., 2014. Rapid Identification of Alleles at the Soybean Maturity Gene E3 using genotyping by Sequencing and a Haplotype-Based Approach. The Plant genome, vol.7, doi: 10.3835/plantgenome2013.10.0034.

Tékeu H., Ngonkeu E.M.L., Djocgoué P.F., Ellis A., Lendzemo V., Springfield L., Moulin L., Klonowska A., Diouf D., Botes W.C. and Béna G., 2017. Genetic diversity of Cameroonian bread wheat (*Triticum aestivum* L.) cultivars revealed by microsatellite markers. *African Journal of Biotechnology*, Vol. 16(36), pp. 1832-1839, 6 September 2017. DOI: 10.5897/AJB2017.16090.

Tekeu H., Ngonkeu E.M.L., Ngoune T.L., Djocgoue P.F., Bell J.M., Mafouasson H., Boyomo O., Petmi C.L., and Fokom R, 2015. Evaluation of maize (Zea mays L.) accessions using Line x Tester Analysis for aluminum and manganese tolerance. International Journal of Biological and Chemical Sciences. Vol 9(4): 2161-2173, August 2015. DOI:http://dx.doi.org/10.4314/ijbcs.v9i4.36.

Thornton F.C., Schaedle M. and Raynal D.L., 1986. Effect of aluminum on the growth of sugar maple in solution culture. *Can. J. For. Res.,* 16**,** 892-896.

Thudi M., Li Y., Jackson S.A., May G.D., and Varshney R.K., 2012. Current state-of-art of sequencing technologies for plant genomics research. *Brief. Funct. Genomics* 11, 3–11. doi: 10.1093/bfgp/elr045.

Torkamaneh D., Belzile F., 2015. Scanning and Filling: Ultra-Dense SNP Genotyping Combining Genotyping-By-Sequencing, SNP Array and WholeGenome Resequencing Data. PLoS ONE 10(7): e0131533. doi: 10.1371/journal.pone.0131533.

Torkamaneh D., Laroche J., Belzile F., 2016. Genome-Wide SNP Calling from Genotyping.

Trethowan R.M., and Mujeeb-Kazi A., 2008. Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. Crop Sci. 48, 1255–1265. doi: 10.2135/cropsci2007.08.0477.

Tsilo T.J., Hareland G.A., Simsek S., Chao S., and Anderson J.M., 2010. Genome mapping of kernel characteristics in hard red spring wheat breeding lines. Theor. Appl. Genet. 121: 717–730.

Utz H.F., and Melchinger A.E., 1996. PLABQTL: a program for composite interval mapping of QTL. *J Quant Trait Loci*, *2*(1), 1-5.

Valluru R., Reynolds P.M., and Salse J., 2014. Genetic and molecular bases of yield-associated traits: a translational biology approach between rice and wheat. Theor. Appl. Genet. 127, 1463–1489. doi: 10.1007/s00122-014-2332-9.

Van Inghelandt D., Reif J.C., Dhillon B.S., Flament P., and Melchinger A.E., 2011. Extent and genome-wide distribution of linkage disequilibrium in commercial maize germplasm. *Theoretical and applied genetics*, *123*(1), 11-20.

Velasquez P.C.J., Lopes de souza C., Narro A.L., Pandey S., 2008. Genetic effects for maize traits in acid and non-acid soils. *Genetics and Molecular Biology,* 31**,** 89-97.

Von Baer E, 2007. wheat breeding for soil acidityand aluminum toxicity. *H.T. Buck et al. (eds.), Wheat Production in Stressed Environments*, 387–393.

Von Uexkull H.R., and Mutert E., 1995. Global extent, development and economic impact of acid soils. Plant and Soil 171:1-15.

Voorrips RE, 2002. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. Journal of Heredity 93 (1):77-78. doi:10.1093/jhered/93.1.77.

Wang E.,Wang J., Zhu X., Hao W., Wang L., Li Q., et al., 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. Nat. Genet. 40, 1370–1374. doi: 10.1038/ng.220.

Wang G.Z., Miyashita N.T., Tsunewaki K., 1997. Plasmon analyses of *Triticum* (wheat) and Aegilops: PCR single-strand conformational polymorphism (PCR-SSCP) analyses of organellar DNAs. Proceedings of the National Academy of Sciences of the United States of America 94 (26):14570-14577. doi:10.1073/pnas.94.26.14570.

Wang J., 2009. Inclusive composite interval mapping of quantitative trait genes. *Acta Agronomica Sinica*, *35*(2), 239-245.

Wang J., 2013. Inclusive Composite Interval Mapping (ICIM) of Quantitative Trait Genes. In *Plant and Animal Genome XXI Conference*. Plant and Animal Genome.

Wang M., Yan J., Zhao J., Song W., Zhang X., Xiao Y., et al., 2012. Genome-wide association study (GWAS) of resistance to head smut in maize. Plant Sci. 196, 125–31. doi: 10.1016/j.plantsci.2012.08. 004 PMID: 23017907.

Wang S., Wong D., Forrest K., Allen A., Chao S., Huang B.E., Maccaferri M., et al., 2014. Characterization of polyploid wheat genome diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnol. J. 12, 787–796.

Wang S., Wong D., Forrest K., Allen A., et al., 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnology Journal, 12, pp. 787–796 doi: 10.1111/pbi.12183.

Waugh R., Jannink J.L., Muehlbauer G.J., and Ramsay L., 2009. The emergence of whole genome association scans in barley. *Current opinion in plant biology*, *12*(2), 218-222.

Weng J., Gu S., Wan X., Gao H., Guo T., Su N., et al., 2008. Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. Cell Res. 18, 1199–1209. doi: 10.1038/cr.2008.307.

Weston L.A., Ryan P.R., Watt M., 2012. Mechanisms for cellular transport and release of allelochemicals from plant roots into the rhizosphere, J. Exp. Bot. 63, 3445–3454.

Wheat Initiative, 2011.http://www.wheatinitiative.org/sites/default/files/ wysiwyg/g20-proposal.pdf.

Williams K., and Sorrells M.E., 2014. Three-dimensional seed size and shape QTL in hexaploid wheat (*Triticum aestivum* L.) populations. Crop Sci. 54, 98–110. doi: 10.2135/cropsci2012.10.0609.

Williams K., Munkvold J.and Sorrells M., 2013. Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (Triticum aestivum L.). Euphytica 190: 99–116.

Winfield M.O., Allen A.M., Burridge A.J., Barker G.L., Benbow H.R., Wilkinson P.A., Coghill J., et al., 2015. High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. Plant Biotechnol. J. 000, 000–000.

Winfield M.O., Wilkinson P.A., Allen A.M., Barker G.L.A., Coghill J.A., Burridge A., Hall A., et al., 2012. Targeted re-sequencing of the allohexaploid wheat exome. Plant Biotechnol. J. 10, 733–742.

Würschum T., 2012. Mapping QTL for agronomic traits in breeding populations. *Theoretical and Applied Genetics*, *125*(2), 201-210.

Xian-Jun S., 2017. Crop Seed Size: BR Matters. Molecular plant.; Volume 10, Issue 5, p668–669, 1 May 2017. DOI: http://dx.doi.org/10.1016/j.molp.2017.04.007.

Yamdjeu A., 2012. Données statistiques sur les importations, exportations et production des produits de base. http://www.acdi c.net.

Yan J., Warburton M., and Crouch J., 2011. Association mapping for enhancing maize (L.) genetic improvement. *Crop Science*, *51*(2), 433-449.

Yan L., Liang F., Xu H., Zhang X., Zhai H., Sun Q., and Ni Z., 2017. Identification of QTL for Grain Size and Shape on the D Genome of Natural and Synthetic Allohexaploid Wheats with Near-Identical AABB Genomes. Front. Plant Sci. 8:1705. doi: 10.3389/fpls.2017.01705.

Yang Z, Li Z, Bickel D.R., 2013. Empirical Bayes estimation of posterior probabilities of enrichment: a comparative study of five estimators of the local false discovery rate. BMC Bioinformatics, 14, 87. doi:10.1186/1471-2105-14-87 PMID: 23497228.

Yu J., and Buckler E.S., 2006. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, *17*(2), 155-160.

Yu M., Chen G., Zhang L., Liu Y., Liu D., Wang J., et al., 2014. QTL mapping for important agronomic traits in synthetic hexaploid wheat derived from Aegiliops tauschii ssp. tauschii. J. Integr. Agric. 13, 1835–1844.doi: 10.1016/S2095-3119(13)60655-3.

Zhang L.-Y., Liu D.-C., Guo X.-L., Yang W.-L., Sun J.-Z, Wang D.-W., et al., 2010. Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. J. Integr. Plant Biol. 52, 996–1007. doi: 10.1111/j.1744-7909.2010.00967.x

Zhang X., Wang J., Huang J., Lan H., Wang C., Yin C., Wu Y., Tang H., Qian Q., Lin J.et al., 2012. Rare allele of OsPPKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. Proc. Natl. Acad. Sci. USA 109: 21534-21539.

Zhao H.H., Fernando R.L., and Dekkers J.C., 2007. Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. *Genetics*, *175*(4), 1975-1986.

Zhao J.L., Wang H.W., Zhang X.C., Du X.Y., Li A.F., and Kong L.R., 2015. Association analysis of grain traits with SSR markers between Aegilops tauschii and hexaploid wheat (Triticum aestivum L.). J. Integr. Agric. 14, 1936–1948.doi: 10.1016/S2095-3119(15)61070-X.

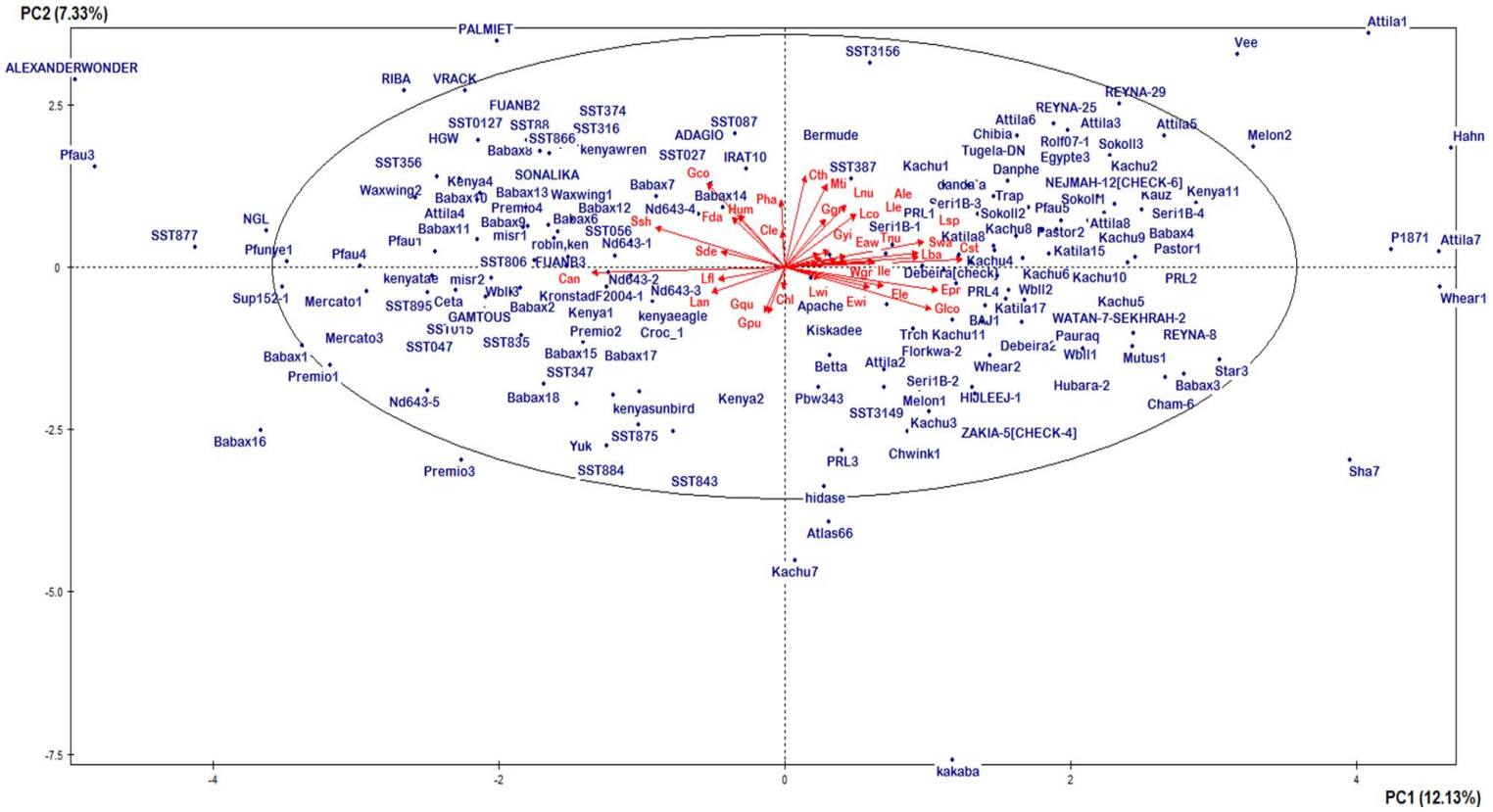
Zhao K., Aranzana M.J., Kim S., Lister C., Shindo C., Tang C., and Nordborg M., 2007. An Arabidopsis example of association mapping in structured samples. *PLoS Genetics, 3*(1), e4.

Zhou L.L., Bai G.H., Ma H.X., Carver B.F., 2007. Quantitative trait loci for aluminum resistance in wheat, Mol. Breed. 19, 153–161.

Zhu C., Gore M., Buckler E.S., and Yu J., 2008. Status and prospects of association mapping in plants. *The plant genome, 1*(1), 5-20.

Zhu Y.L., Song Q.J., Hyten D.L., van Tassell C.P., Matukumalli L.K., Grimm D.R., et al., 2003. Single-nucleotide polymorphisms in soybean. *Genetics* 163, 1123–1134. doi: 10.1126/science.2047873.

APPENDIX



Appendix 1. GT Biplot of 179 wheat breeding lines with 39 wheat descriptors

Appendix 2

Questionnaire Sheet

Socio-economic aspects of agro-ecological zones and limiting factors of wheat and productivity improvement strategies

No plug.....

I. LOCATION OF STUDY AREA

Village. ..... ……..
District.....
Department.....
Region.....
Geographic coordinates (GPS).....

II. IDENTIFICATION FARMER

1. Name and Surname: ....
Sex..... ..... Age.....

2. Level of education:

Primary education (0 to 6 years)
Secondary education (7 to 13 years)
University studies (14 years and older).

3. Professional Training.

Agricole. Other (specify).....

4. Household status

Head of household. Window Bride. Child. Other (specify).....

5. Origin of labor

Family. Association. Official. Temporary. Other (please specify).....

III. CULTURAL PRACTICES

1. What major crops do you practice? (List three main plants)

.....

2. What plant do you practice before that of wheat on the same plot (previous crop **or** intercropping)?

2.1. Maize
2.2. Potato
2.3. Bean
2.4. Sorghum
2.5. Fallow
2.6. Other (please specify).....
2.7. No culture

3. What plant do you collect at the same time as the wheat?

……..

4. What are your different dishes obtained after processing of wheat

4.1. Donuts

4.2. Porridge
4.3. Bread
4.4. Cakes
4.5. Other (please specify).....

5. Wheat growing periods in the village (month of year)
Months of sowing: .....
Harvest Month: .....
Pile driving period: .....
Cycle time ……..

6. Methods of managing soil fertility
Do you use fertilizer in your wheat fields?
1 = yes
2 = no

If so, what types of fertilizer do you use?
1 = organic fertilizer (cow dung, droppings of chickens, compost)
2 = chemical fertilizers (NPK, Urea, sulfated fertilizer)
3 = other (specify) .....

Are you satisfied with the yields obtained using this fertilizer?
1 = yes
2 = no

IV. ORIGINS OF GROWN SEEDS AND SELECTION METHODS

Where are the seeds you grow?	Replies (1 = yes, 2 = no)
Only seeds from my own field (previous harvest)	
Mixture of seeds from my own field with those from other farmers	
Seed from other farmers harvest only	
Seeds bought at the market only	
Mixture of seeds bought at the market with my seeds	
Seed from research institutes	
Seeds from selecting the best panicles for next season	
"The bulk" (Seeds for consumption)	

V. MOVEMENT OF SEEDS: MOMENTS, ACTORS AND NATURE OF EXCHANGE EQUIPMENT

Moments of exchange	Nature of the material exchanged (code 1)	Receiving? (Code 2)	Use (Code 3)
harvest			
threshing			
visits			
Ceremonies (parties, weddings, birth of a baby, mourning ...)			
Markets (purchase, exchange, ..)			

Code1: 1 = seeds; 2 = panicles
Code 2: 1 = family member; 2 = neighbors; 3 = in-laws; 4 = foreign to the village; 5 = friend; 6 = other knowledge (specify).
Code 3: 1 = food; 2 = seed

Names of the variety	Type of variety (code1)	How many years has this variety was introduced in the village? (if more than 30 to 30 years)	Who introduced this variety in the village? (code2)	Are the seeds of this variety available in the village? 1 = yes 0 = No	Are the seeds of this variety available somewhere out of this village? 1 = yes 0 = No	This variety is it still grown in the village? 1 = yes 0 = No

VI. IDENTIFICATION OF WHEAT VARIETY INTRODUCED IN THE VILLAGE

What are the varieties of wheat that have been introduced in this village

(Note: The full list of varieties should be stopped before asking other questions).

Code1- variety type: 1 = local variety; 2 = improved variety

Code2- Who introduced the variety: 1 = village peasant, 2 = farmer from another village, 3 = NCRE Project, 4 = National Extension Service, 5 = SODEBLE, 6 = NGO 6 = Peasant Organization, 7 = private structure, 8 = other (specify)

VII. COMMUNITY ASSESSMENT OF VARIETIES GROWN IN THE VILLAGE

NB: List first the most important varieties, up to 7 prior to evaluation

Variety code	1	2	3	4	5	6	7
Name of the variety							
Evaluation Code: 0: does not know anything; 1 = Low / short / bad; 2 = average; 3 = High (strong / good / good)							
Agronomic and morphological characteristics							
Production (grain yield)							
Undamaged by birds							
Weed resistance							
Lodging resistance							
Size of the panicle							
length of peduncle							
Tasks on the leaves (disease)							
Resistance to stalk rot							
Resistance to insect attack							
Resistance to attack rodents							
Resistance to attack by birds							
Post-harvest characteristics							
Facilitation ginning with Moulinex							
Threshing facilitation							
Shelling facilitation (manual shelling)							
Grain color (wheat)							
Culinary and sensory characteristics							
taste							
Arom (parfum)							
Ability to swell							

Classification of varieties (from best to worst) depending on the perspective of wheat growers							

Note: Use code 0 for the case (s) farmer (s) knows nothing about the feature set.

VIII. COMMUNITY ASSESSMENT OF VARIETIES GROWN

(The most popular features in wheat varieties)

Do you (farmer), what are the most important characteristics of a variety that you want to grow and could meet your expectations? (use codes and list in descending order of importance)

Code: 1 = high performance; 2 = tolerant of poor soils; 3 = Short dormancy; 4 = Early maturity / short cycle; 5 = tolerant diseases; 6 = many tillers; 7 = resistance to weeds; 8 = Easy to beat (threshing); 9 = Easy to shell or peel (hand shelling); 11 = Easy to maintain; 12 = Good market selling prices; 9 = Large grains; 10 = tapered Grains; 11 = short grain; 12 = long grain; 13 = white color; 14 = Other color (yellow, red, etc.); 15 = Good for flour swelling; 16 = Other (specify)

A) For men

1) /__/__/ **2)** /__/__/ **3)** /__/__/

4) /__/__/ 5) /__/__/ 6) /__/__/

B) For women

1) /__/__/ **2)** /__/__/ **3)** /__/__/

4) /__/__/ 5) /__/__/ 6) /__/__/

IX. VARIETY ABANDONED OR LOST SINCE THE YEAR 2000

Names of known varieties, but which are no longer cultivated	1 = Abandoned 2 = lost	Si abandonnée, précisez l'année et raison d'abandon	if lost, would you have again 1 = yes, 0 = no	If yes / no give reasons
1)-				
2)-				
3)-				
4)-				
5)-				
6)-				
7)-				

Code: 1 = high (strong) 2 = average; 3 = low; 0 = does not exist (Never Experienced)

X. KNOWLEDGE AND KEY MANAGEMENT AND BIOTIC STRESS abiotic

Name the main biotic and abiotic stresses in the village	Have you had this problem in your fields at least once 1 = yes, 2 = no	Impact of stress in the field (code)	Affected performance ratio when the problem occurred (%)	What methods did you use to manage this constraint
	abiotic constraints			
Decline in soil fertility (acid soil, poor in nutrients)				
Climate change (erratic rainfall, heat, cold)				
	Biotic constraints			
Lack of improved seeds				
Pest pressure (birds, nematodes, insects,)				
weeds				
Inadequate labor				
Diseases (rust, Fusarium, Septoria leaf blight)				
	Socio-economic constraints			
Weakness in wheat prices on the market				
Local wheat more expensive than imported wheat				
Unavailability of extension services				
Effectiveness of these services is insufficient				
Marginalization of growing wheat				

Appendix 3

Sample collection kit and protocol

Contents of the kit

1 punch

1 cutting mat

1 empty 2ml tube (for cleaning the punch)

Warning: Put 1 ml of water in the tube before going in the field to collect samples.

1 pair of tweezers

1 large Ziploc bag

1 bottle containing 200g of desiccant

2sheets of porous tape

2 plates, each containing twelve strips of eight capped tubes taped to a blue block

Warning: Do not remove the strips from the block, as the orientation of each strip is crucial to the correct identification of the samples (see examples in Fig.1).

Protocol

In the field:

Warning: Keep the plates containing the leaf samples on ice during the collection process.

1- Remove the caps from the first strip of tubes (tubes 1A to 1H).
2- Place a leaf from the first plant on the brown side of the cutting pad (Fig.2).
Warning. You don't need to remove the leaf from the plant.
3- Press the punchon the leaf and twist the punch back and forthto collect aleaf disk. Repeat seven times to gather **a total of 8 leaf disks** inside the punch (Fig.2).
Warning: Collect samples from the leaf blade and not from the main vein.
4- Insertthehead of the punch about 1 cm inside the tube located at position 1A on the plateand press the ejector (top of the punch) to simultaneously expel the8leaf disks in the tube (see Fig.1 for tube location).
Warning. You will need to press the small release button on the side of the punch to allow the ejector to go back inside the punch for the next step.
5- Write down in your field book the identification of the tube corresponding to this plant.
6- Repeat steps 2 to 5 each time with a different plant to fill the other seven tubes from the first strip (tubes 1B to 1H).
7- Replace the caps back onto the first strip of tubes.
8- Remove the caps from the nextstrip of tubes (tubes 2A to 2H).
9- Repeat steps 2 to 8until you have collected all thesamplesfrom the population.
Warning. You will need to clean the blade of the punch between each plate or sooner if you see any leaf debris sticking to it. To clean the blade, insert the head of the punch in the 2 ml tube containing water to put the blade into the water and press the ejector up and down a few times (Fig.3). Remove the punch from the

tube and strongly shake it 3 times to eliminate all the water. Make sure that there is no water left in the punch before using it again.

Back at your workplace:

Warning: If you don't proceed immediately with step 10, the plates with the leaf samples can be kept in the refrigerator (at 4°C) for several hours. DO NOT FREEZE THE SAMPLES.

10- Remove the plates from the ice and pat themdry (especially the base that was in contact with the ice) with a paper towel.
Warning: Residual water may impair the drying process.
11- Carefully remove the caps from all the strips of tubes from a plate. Make sure that allleaf samples stay in their tubes.
12- Remove the plasticfilmcovering the adhesive side of one sheet of porous tape.
13- Cover the opening of the tubes with the sheet of porous tape, making sure that it sticks well on the opening/ridge of each tube. Fold the excess tape on the side of the outermosttubes.
14- Create one 3-mm hole per tube by punching through the porous tapewith the tweezers and moving them back and forth a few times (Fig.4).
Warning: If the holes are smaller than 2 mm, this will impair the drying process.
15- Add half of the desiccant from the bottle (about 100g) in the large Ziploc bag.
16- Place the two plates in the bag. Place the plates upside-down, with the porous tape directly on the desiccant (Fig.4).
17- Remove as much air as possible from the bag and carefully seal the bag, making sure that the zipis sealed all the way so that the bagbecomesanairtight container.
18- After 24 and 48 hours, check the color of the desiccant. If all or most of it is pink or lilac (Fig.5), replace it with the additional 100g of fresh desiccant (blue) provided in the bottle.
19- Send the bag containing the 2 plates as well asall the caps back for DNA extraction and sequencing:

Fig.s

Fig. 1. Identification of tubes using a column and row system.

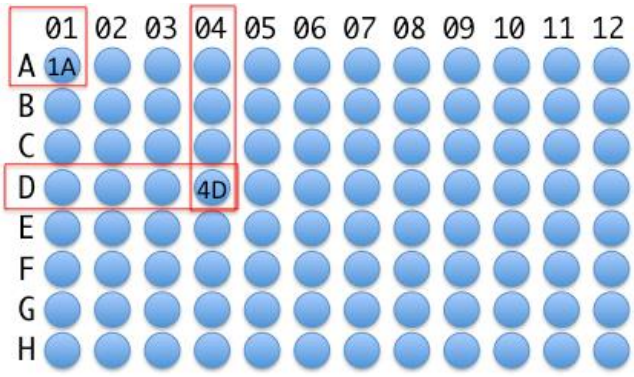


Fig. 2. Collecting leaf samples.

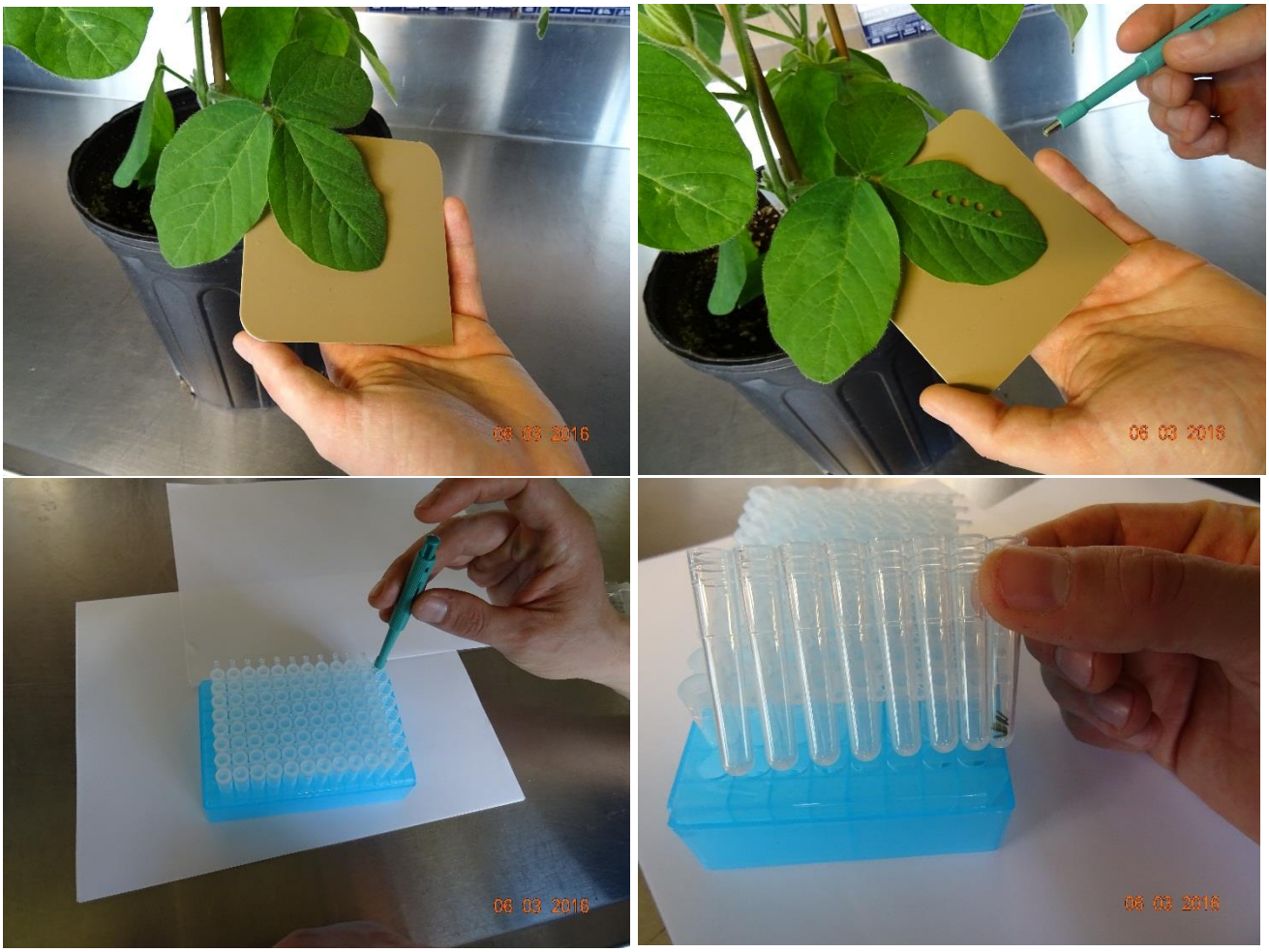


Fig. 3. Cleaning the punch.

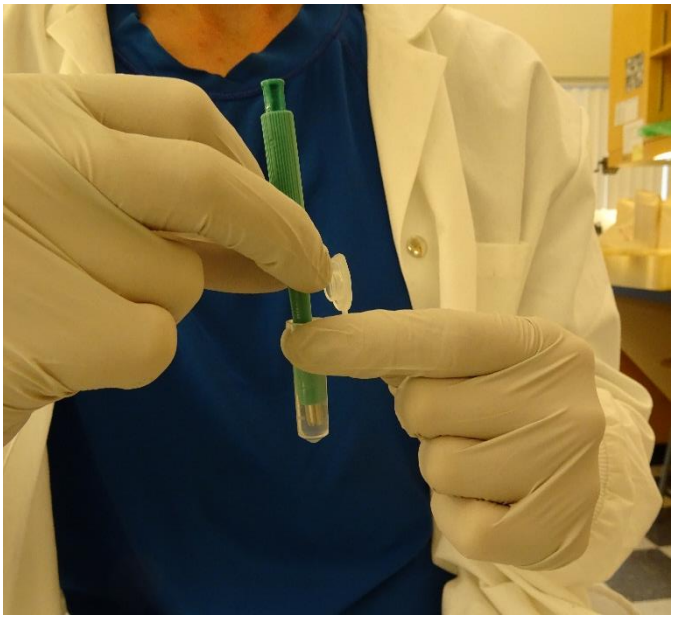


Fig. 4. Preparing plates for drying the leaf samples.

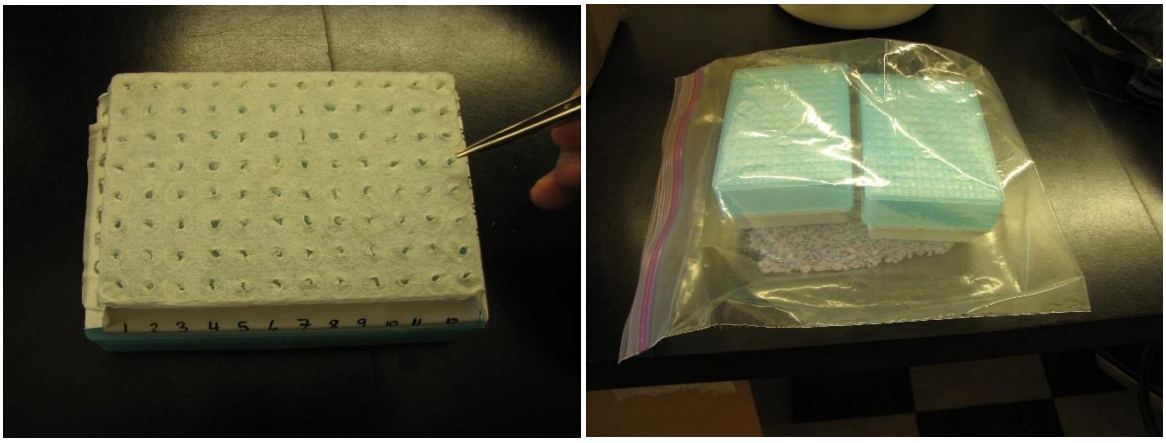


Fig. 5. Color change of the desiccant (right: blue = dry/active; left: lilac = moist/inactive).

Appendix 4. Physico-chemical composition of potted soils for the evaluation of acid-tolerant wheat lines in the greenhouse

Elements	Values
pH (KCl)	5.7
Resistance (Ohms)	3520
Texture	Sand
Calcium (cmol(+)/kg)	0.56
Magnesium (cmol(+)/kg)	0.08
Potassium (mg/kg)	16
Sodium (mg/kg)	5
P(citric acid) (mg/kg)	39
Total cations (cmol(+)/kg)	0.71
Copper (mg/kg)	0.07
Zinc (mg/kg)	0.21
Manganese (mg/kg)	3.98
Aluminium (mg/kg)	0.00
Boron (mg/kg)	0.01
Carbon (%)	0.04
Sulfur (mg/kg)	5.9
P(Olsen) (mg/kg)	2
Coarse sand (%)	35
Medium sand (%)	32
Fine sand(%)	25
Clay (%)	4
Silk (%)	4
Klip (%)	44

Preparation of Mix for PCR reaction

	x 1µl	x13	x26	x19
dH_2O (PCR grade water)	5,75	74,75	149,5	109,25
2G Fast Multiplex Mix	6,25	81,25	162,5	118,75
Primer F (10µM)	0,5	6,5	13	9,5
Primer R (10µM)	0,5	6,5	13	9,5
Template DNA (~100 ng/µl)	1			
Total volume	**14**			

PCR reaction

94 °C……..3min — Denaturation

94 °C……1min
Ta (°C)……..1min — Annealing
72 °C……..2min

72 °C……..10min — Extention

2G Fast Multiplex PCR contain all the components for Fast Multiplex PCR except primers and template DNA;

The novel 2G Fast Hotstart DNA polymerase (supplied at 1 U/25µl rxn) offers faster cycling tomes and and improved reaction efficiency for the even amplification of all target fragment;

2G Fast Multiplex contains 0.2 mM of each dNTP and 3.0mM $MgCl_2$ at 1x final concentration;

Use 0.2 µM of eact primer, and 10-100 ng of template DNA.

Anneal for 30sec at 60 °C;

Microsatellite Solutions and Dilutions

I. Gel mix and associated reagents

40% Acrylamide stock solution

Add the following to a 250 ml Shott flask:
- 76 g Acrylamide
- 4 g Bis-acrylamide

Make up to 200 ml with H2O.
Cover the flask with aluminum foil and store @ 4 °C for up to 2 weeks.

6% sequencing gel mix (6 M urea, 1 X TBE)

Add the following to a 250 ml Schott flask:
- 37.5 ml 40% acrylamide stock solution
- 90.09 g urea
- 50 ml 5 X TBE

Stir until urea is completely dissolved and make up to 250 ml with dH20.
Cover the flask with aluminum foil and store @ 4 °C for up to 1 week.

10% Ammonium persulphate

Dissolve 0.1 g APS in 1 ml dH2O in a 2.2 ml eppie. Aliquot and store for up to 1 month.

GEL:

Prepare the gel by adding 800 µl 10% APS and 160 µl TEMED to 160 ml of 6% gel mix in a glass beaker. Mix well and cast the gel. Leave at least 1 H to set.

II. Plate preparation

Plate glue

Dilute 125 µl plate glue into 25 ml 100% ethanol. Dilute this stock 1:3 (500 µl: 1500 µl) in 100 % ethanol. Add 1740 µl plate glue to 140 µl **10 %** Acetic acid.

Long glass plate: Clean with 100 % ethanol and use a kimwipe to wipe the entire plate with "C-thru". Leave entire plate for 3 min until dry and wipe entire plate with dry kimwipe until shining.

Short glass plate: Clean with 100 % ethanol and rub entire plate with plate glue (± 2. 0 ml). Leave 30 s to dry. Wipe plate immediately.

III. Loading of samples

Pre run gel at 70 W for 30 min.

Add equal volumes of AFLP loading buffer (98% formamide, 10mM EDTA pH 8.0, 0.05% w/v bromo phenol blue, 0.05% w/v xylene cyanol FF) to PCR samples. Denature samples at 95°C for 5 min and immediately quench on ice. Load 12-15 µl of each sample and 1 µl 100 bp ladder.

Perform electrophoresis at a constant power of 70 W for approximately 5 ½ hours.

IV. Silver Staining

Fixing solution (Fresh)
210 ml Ethanol
10.5 ml Acetic acid
1879.5 ml dH2O

Developing solution (Fresh)
31.5 g NaOH
2100 ml dH2O
8.505 ml Formaldehyde (add just before use)

Staining solution (Use up to 3 times)
2.1 g AgNO3
2100 ml dH2O

(Add 10g NaCl to 2 L silver stain to precipitate and dispose after use)

Staining procedure

Fixing solution (20 min on belly dancer)

↓

2 X 5 min 2 L dH2O rinses (on belly dancer)

↓

Staining solution (20 min on belly dancer)

↓

10 second dH20 rinse

↓

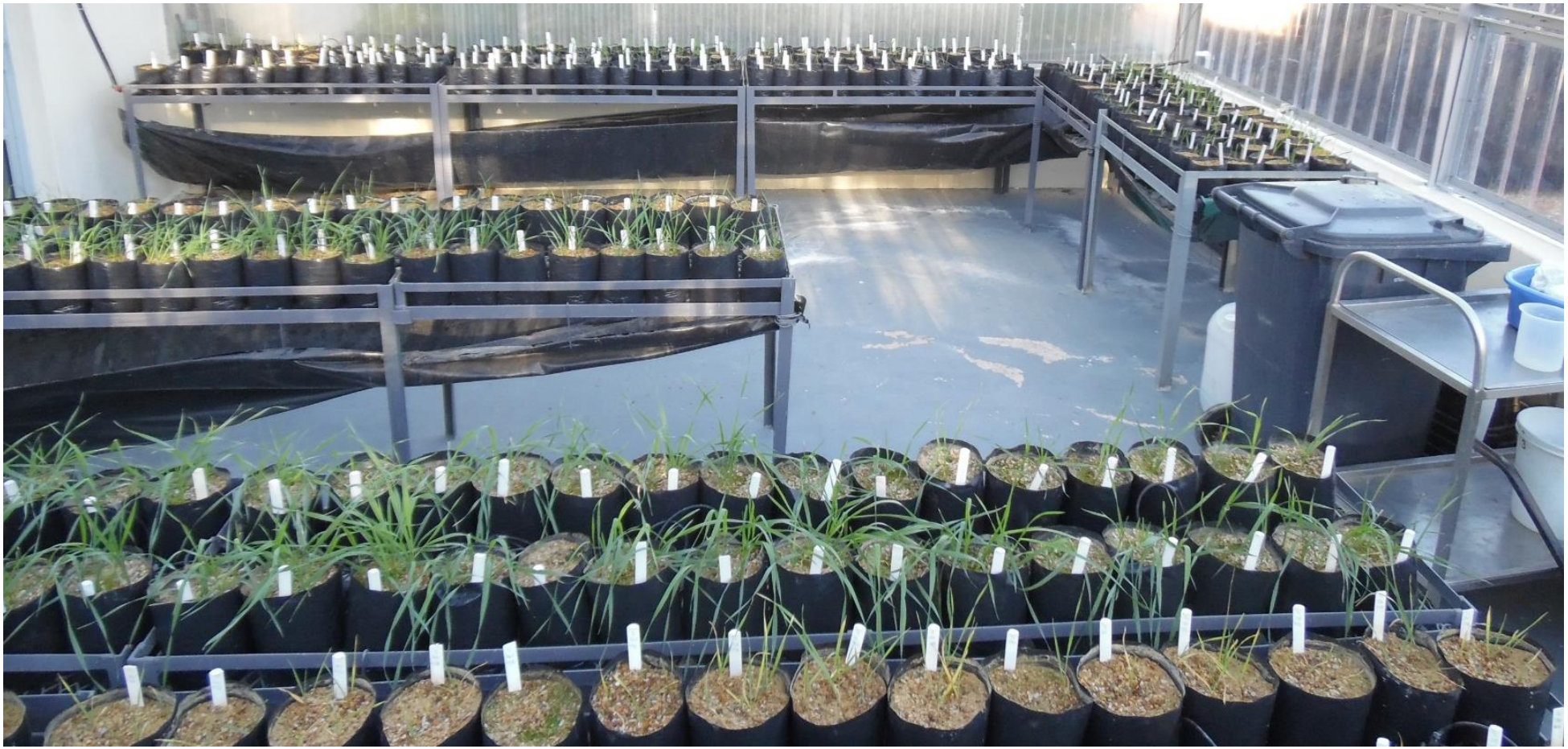
Developing solution-until bands appear (belly dancer)

↓

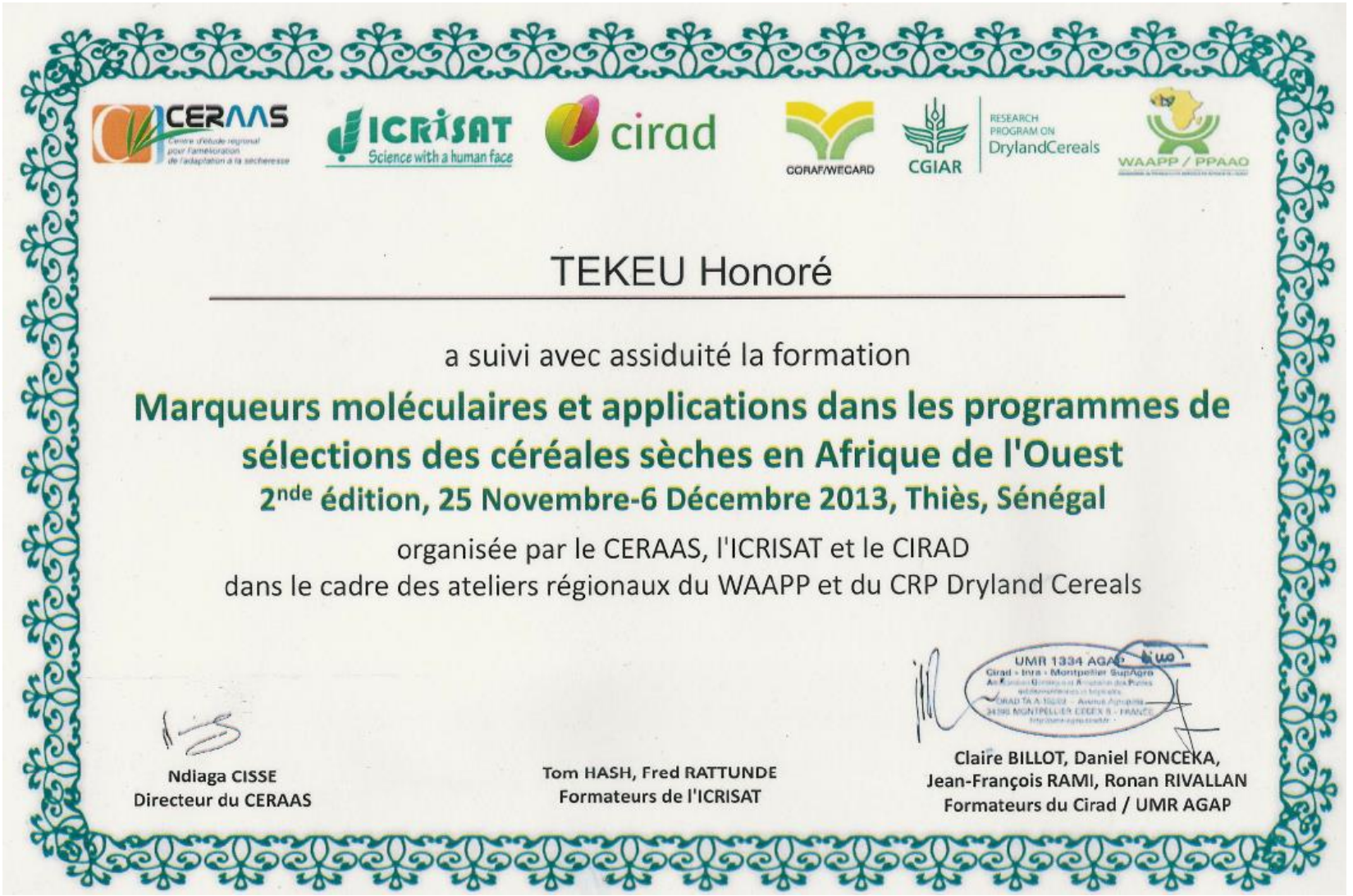
dH20 rinse

↓

Take digital photo and/or cover with plastic film



Appendix 5. Global system for phenotyping wheat lines for tolerance to acid soils in the greenhouse



Appendix 6. Certificate of completion, Senegal 2013

COMITE CAMEROUNAIS DES BIOSCIENCES

(PRIX C.C.B)

(CBS Excellence Award)

Le Comité Camerounais de Biosciences (CCB) certifie que,
The Cameroon Bioscience Society (CBS), hereby certifies that,

Mme/Mlle/M. : ..TEKEU..Honoré.....

a obtenu le 1er Prix de la meilleure Communication orale jeune chercheur lors de la 21ème Conférence Annuelle organisée du 04 au 06 Décembre 2014 à l'Université de Yaoundé 1/ENS-Cameroun.
Has obtained excellence price for the best oral communication (1st) at the 21st annual conference of the Society organized from the 04th to the 06th of December at the University of Yaounde 1/ENS-Cameroon.

Thème: Biosciences, OGMs et Biorévolution
Bioscience, GMO and Biorevolution

En foi de quoi la présente attestation lui est délivrée pour servir et valoir ce que de droit.
In testimony whereof this certificate is issued to serve any useful purpose.

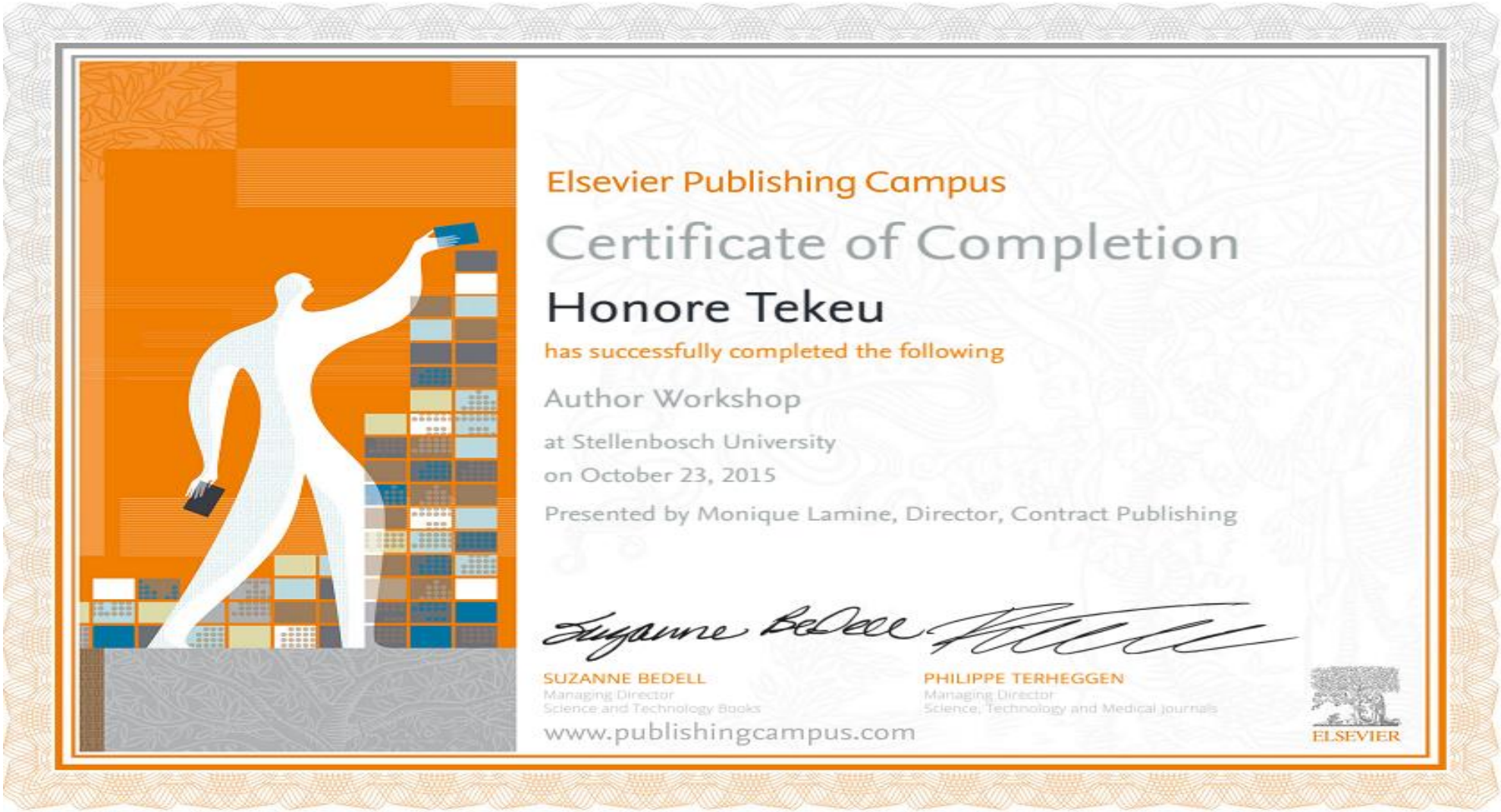
Le Président du Comité local d'organisation

Pr. Elie Fokou
Nutrition et Sciences Alimentaires

Le Président du CCB

Nicolas Niemenak
Associate Professor

Appendix 7. Winner of the 1st Prize of Excellence for the Best Oral Communication young researcher at the 21st annual conference of BIOSCIENCES 2014



Appendix 8. Certicate of completion, Elsevier 2015

COMITE CAMEROUNAIS DE BIOSCIENCES (CCB)

Deuxième (2ème) Prix Communication Orale
Second (2nd) Prize of Oral Communication

Le Comité Camerounais de Biosciences (CCB) certifie que,
The Cameroon Bioscience Society (CBS), hereby certifies that,

Mme/Mlle/M. : .....TEKEU Honoré.....

a reçu le Deuxième (2ème) Prix Communication Orale de la 22ème Conférence Annuelle organisée du 03 au 05 Décembre 2015 à l'IRAD de Nkolbisson, Yaoundé - Cameroun.
Received the Second (2nd) Prize of Oral Communication of the 22nd annual conference organized from December 3rd to 5th 2015 at IRAD of Nkolbisson, Yaounde - Cameroon.

Thème: Biosciences Innovantes pour une Agriculture Compétitive et Durable
Innovative Biosciences For a Sustainable and Competitive Agriculture

En foi de quoi la présente attestation lui est délivrée pour servir et valoir ce que de droit.
In testimony whereof this certificate is issued to serve any useful purpose.

Le Président du Comité Local d'Organisation

Ngomkeu M. Eddy L. (PhD)

Le Président du CCB

Nicolas Niemenak
Associate Professor

Appendix 9. Winner of the 2nd Prize of Excellence for the Best Oral Communication at the 22nd annual conference of BIOSCIENCES 2015

PUBLICATION

academicJournals

Vol. 16(36), pp. 1832-1839, 6 September, 2017
DOI: 10.5897/AJB2017.16090
Article Number: D0304ED65883
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Genetic diversity of Cameroonian bread wheat (*Triticum aestivum* L.) cultivars revealed by microsatellite markers

Honoré Tékeu[3], Eddy M. L. Ngonkeu[1,3], François P. Djocgoué[3], Aletta Ellis[2], Venasius Lendzemo[1], Lezaan Springfield[2], Lionel Moulin[5], Agnieszka Klonowska[5], Diégane Diouf[4], Willem C. Botes[2]* and Gilles Béna[5]

[1]Institute of Agricultural Research for Development, P. O. Box 2123, Yaoundé, Cameroon.
[2]Department of Genetics, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, 7602, South Africa.
[3]Department of Plant Biology, Faculty of Science, University of Yaoundé I, P. O. Box 812, Yaoundé, Cameroon.
[4]Laboratoire Commun de Microbiologie de Dakar, Centre de recherche IRD/ISRA/UCAD, Route des Hydrocarbures, Bel Air, P. O. Box 1386, Dakar, Sénégal.
[5]IRD, CIRAD, Laboratoire des Interactions Plantes-Microorganismes-Environnement (IPME), Université de Montpellier, 911 Avenue Agropolis 34394 Montpellier, France.

Received 29 May, 2017; Accepted 4 August, 2017

The assessment of genetic diversity is a key prerequisite for studying the adaptation of populations to new environmental conditions, and therefore for the selection of new varieties. The present investigation aimed to estimate the levels and genetic structure within bread wheat varieties grown in Cameroon. Thus, genetic diversity was assessed in 17 hexaploid wheat cultivars, using 11 microsatellite markers. Genetic resources were collected in the Northwest, Adamawa and North Regions. All pairs of specific marker loci used gave amplifications with allelic variations of size on all DNA of wheat accessions. A total of 77 alleles were detected among cultivars and the number of alleles per locus ranged from 2 to 13 with an average of 7, comparable to those observed in most previous studies. Gene diversity ranged from 0.46 (Xgdm 125) to 0.90 (Xgwm 177) with an average of 0.88, increasing with the number of alleles, with a correlation coefficient of 0.88 (Adamawa) and 0.76 (Northwest). Microsatellite markers used had an average value of polymorphic information content (PIC) of 0.69, indicating that these markers are highly informative in this study. These markers are valid and will make a contribution to the studies in hexaploid wheat. Moreover, cluster analysis at a genetic similarity of 80% and the principal component analysis, where the first two components explaining 59.86% of variation which structured 17 accessions in 5 main distinct groups. This high diversity revealed among wheat accessions, grown in Cameroon could be used in the breeding programs.

Key words: Genetic diversity, bread wheat (*Triticum aestivum* L.), simple sequence repeats (SSR), Cameroon.

INTRODUCTION

The global demand for wheat yields has been estimated to increase by 50% in 2050, in order to feed the world's

growing population (Grassini et al., 2013; Allen et al., 2017). To meet this demand, wheat production should be increased through agricultural intensification in cropping regions areas. To this end, Rajaram and Hettel (1994) had delimited 12 Mega environments (MEs) for wheat cultivation, where three correspond to several agroecological areas in Cameroon. Among them, the main Cameroonian areas are in the North, North-West and Adamawa Regions.

Wheat production is a means of subsistence for many families in Cameroon. Indeed, the cultivation of wheat has started in Cameroon since 1975, through the Development Society for the Cultivation and Processing of Wheat (SODEBLE). Located in Wassandé (Adamawa's Region), the SODEBLE has grown wheat, converted wheat into flour, marketed and carried trials, in order to improve the production technics. Before its closure, this Company had produced high yielding wheat lines resistant to major fungal diseases (Monthé Biris and Habas, 1980). Twelve of these bread wheat varieties were evaluated for the agronomic traits in the North-West Region (Ayuk-Takem, 1984).

In addition, Ayuk-takem (1984) evaluated the agronomic characteristics of 12 varieties of bread wheat in the Northwest Region to identify high yielding varieties for Bui and other agro-ecological zones in high altitudes in Cameroon. The study showed that the local variety (IRAB-1) had the highest yield (4.1 t/ha), but with a non-significant difference with the varieties Chris Mutageneuse (3.5t/ha) and wheat Blésil 430 (4 t/ha). However, the yields of these three varieties were significantly better compared to all other tested varieties. These varieties had also been subjected to various agronomic tests in 1985/1986. In doing so, certain varieties had not been made available to Cameroonian farmers. Until today, beyond these agronomic evaluations, no studies have ever been carried out on the genetic variability of wheat cultivars grown in Cameroon.

Evaluating the genetic diversity is a prerequisite for studying the adaptation of populations to new environmental conditions and hence for the selection of new varieties. The loss of genetic diversity due to modern breeding practices has been reported in several studies (Fu et al., 2005). Several authors have shown that the narrowness of crop genetic diversity could lead to increased susceptibility to diseases and pests, as well as the inability of plants to respond to different environmental constraints (Gorji and Zolnoori, 2011). Therefore, it is necessary to estimate the level of genetic diversity within existing varieties to serve as a base for strategies development geared in the management and exploitation of genetic resources.

In this context, the use of molecular markers to assess genetic diversity is necessary because, unlike phenotypic markers, they are independent from environmental effects (Reza et al., 2015). Several markers, independently or in combination with others, were efficiently used for wheat genetic diversity analyses, including morphological traits (Sonmezoglu et al., 2012). Randomly amplified polymorphic DNAs (RAPDs) (Mukhtar et al., 2002), amplified fragment length polymorphisms (AFLPs) (Reza et al., 2015), restriction fragment length polymorphism (RFLPs) (Bohn et al., 1999) and diversity array technology (DArT) markers have recently been developed and used for genetic diversity assessment and mapping (Ryan et al., 2009), as well as Single nucleotide polymorphisms (SNPs) (Froese and Carter, 2016).

On the other hand, the use of simple sequence repeats (SSRs) markers combines with many desirable marker properties such as abundance, high levels of polymorphism (unlike RFLP), very good reproducibility (compared to RAPD), and co-dominance (contrary to the AFLP for which codominance is not exploitable), but also an even coverage of the genome and the specificity of amplification. In wheat, SSRs markers have been used successfully in a wide range of applications such as genotype identification (Prasad et al., 2000), diversity studies (Akfirat and Uncuoglu, 2013) and genetic mapping. This present study aimed to assess the level of genetic diversity of bread wheat accessions grown in Cameroon.

MATERIALS AND METHODS

Plant material and genomic DNA extraction

The plant material consists of 17 cultivars of bread wheat (*Triticum aestivum* L.) collected in six villages located in two Regions of Cameroon (Table 1). Among them, 11 accessions were collected in five villages of Northwest and 6 were collected from one village (Wassande) of Adamawa region. The cultivars of the North West are mainly local seeds, whereas those of Adamawa were originally given by the SODEBLE and some others were imported from Tchad. In our study, we collected all materials used by farmers in those regions.

An adjusted Doyle and Doyle (1990) protocol was used to extract genomic DNA (gDNA) from seedlings at the two to three leaf stage.

Microsatellite markers and PCR amplification

Eleven wheat microsatellite markers for 11 loci located in the chromosomes 1A, 2A, 2D, 3A, 3B, 4D, 5D, 6B and 7D, were used for genetic diversity analysis. Xgwm and Xwmc markers were obtained, respectively from Röder et al. (1998) and Somers and Isaac (2004; Grain Genes).

PCR reactions were carried out in 14 µl reaction mixtures of KAPA2GTM Fast Multiplex PCR Mix, 6.25 µM of each forward and

*Corresponding author: E-mail: wcb@sun.ac.za.

Table 1. Wheat cultivars used and their origins in Cameroon.

S/N	Samples' names	Local name	Village	Region
1	Ngm 2	Ngm 2	Wassande	Adamawa
2	Fuanb2	Fuanb2	Fuanentui	Northwest
3	Babankit	Babankit	Smal Babanki	Northwest
4	Alexander wonder	Alexander wonder	Boyo	Northwest
5	Fuanb1	Fuanb1	Fuanentui	Northwest
6	Sonalika	Sonalika	Wassande	Adamawa
7	Fuanb3	-	Fuanentui	Northwest
8	Fuanb4	-	Fuanentui	Northwest
9	HGW	Hard wheat	Abongphen	Northwest
10	BBT2		Abongphen	Northwest
11	WASSANDE 2	WASSANDE 2	Wassande	Adamawa
12	Vrack	Vrack	Bambui	Northwest
13	Ngderem4	-	Wassande	Adamawa
14	Ngderem1	Ngderem1	Wassande	Adamawa
15	Ngderem3	-	Wassande	Adamawa
16	IRAT 10	IRAT 10	Bambui	Northwest
17	RIBA	RIBA	Boyo	Northwest

reverse primer, 1 µl gDNA and dH_2O. The PCR cycling conditions was set at 94°C for 3 min of denaturation, followed by 45 cycles of 1 min at 94°C, 1 min at the annealing temperature (Ta), 2 min at 72°C and then 72°C for 10 min for extension.

The PCR products were electrophoresed on 6% non-denaturing polyacrylamide gels containing 1xTBE (Tris Borate EDTA). The amplified band sizes for each SSR locus were determined on the basis of their migration relative to the 50 bp marker.

Data analysis

The molecular diversity within all accessions was estimated for each SSR locus, using the Power Marker 3.25 software (Liu and Muse, 2005). To measure the informative character of the SSR markers, the PIC for each marker was calculated using the formula of Nei (1973):

$$PIC = 1-\Sigma_{i=1}^{k} P_i^2$$

Where, k is the total number of alleles detected per locus and Pi the frequency of the allele i in all 17 accessions.

Genetic similarity (GS; Dice, 1945) was calculated as:

$$GS = 2N_{ij}/(N_i + N_j)$$

Where, N_{ij} is the number of fragment common to individual i and j, and (Ni + Nj) is the total number of fragment in both individuals.

Genetic distance (GD) among group pairs was calculated following Nei and Li (1979),

$$(GD_{xy}) = 1-(2N_{xy}/N_x + N_y)$$

The dendrogram was constructed using the method based on the genetic distance (SAHN method, UPGMA algorithm) of the 17 accessions and using the software Statistica 12. To calculate allele frequency (A_{xy}) from one of the variation to another in each locus, the formula of Khlestkina et al. (2004) was used:

$$A_{xy} = \Sigma|P_{xi} - P_{yi}|N_{xy}$$

Where, P_{xi} and P_{yi} are the frequencies of the i^{th} allele in regions X and Y, respectively, and N_{xy} is the total number of alleles for the two groups X and Y. The allelic frequency variation was calculated separately for each of the 11 loci and then for all of them as an average. All fragments were used to generate GS matrix for Principal Component Analysis (Sneath and Sokal, 1973).

RESULTS

Characteristics of markers and genetic diversity

All pairs of primers specific for SSR locus used resulted in a positive amplification with allelic variations in size on all DNA of wheat accessions. A total of 77 microsatellite alleles were detected. The number of alleles per locus varied from 2 (Xgwm 125 and Xgwm 331) to 13 (Xwmc 177), with an average of 7 alleles per locus. Genetic diversity for microsatellite loci ranged from 0.46 (Xgdm 125) to 0.90 (Xgwm 177) with an average of 0.88. The polymorphism information Content (PIC) varied from 0.25 (Xwmc 331) to 0.89 (Xwmc 177), with an average of 0.69 (Table 2).

The results indicated a significant correlation ($P < 0.01$) between gene diversity and number of alleles across wheat accessions in both Regions (Figure 1). The correlation coefficient between these two variables over the 11 loci were 0.88 (Adamawa) and 0.76 (Northwest).

Genetic relationship and diversity among different geographical regions

Genetic distance value (GD) indicates that some

Table 2. Description of SSR Markers.

Locus	Chromosome position	Primers sequences	Repeat	Bases expected	Annual temperature	Alleles frequency	Number of alleles	Gene diversity	PIC
Xwmc 11	1A, 3A	5' TTGTGATCCTGGTTGTGTTGTGA 3' 5' CACCCAGCCGTTATATATGTTGA 3'	CT	177	61	0.29	8	0.83	0.81
Xwmc 59	1A. 6A	5' TCATTCGTTGCAGATACACCAC 3' 5' TCAATGCCCTTGTTTCTGACCT 3'	(CA)19	197	58	0.18	10	0.89	0.87
Xwmc 177	2A	5' AGGGCTCTCTTTAATTCTTGCT 3' 5' GGTCTATCGTAATCCACCTGTA 3'	(CA)21	184	52	0.18	13	0.90	0.89
Xgwm 190	5D	5' GTGCTTGCTGAGCTATGAGTC 3' 5' GTGCCACGTGGTACCTTTG 3'	(CT)22	201-253	55	0.18	9	0.87	0.86
Xgwm 437	7D	5' GATCAAGACTTTTGTATCTCTC 3' 5' GATGTCCAACAGTTAGCTTA 3'	(CT)24	109-111	47	0.18	10	0.88	0.87
Xgwm 539	2D	5' CTGCTCTAAGATTCATGCAACC 3' 5' GAGGCTTGTGCCCTCTGTAG 3'	(GA)27	143-157	60	0.24	8	0.83	0.81
Xdgm 125	4D	5' GCAGGCGTGTTACTCCAAGT 3' 5' CCGAGGTGGATAGGAGGAAA 3'	-	-	60	0.65	2	0.46	0.35
Xwmc 331	4D	5' CCTGTTGCATACTTGACCTTTTT 3' 5' GGAGTTCAATCTTTCATCACCAT 3'	-	128	61	0.82	2	0.29	0.25
Barc 133	3B	5' AGCGCTCGAAAAGTCAG 3' 5' GGCAGGTCCAACTCCAG 3'	(CT)24	-	-	0.65	4	0.52	0.47
Xgwm 133	6B	5' ATCTAAACAAGACGGCGGTG 3' 5' ATCTGTGACAACCGGTGAGA 3'	(CT)39	-	-	0.35	4	0.72	0.67
Xgwm 644	6B	5' GTGGGTCAAGGCCAAGG 3' 5' AGGAGTAGCGTGAGGGGC 3'	(GA)20	-	-	0.29	7	0.79	0.76
Mean	-	-	-	-	-	0.36	7	0.72	0.69

Xgwm and Xwmc markers were obtained respectively from Röder et al. (1998) and Somers and Isaac (2004; Grain Genes); PIC, Polymorphism information content.

accessions are closely related. The GD over accessions in all regions ranged from 0.18 (between Wassande2 and NGDEREM3) to 1 with a mean of 0.8 (80%). So, at 80% of genetic divergence, the 17 wheat cultivars studied were structured into 5 main groups (A, B, C, D and E) in the dendrogram based on the UPGMA analysis using SSR data (Figure 2). Group A included 4 cultivars (Alexander wonder, Riba, Vrack and FUANB3).

Very close to 80% of genetic dissimilarity, the group B could be divided into two subgroups: subgroup B1 contained 6 accessions (BABANKIT, FUANB1, FUANB4, FUANB2, BBT2 and HGW) while subgroup B2 included only one cultivar (IRAT 10). It is noteworthy that the two varieties FUANB1 and FUANB4 are identical. Group C contained 4 cultivars (Ngderem1, Ngderem3, WASSANDE 2 and NGM2) while Groups D and E each contained 1 cultivar, respectively (Ngderem4 and SONALIKA).

Furthermore, the principal component analysis (PCA) for the six-collection village split the accessions into five clearly distinct groups. The first two principal components had Eigen values of 6.36 and 3.81. The PCA grouped the 17 wheat accessions into various components with the first two explaining 59.86 and 37.44% of the total variation. Accessions from each village were approximal clustered together (Figure 3). So, 80% of the genetic material from the same geographical village could be clustered in specific groups.

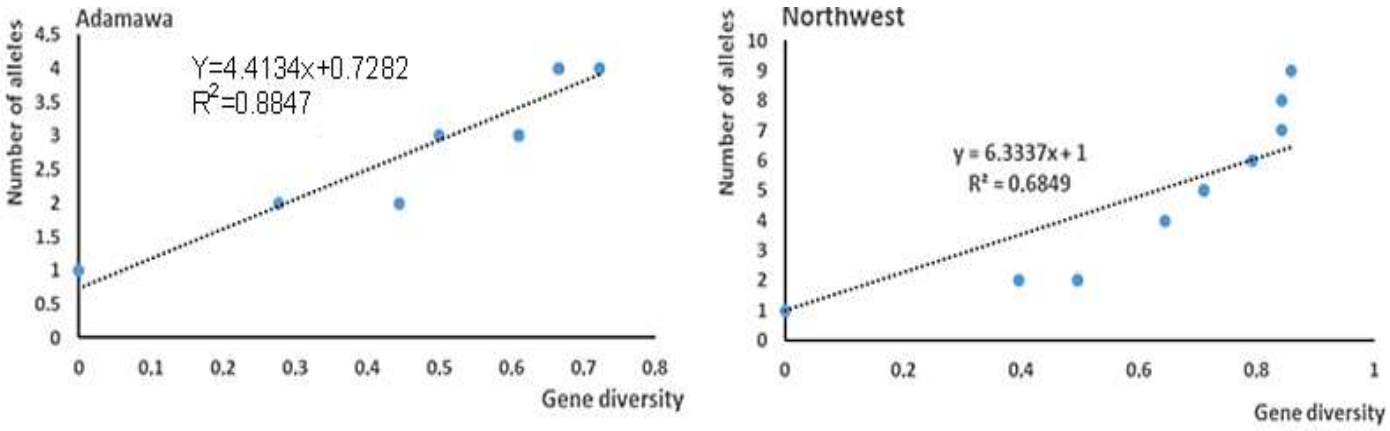


Figure 1. Correlation between gene diversity and the number of alleles over 11 microsatellite loci in hexaploid wheat.

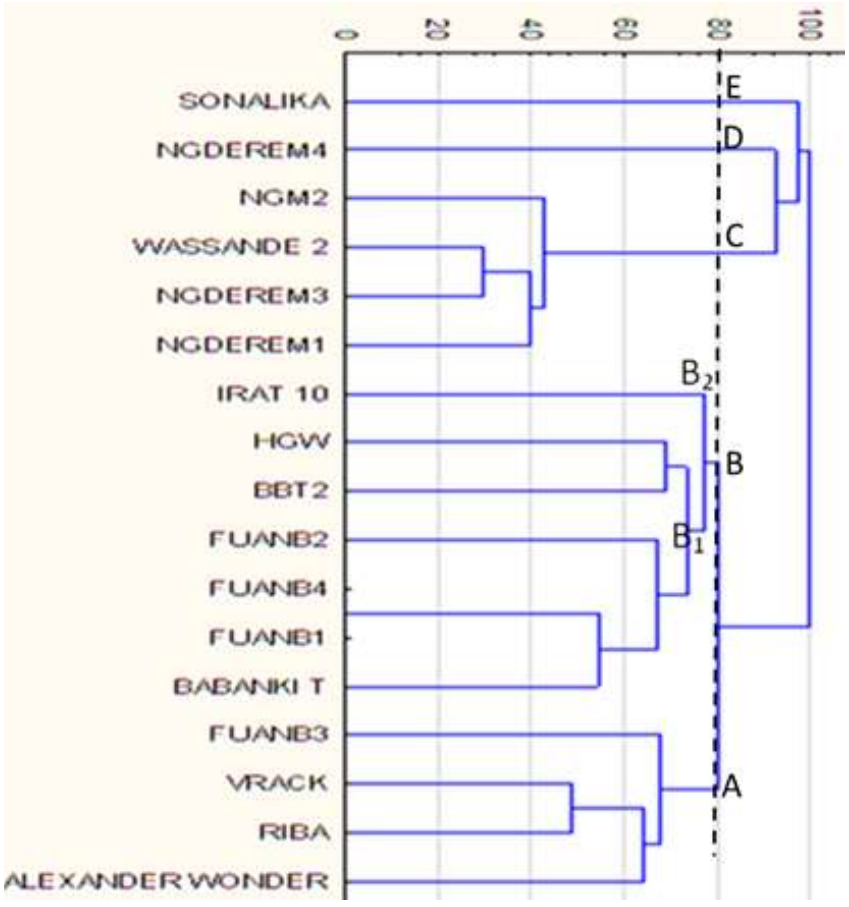


Figure 2. Grouping according to the dissimilarities between 17 accessions of hexaploid wheat on the basis of the SSR profiles of 11 loci.

Accessions were then analyzed separately according to their region of origin (Adamawa and Northwest). A comparison of the genetic diversity of wheat accessions was done between two germplasm pools. The mean of gene diversity, number of alleles per locus, total number of alleles and the number of accessions carrying rare alleles were higher in Northwest, compared to those in Adamawa Region (Table 3). These results suggest that

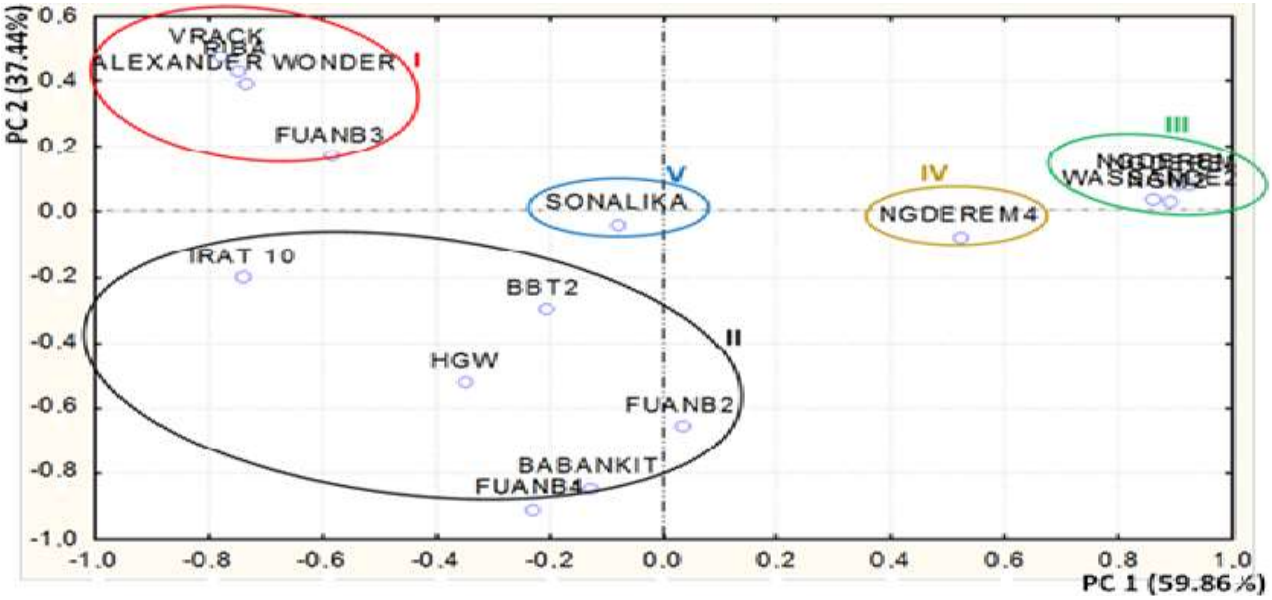


Figure 3. Principal component analysis of 17 hexaploid wheat accessions from 6 villages in Cameroon. The grouping is based on Dice's similarity coefficients.

Table 3. Analysis of geographical regions.

Item	Northwest	Adamawa
Number of accessions	n = 11	n= 6
Total number of alleles	51	34
Average number of alleles per marker	4.64	3.09
Number of rare alleles	5	6
Mean of PIC-values	0.57	0.48
Mean of Gene diversity	0.62	0.54

the Northwest area exhibited greater genetic diversity than Adamawa region, even after taking into account the effect of collection size.

DISCUSSION

Diversity of SSR markers

In the present study, 11 microsatellite markers revealing 77 alleles allowed to discriminate 17 cultivars of hexaploid wheat collected in Cameroon. The number of alleles per locus ranged from 2 to 13 with an average of 7. Röder et al. (2002) detected an average of 10.5 alleles per locus from 502 recent European wheat varieties, using 19 microsatellite markers. Khaled et al. (2015) used 17 SSR markers to assess genetic diversity of 33 genotypes of hexaploid wheat from Egypt and detected an average of 5.59 alleles per locus.

The average number of alleles per locus in the study is thus comparable to those observed in previous studies. In addition, the microsatellite markers we used had an average PIC value of 0.69, which means that these markers are highly informative in our study. Indeed, Botstein et al. (1980) reported that a PIC value higher than 0.5 is considered to be a sign of a very informative marker, while 0.5> PIC> 0.25 corresponding to an informative marker. In previous studies, Röder et al. (2002) found an average PIC value of 0.67 in 500 genotypes. The choice of these SSR loci is therefore relevant for our study.

Genetic relationship between wheat cultivars

Cluster analysis discriminated all cultivars of Cameroonian hexaploid wheat into five main groups. Most cultivars clustered according to their geographical location. Indeed, all accessions from groups B and C are cultivated in the Northwest and Adamawa Regions,

respectively. Similarly, the 4 varieties of group A were collected in the Northwest Region. Moreover, the variety NGDEREM4 of group D comes from Adamawa and the SONALIKA variety of group E was introduced in Cameroon since 1975 through the SODEBLE Company, from Mexico. Huang et al. (2002) reported that the genetic diversity of hexaploid wheat was not completely related to geographic distribution. They also reported that, these results might be explained by the fact that similar genetic variation occurred independently in the different geographic regions or that artificial transfer of accessions from one region to others resulted in a false determination of the geographic origin.

Similar results were obtained by Khaled et al. (2015) in hexaploid wheat genotypes grown in Egypt. On the other hand, Al-Khanjari et al. (2007) found that all local varieties of wheat from the same geographical area clustered in the same group. In our case, we can hypothesize that the genetic proximity of the cultivars based on their geographical origin results from a local selection and diversification, coupled with weak or non-existent exchanges of seeds between regions, inducing a geographical structuration and a form of isolation by distance.

The overall gene diversity increased with the number of alleles at a given locus. We found significant correlation between gene diversity and the number of alleles in Adamawa ($r = 0.88$, $P < 0.01$) and Northwest ($r = 0.76$, $P < 0.01$). Therefore, the number of alleles could be used for the assessment of genetic diversity in hexaploid wheat. Similar results were found by Huang et al. (2002) in a set of 24 microsatellite markers used to characterize 998 accessions of hexaploid wheat germplasm. Consequently, these authors reported that the characterization of a reliable correlation coefficient needs a large sample size. The results in the present study disagrees with those reported by Prasad et al. (2000) who indicated that the polymorphism information content value was not correlated with the number of alleles in 55 wheat accessions. According to Huang et al. (2002), the number of alleles was also correlated with the repeat number of microsatellite DNA and its relative distance from the centromere. It has been suggested that the three mechanisms for creating a new allele at SSR loci are slippage replication (Tachida and Iizuka, 1992), unequal crossing-over and genetic recombination (Harding et al., 1992). The value of genetic distance (GD) indicated that some accessions were closely related. Averages of GD over accessions in all regions were ranged from 0.18 (between Wassande2 and NGDEREM3) to 1 with a mean of 0.8 (80%). The high GD coefficient values indicate the presence of high gene diversity in the accessions.

The mean of gene diversity was relatively higher in Northwest (0.62) compared to Adamawa (0.54). These results suggested that the Northwest exhibited greater genetic diversity than Adamawa region. The Northwest was the presumed center of origin of hexaploid wheat in Cameroon and Adamawa was the sites were the SODEBLE was established. The results obtained in our study provided new information on the relationships between the Cameroonian bread wheat cultivars.

The set of the used microsatellite markers showed a high level of polymorphism and sufficient information to discriminate the cultivars of hexaploid wheat grown in Cameroon. Generally, our study provides a first description about the molecular genetic diversity of Cameroonian wheat varieties. The results are consistent with expectations and provide a first base for further investigations. The important level of the genetic diversity reported in the present study should be taken into account in developing wheat breeding programs in agro-ecological zones of Cameroon. Morphological and phenotypic studies will also be required to couple our results of molecular analyzes.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGMENTS

This collaborative work was supported by the « MIC-CERES » project (Microbial eco-compatible strategies for improving wheat quality traits and rhizospheric soil sustainability) jointly supported by Agropolis Fondation (grant AF Project ID 1301-003) through the "Investissements d'avenir" programme (with reference number ANR-10-LABX-0001-01) and Fondazione Cariplo (grant FC Project ID 2013-1888), and European Commission through PAFROID projects (*INTRA-ACP Program, lot Africa for an action N°2013- 4644 / 001 – 001 with the reference: 384201-EM-1-2013-1-MG-INTRA_ACP*). The authors are grateful to the Plant Breeding Laboratory staff of Stellenbosch University for their technical support during the molecular analyses in laboratory. They would like to thank Dr. Apollinaire Tagne, Yves H. Tchiechoua, Adrienne N. Ngom, Aga Valentine Enwe and Tume George Tardzenyuy for their technical assistance during the samples collection.

REFERENCES

Akfirat FS, Uncuoglu AA (2013). Genetic diversity of winter wheat (*Triticum aestivum* L.) revealed by SSR markers. Biochem. Genet. 51:223-229.

Al-Khanjari S, Hammer K, Buerkert A, Röder MS (2007). Molecular diversity of Omani wheat revealed by microsatellites: II. Hexaploid landraces. Genet. Resour. Crop Evol. 54:1407-1417.

Allen AM, Winfield MO, Burridge AJ, Downie RC, Benbow HR, Barker GL, Wilkinson PA, Coghill J, Waterfall C, Davassi A, Scopes G (2017). Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). Plant Biotechnol. J. 15(3):390-401.

Ayuk-Takem J (1984). Performance de 12 variétés de blé (*Triticum aestivum* L.) en croissance à UPPER FARM (1982 m). Rapport d'activités annuelles de IRA (Institut de la Recherche Agronomique) Bambui, pp.1982-1986.

Bohn M, Utz HF, Melchinger AE (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs and SSRs and their use for predicting progeny variance. Crop Sci. 39:228-237.

Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

Dice LR (1945). Measures of the amount of ecologic association between species. Ecology 26:297-302.

Doyle JD, JL Doyle (1990). Isolation of plant DNA from fresh tissue. BRL Focus 12:13-15.

Froese PS, Carter AH (2016). Single Nucleotide Polymorphisms in the Wheat Genome Associated with tolerance of Acidic Soils and Aluminum Toxicity. Crop Sci. 56:1662-1677.

Fu YB, Peterson GW, Richards KW, Somers D, DePauw RM, Clarke JM (2005). Allelic reduction and genetic shift in the Canadian hard red spring wheat germplasm released from 1845 to 2004. Theor. Appl. Genet. 110:1505-1516.

Gorji AH, Zolnoori M (2011). Genetic diversity in Hexaploid Wheat Genotypes using Microsatellite Markers. Asian J. Biotechnol. 3(4):368-377.

Grassini P, Eskridge KM, Cassman KG (2013). Distinguishing between yield advances and yield plateaus in historical crop production trends. Nat. Commun. 4:2918-000.

Harding RM, Boyce AJ, Clegg JB (1992). The evolution of tandemly repetitive DNA: recombination rules. Genetics 132:847-859.

Huang XQ, Börner A, Röder MS, Ganal MW (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor. Appl. Genet. 105:699-707.

Khaled FMS, Röder MS, Börner A (2015). Assessing genetic diversity of Egyptian hexaploid wheat (*Triticum aestivum* L.) using microsatellite markers. Genet. Resour. Crop Evol. 62:377-385.

Khlestkina EK, Röder MS, Efremova TT, Börner A, Shumny VK (2004). The genetic diversity of old and modern Siberian varieties of common spring wheat as determined by microsatellite markers. Plant Breed. 123:122-127.

Liu K, Muse SV (2005). PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics 21(9):2128-2129.

Mukhtar MS, Rehman M, Zafar Y (2002). Assessment of genetic diversity among wheat (*Triticum aestivum* L.) cultivars from a range of localities across Pakistan using random amplified polymorphic DNA (RAPD) analysis. Euphytica 28:417-425.

Nei M (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321-3323.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of endonucleases. Proc. Natl. Acad. Sci. USA 76:5269-5273.

Prasad M, Varshney RK, Roy JK, Balyan HS, Gupta PK (2000). The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theor. Appl. Genet. 100:584-592.

Rajaram S, Hettel GP (1994). Wheat Breeding at CIMMYT: Wheat special report N° 29. Ciudad Obregon, Sonora, Mexico 21-25 March, 1994.

Reza MD, Goodarz N, Elham Aram S (2015). Investigation of genetic diversity of some durum and bread wheat genotypes using SSR markers. J. Biodivers. Environ. Sci. 6(3):24-32.

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998). A microsatellite map of wheat. Genetics 149:2007-2023.

Röder MS, Wendehake K, Korzun V, Bredemeijer G, Laborie D, Bertrand L, Isaac P, Rendell S, Jackson J, Cooke RJ, Vosman B, Ganal MW (2002). Construction and analysis of a microsatellite-based database of European wheat varieties. Theor. Appl. Genet. 106:67-73.

Ryan PR, Raman H, Gupta S, Horst WJ, Delhaize E (2009). A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. Plant Physiol. 149 340-351.

Sneath PHA, Sokal RR (1973). Numerical Taxonomy. Freeman. San Francisco. P. 573.

Somers DJ, Isaac P (2004). SSRs from the Wheat Microsatellite Consortium. Available at< wheat. pw. usda. gov/ggpages/SSR/WMC> Accessed on March 10 (2004): 2010.

Sonmezoglu OA, Bozmaz B, Yildirim A, Kandemir N, Aydin N (2012). Genetic characterization of Turkish bread wheat landraces based on microsatellite markers and morphological characters. Turk. J. Biol. 36:589-597.

Tachida W, Iizuka M (1992). Persistence of repeated sequences that evolve by replication slippage. Genetics 131:471-478.