THE UNIVERSITY OF YAOUNDE I
****
FACULTY OF SCIENCE
****
CENTRE FOR RESEARCH AND TRAINING IN GRADUATE STUDIES IN LIFE, HEALTH AND ENVIRONMENTAL SCIENCES
****

UNIVERSITÉ DE YAOUNDÉ I
****
FACULTÉ DES SCIENCES
****
CENTRE DE RECHERCHE ET DE FORMATION DOCTORALE EN SCIENCES DE LA VIE, SANTE ET ENVIRONNEMENT
****

DEPARTMENT OF PLANT BIOLOGY
DÉPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VÉGÉTALES

Screening peanut (*Arachis hypogaea* L.) germplasm for quality traits and mapping QTL associated with yield-traits in an advanced backcross population

Thesis submitted in partial fulfilment for the requirements for award of a Doctorate (Ph.D.) Degree in Plant Biology
Option: Plant Biotechnology

BIBLIOTHEQUE UNIVERSITAIRE UY1

By:

Fentanesh Chekole KASSIE

MSc.in Plant Biotechnology
Registration number: 18W4585

Supervised by:

Hermine BILLE NGALLE
Associate Professor

and

Joseph Martin BELL
Professor

Year 2024

UNIVERSITÉ DE YAOUNDÉ I

THE UNIVERSITY OF YAOUNDE I

FACULTÉ DES SCIENCES

FACULTY OF SCIENCE

DÉPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VÉGÉTALES

DEPARTEMENT OF PLANT BIOLOGY

ATTESTATION DE CORRECTION

****

Nous soussignés, membres du Jury de soutenance de la thèse de **Doctorat/PhD** en **Biologie des Organismes Végétaux,** Option : **Biotechnologies Végétales,** soutenue le **02 Août 2024** par Madame **Fentanesh CHEKOLE KASSIE,** Msc. In Plant Biotechnology, Matricule : **18W4585,** intitulée « **Screening peanut (*Arachis hypogaea* L.) germplasm for quality traits and mapping QTL associated with yield-traits in an advanced backcross population** », certifions qu'elle a effectué les corrections conformément aux remarques et recommandations du Jury.

En foi de quoi, nous lui délivrons cette attestation de correction pour servir et valoir ce que de droit. /-

Rapporteurs

BELL Joseph Martin
Professeur

NGALLE BILLE Hermine
Maître de Conférences

Membres

NDONGO BEKOLO
Professeur

NYASSE Salomon
Directeur de Recherche

ANGONI Hyacinthe
Maître de Conférences

Président du Jury

AMBANG Zachée
Professeur

PROTOCOL LIST

UNIVERSITÉ DE YAOUNDÉ I

Faculté des Sciences

Division de la Programmation et du Suivi des Activités Académiques

THE UNIVERSITY OF YAOUNDE I

Faculty of Science

Division of Programming and Follow-up of Academic Affaires

LISTE DES ENSEIGNANTS PERMANENTS LIST OF PERMANENT TEACHING STAFF

ANNÉE ACADEMIQUE 2023/2024

(Par Département et par Grade)

DATE D'ACTUALISATION 04 Juin 2024

ADMINISTRATION

DOYEN : Owono Owono Luc Calvin, *Professeur*
VICE-DOYEN / DPSAA : ATCHADE Alex de Théodore, *Professeur*
VICE-DOYEN / DSSE : NYEGUE Maximilienne Ascension, *Professeur*
VICE-DOYEN / DRC : NOUNDJEU Pierre, *Maître de Conférences*
Chef Division Administrative et Financière : NDOYE FOE Florentine Marie Chantal, *Maître de Conférences*
Chef Division des Affaires Académiques, de la Recherche et de la Scolarité DAARS : AJEAGAH Gideon AGHAINDUM, *Professeur*

1- DÉPARTEMENT DE BIOCHIMIE (BC) (43)			
N°	**NOMS ET PRÉNOMS**	**GRADE**	**OBSERVATIONS**
1.	BIGOGA DAIGA Jude	Professeur	En poste
2.	FEKAM BOYOM Fabrice	Professeur	En poste
3.	KANSCI Germain	Professeur	En poste
4.	MBACHAM FON Wilfred	Professeur	En poste
5.	MOUNDIPA FEWOU Paul	Professeur	*Chef de Département*
6.	NGUEFACK Julienne	Professeur	En poste
7.	NJAYOU Frédéric Nico	Professeur	En poste
8.	OBEN Julius ENYONG	Professeur	En poste
9.	ACHU Merci BIH	Maître de Conférences	En poste
10.	ATOGHO Barbara MMA	Maître de Conférences	En poste
11.	AZANTSA KINGUE GABIN BORIS	Maître de Conférences	En poste
12.	BELINGA née NDOYE FOE F. M. C.	Maître de Conférences	*Chef DAF / FS*
13.	DAKOLE DABOY Charles	Maître de Conférences	En poste
14.	DONGMO LEKAGNE Joseph Blaise	Maître de Conférences	En poste
15.	DJUIDJE NGOUNOUE Marceline	Maître de Conférences	En poste
16.	DJUIKWO NKONGA Ruth Viviane	Maître de Conférences	En poste
17.	EFFA ONOMO Pierre	Maître de Conférences	*VD/FS/Univ Ebwa*

18.	EWANE Cécile Annie	Maître de Conférences	En poste
19.	KOTUE TAPTUE Charles	Maître de Conférences	En poste
20.	LUNGA Paul KEILAH	Maître de Conférences	En poste
21.	MANANGA Marlyse Joséphine	Maître de Conférences	En poste
22.	MBONG ANGIE M. Mary Anne	Maître de Conférences	En poste
23.	MOFOR née TEUGWA Clotilde	Maître de Conférences	*Doyen FS / UDs*
24.	NANA Louise épouse WAKAM	Maître de Conférences	En poste
25.	NGONDI Judith Laure	Maître de Conférences	En poste
26.	Palmer MASUMBE NETONGO	Maître de Conférences	En poste
27.	PECHANGOU NSANGOU Sylvain	Maître de Conférences	En poste
28.	TCHANA KOUATCHOUA Angèle	Maître de Conférences	En poste
29.	AKINDEH MBUH NJI	Chargé de Cours	En poste
30.	BEBEE Fadimatou	Chargée de Cours	En poste
31.	BEBOY EDJENGUELE Sara Nathalie	Chargé de Cours	En poste
32.	FONKOUA Martin	Chargé de Cours	En poste
33.	FOUPOUAPOUOGNIGNI Yacouba	Chargé de Cours	En poste
34.	KOUOH ELOMBO Ferdinand	Chargé de Cours	En poste
35.	MBOUCHE FANMOE Marceline Joëlle	Chargé de Cours	En poste
36.	OWONA AYISSI Vincent Brice	Chargé de Cours	En poste
37.	WILFRED ANGIE ABIA	Chargé de Cours	En poste
38.	BAKWO BASSOGOG Christian Bernard	Assistant	En Poste
39.	ELLA Fils Armand	Assistant	En Poste
40.	EYENGA Eliane Flore	Assistant	En Poste
41.	MADIESSE KEMGNE Eugenie Aimée	Assistant	En Poste
42.	MANJIA NJIKAM Jacqueline	Assistant	En Poste
43.	WOGUIA Alice Louise	Assistant	En Poste

2- DÉPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE ANIMALES (BPA) (49)			
1.	AJEAGAH Gideon AGHAINDUM	Professeur	*DAARS/FS*
2.	DIMO Théophile	Professeur	En Poste
3.	DJIETO LORDON Champlain	Professeur	En Poste
4.	DZEUFIET DJOMENI Paul Désiré	Professeur	En Poste
5.	ESSOMBA née NTSAMA MBALA	Professeur	*CD et Vice Doyen/FMSB/UYI*
6.	KEKEUNOU Sévilor	Professeur	En poste
7.	NJAMEN Dieudonné	Professeur	En poste
8.	NOLA Moïse	Professeur	En poste
9.	TAN Paul VERNYUY	Professeur	En poste
10.	TCHUEM TCHUENTE Louis Albert	Professeur	Inspecteur de service / Coord.Progr./MINSANTE
11.	ZEBAZE TOGOUET Serge Hubert	Professeur	En poste
12.	ALENE Désirée Chantal	Maître de Conférences	*Vice Doyen/ Uté Ebwa*
13.	ATSAMO Albert Donatien	Maître de Conférences	En poste
14.	BILANDA Danielle Claude	Maître de Conférences	En poste
15.	DJIOGUE Séfirin	Maître de Conférences	En poste
16.	GOUNOUE KAMKUMO Raceline épse FOTSING	Maître de Conférences	En poste
17.	JATSA BOUKENG Hermine épse MEGAPTCHE	Maître de Conférences	En Poste
18.	KANDEDA KAVAYE Antoine	Maître de Conférences	En Poste
19.	LEKEUFACK FOLEFACK Guy B.	Maître de Conférences	En poste
20.	MAHOB Raymond Joseph	Maître de Conférences	En poste
21.	MBENOUN MASSE Paul Serge	Maître de Conférences	En poste
22.	MEGNEKOU Rosette	Maître de Conférences	En poste

23.	MOUNGANG Luciane Marlyse	Maître de Conférences	En poste
24.	NOAH EWOTI Olive Vivien	Maître de Conférences	En poste
25.	MONY Ruth épse NTONE	Maître de Conférences	En Poste
26.	MVEYO NDANKEU Yves Patrick	Maître de Conférences	En Poste
27.	NGUEGUIM TSOFACK Florence	Maître de Conférences	En poste
28.	NGUEMBOCK	Maître de Conférences	En poste
29.	TAMSA ARFAO Antoine	Maître de Conférences	En poste
30.	TOMBI Jeannette	Maître de Conférences	En poste
31.	AMBADA NDZENGUE GEORGIA ELNA	Chargé de Cours	En poste
32.	BASSOCK BAYIHA Etienne Didier	Chargé de Cours	En poste
33.	ETEME ENAMA Serge	Chargé de Cours	En poste
34.	FEUGANG YOUMSSI François	Chargé de Cours	En poste
35.	FOKAM Alvine Christelle Epse KENGNE	Chargé de Cours	En poste
36.	GONWOUO NONO Legrand	Chargé de Cours	En poste
37.	KOGA MANG DOBARA	Chargé de Cours	En poste
38.	LEME BANOCK Lucie	Chargé de Cours	En poste
39.	MAPON NSANGOU Indou	Chargé de Cours	En poste
40.	METCHI DONFACK MIREILLE FLAURE EPSE GHOUMO	Chargé de Cours	En poste
41.	NGOUATEU KENFACK Omer Bébé	Chargé de Cours	En poste
42.	NJUA Clarisse YAFI	Chargée de Cours	*Chef Div. Uté Bamenda*
43.	NWANE Philippe Bienvenu	Chargé de Cours	En poste
44.	TADU Zephyrin	Chargé de Cours	En poste
45.	YEDE	Chargé de Cours	En poste
46.	YOUNOUSSA LAME	Chargé de Cours	En poste
47.	KODJOM WANCHE Jacguy Joyce	Assistante	En poste
48.	NDENGUE Jean De Matha	Assistant	En poste
49.	ZEMO GAMO Franklin	Assistant	En poste

3- DÉPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VÉGÉTALES (BPV) (32)			
1.	AMBANG Zachée	Professeur	*Chef de Département*
2.	DJOCGOUE Pierre François	Professeur	En poste
	Joseph Martin BELL	Professeur	
3.	MBOLO Marie	Professeur	En poste
4.	MOSSEBO Dominique Claude	Professeur	En poste
5.	NDONGO BEKOLO	Professeur	En poste
6.	ZAPFACK Louis	Professeur	En poste
7.	ANGONI Hyacinthe	Maître de Conférences	En poste
8.	BIYE Elvire Hortense	Maître de Conférences	En poste
9.	MAHBOU SOMO TOUKAM. Gabriel	Maître de Conférences	En poste
10.	MALA Armand William	Maître de Conférences	En poste
11.	MBARGA BINDZI Marie Alain	Maître de Conférences	*DAAC /UDla*
12.	NGALLE Hermine BILLE	Maître de Conférences	En poste
13.	NGONKEU MAGAPTCHE Eddy L.	Maître de Conférences	*CT / MINRESI*
14.	TONFACK Libert Brice	Maître de Conférences	En poste
15.	TSOATA Esaïe	Maître de Conférences	En poste
16.	ONANA JEAN MICHEL	Maître de Conférences	En poste

17.	DJEUANI Astride Carole	Chargé de Cours	En poste
18.	GONMADGE CHRISTELLE	Chargée de Cours	En poste
19.	MAFFO MAFFO Nicole Liliane	Chargé de Cours	En poste
20.	MANGA NDJAGA JUDE		En poste
21.	NNANGA MEBENGA Ruth Laure	Chargé de Cours	En poste
22.	NOUKEU KOUAKAM Armelle	Chargé de Cours	En poste
23.	NSOM ZAMBO EPSE PIAL ANNIE CLAUDE	Chargé de Cours	*En détachement/UNESCO MALI*
24.	GODSWILL NTSOMBOH NTSEFONG	Chargé de Cours	En poste
25.	KABELONG BANAHO Louis-Paul-Roger	Chargé de Cours	En poste
26.	KONO Léon Dieudonné	Chargé de Cours	En poste
27.	LIBALAH Moses BAKONCK	Chargé de Cours	En poste
28.	LIKENG-LI-NGUE Benoit C	Chargé de Cours	En poste
29.	TAEDOUNG Evariste Hermann	Chargé de Cours	En poste
30.	TEMEGNE NONO Carine	Chargé de Cours	En poste
31.	DIDA LONTSI Sylvere Landry	Assistant	En poste
32.	METSEBING Blondo-Pascal	Assistant	En poste

4- DÉPARTEMENT DE CHIMIE INORGANIQUE (CI) (27)			
1.	GHOGOMU Paul MINGO	Professeur	*Ministre Chargé de Mission PR*
2.	NANSEU NJIKI Charles Péguy	Professeur	En poste
3.	NDIFON Peter TEKE	Professeur	*CT MINRESI*
4.	NENWA Justin	Professeur	En poste
5.	NGOMO Horace MANGA	Professeur	*Vice Chancelor/UB*
6.	NJIOMOU C. épse DJANGANG	Professeur	En poste
7.	NJOYA Dayirou	Professeur	En poste
8.	ACAYANKA Elie	Maître de Conférences	En poste
9.	EMADAK Alphonse	Maître de Conférences	En poste
10.	KAMGANG YOUBI Georges	Maître de Conférences	En poste
11.	KEMMEGNE MBOUGUEM Jean C.	Maître de Conférences	En poste
12.	KENNE DEDZO GUSTAVE	Maître de Conférences	En poste
13.	MBEY Jean Aime	Maître de Conférences	En poste
14.	NDI NSAMI Julius	Maître de Conférences	*Chef de Département*
15.	NEBAH Née NDOSIRI Bridget NDOYE	Maître de Conférences	*Sénatrice/SENAT*
16.	NYAMEN Linda Dyorisse	Maître de Conférences	En poste
17.	PABOUDAM GBAMBIE AWAWOU	Maître de Conférences	En poste
18.	TCHAKOUTE KOUAMO Hervé	Maître de Conférences	En poste
19.	BELIBI BELIBI Placide Désiré	Maître de Conférences	*Chef Service/ ENS Bertoua*
20.	CHEUMANI YONA Arnaud M.	Maître de Conférences	En poste
21.	KOUOTOU DAOUDA	Maître de Conférences	En poste
22.	MAKON Thomas Beauregard	Chargé de Cours	En poste
23.	NCHIMI NONO KATIA	Chargée de Cours	En poste
24.	NJANKWA NJABONG N. Eric	Chargé de Cours	En poste
25.	PATOUOSSA ISSOFA	Chargé de Cours	En poste
26.	SIEWE Jean Mermoz	Chargé de Cours	En Poste
27.	BOYOM TATCHEMO Franck W.	Assistant	En Poste

5- DÉPARTEMENT DE CHIMIE ORGANIQUE (CO) (34)			
1.	Alex de Théodore ATCHADE	Professeur	Vice-Doyen / DPSAA
2.	DONGO Etienne	Professeur	Vice-Doyen/FSE/UYI
3.	NGOUELA Silvère Augustin	Professeur	Chef de Département UDS
4.	PEGNYEMB Dieudonné Emmanuel	Professeur	Directeur/ MINESUP/ Chef de Département
5.	MBAZOA née DJAMA Céline	Professeur	En poste
6.	MKOUNGA Pierre	Professeur	En poste
7.	TCHOUANKEU Jean-Claude	Professeur	En poste
8.	AMBASSA Pantaléon	Maître de Conférences	En poste
9.	EYONG Kenneth OBEN	Maître de Conférences	En poste
10.	FOTSO WABO Ghislain	Maître de Conférences	En poste
11.	KAMTO Eutrophe Le Doux	Maître de Conférences	En poste
12.	KENMOGNE Marguerite	Maître de Conférences	En poste
13.	KOUAM Jacques	Maître de Conférences	En poste
14.	MVOT AKAK CARINE	Maître de Conférences	En poste
15.	NGO MBING Joséphine	Maître de Conférences	Chef de Cellule MINRESI
16.	NGONO BIKOBO Dominique Serge	Maître de Conférences	C.E.A/ MINESUP
17.	NOTE LOUGBOT Olivier Placide	Maître de Conférences	DAAC/Uté Bertoua
18.	NOUNGOUE TCHAMO Diderot	Maître de Conférences	En poste
19.	TABOPDA KUATE Turibio	Maître de Conférences	En poste
20.	TAGATSING FOTSING Maurice	Maître de Conférences	En poste
21.	ZONDEGOUMBA Ernestine	Maître de Conférences	En poste
22.	MESSI Angélique Nicolas	Chargé de Cours	En poste
23.	MUNVERA MFIFEN Aristide	Chargé de Cours	En poste
24.	NGNINTEDO Dominique	Chargé de Cours	En poste
25.	NGOMO Orléans	Chargée de Cours	En poste
26.	NONO NONO Éric Carly	Chargé de Cours	En poste
27..	OUAHOUO WACHE Blandine M.	Chargée de Cours	En poste
28.	OUETE NANTCHOUANG Judith Laure	Chargée de Cours	En poste
29.	SIELINOU TEDJON Valérie	Chargé de Cours	En poste
30.	TCHAMGOUE Joseph	Chargé de Cours	En poste
31.	TSAFFACK Maurice	Chargé de Cours	En poste
32.	TSAMO TONTSA Armelle	Chargé de Cours	En poste
33.	TSEMEUGNE Joseph	Chargé de Cours	En poste
34.	NDOGO ETEME Olivier	Assistant	En poste

6. DÉPARTEMENT DES ENERGIES RENOVELABLES (ER) (1)			
1	BODO Bertrand	Professeur	*Chef de*

7- DÉPARTEMENT D'INFORMATIQUE(IN) (22)			
1.	ATSA ETOUNDI Roger	Professeur	*Chef de Division MINESUP*
2.	FOUDA NDJODO Marcel Laurent	Professeur	*Inspecteur Général/ MINESUP*

3.	NDOUNDAM Réné	Maître de Conférences	En poste
4.	TSOPZE Norbert	Maître de Conférences	En poste
5.	ABESSOLO ALO'O Gislain	Chargé de Cours	*Chef de Cellule MINFOPRA*
6.	AMINOU HALIDOU	Chargé de Cours	*Chef de Département*
7.	DJAM Xaviera YOUH - KIMBI	Chargé de Cours	En Poste
8.	DOMGA KOMGUEM Rodrigue	Chargé de Cours	En poste
9.	EBELE Serge Alain	Chargé de Cours	En poste
10.	HAMZA Adamou	Chargé de Cours	En poste
11.	JIOMEKONG AZANZI Fidel	Chargé de Cours	En poste
12.	KOUOKAM KOUOKAM E. A.	Chargé de Cours	En poste
13.	MELATAGIA YONTA Paulin	Chargé de Cours	En poste
14.	MESSI NGUELE Thomas	Chargé de Cours	En poste
15.	MONTHE DJIADEU Valery M.	Chargé de Cours	En poste
16.	NZEKON NZEKO'O ARMEL JACQUES	Chargé de Cours	En poste
17.	OLLE OLLE Daniel Claude Georges Delort	Chargé de Cours	*Sous-Directeur ENSET Ebolowa*
18.	TAPAMO Hyppolite	Chargé de Cours	En poste
19.	BAYEM Jacques Narcisse	Assistant	En poste
20.	EKODECK Stéphane Gaël Raymond	Assistant	En poste
21.	MAKEMBE. S . Oswald	Assistant	*Directeur CUTI*
22.	NKONDOCK. MI. BAHANACK.N.	Assistant	En poste

8- DÉPARTEMENT DE MATHÉMATIQUES (MA) (33)			
1.	AYISSI Raoult Domingo	Professeur	*Chef de Département*
2.	KIANPI Maurice	Maître de Conférences	En poste
3.	MBANG Joseph	Maître de Conférences	En poste
4.	MBEHOU Mohamed	Maître de Conférences	En poste
5.	MBELE BIDIMA Martin Ledoux	Maître de Conférences	En poste
6.	NOUNDJEU Pierre	Maître de Conférences	*Chef Service des Programmes & Diplômes/FS/UYI*
7.	TAKAM SOH Patrice	Maître de Conférences	En poste
8.	TCHAPNDA NJABO Sophonie B.	Maître de Conférences	*Directeur/AIMS Rwanda*
9.	TCHOUNDJA Edgar Landry	Maître de Conférences	En poste
10.	AGHOUKENG JIOFACK Jean Gérard	Chargé de Cours	*Chef Cellule MINEPAT*
11.	BOGSO ANTOINE Marie	Chargé de Cours	En poste
12.	BITYE MVONDO Esther Claudine	Chargé de Cours	En poste
13..	CHENDJOU Gilbert	Chargé de Cours	En poste
14.	DJIADEU NGAHA Michel	Chargé de Cours	En poste
15	DOUANLA YONTA Herman	Chargé de Cours	En poste
16.	KIKI Maxime Armand	Chargé de Cours	En poste
17.	LOUMNGAM KAMGA Victor	Chargé de Cours	En poste
18..	MBAKOP Guy Merlin	Chargé de Cours	En poste
19	MBATAKOU Salomon Joseph	Chargé de Cours	En poste
20..	MENGUE MENGUE David Joël	Chargé de Cours	*Chef Dpt /ENS Université d'Ebolowa*

21	MBIAKOP Hilaire George	Chargé de Cours	En poste
22.	NGUEFACK Bernard	Chargé de Cours	En poste
23.	NIMPA PEFOUKEU Romain	Chargée de Cours	En poste
24.	OGADOA AMASSAYOGA	Chargée de Cours	En poste
25.	POLA DOUNDOU Emmanuel	Chargé de Cours	*En stage*
26.	TENKEU JEUFACK Yannick Léa	Chargé de Cours	*En stage*
27.	TCHEUTIA Daniel Duviol	Chargé de Cours	En poste
28.	TETSADJIO TCHILEPECK M. Eric.	Chargé de Cours	En poste
29.	FOKAM Jean Marcel	Assistant	En poste
30.	GUIDZAVAI KOUCHERE Albert	Assistant	En poste
31.	MANN MANYOMBE Martin Luther	Assistant	En poste
32.	MEFENZA NOUNTU Thiery	Assistant	En poste
33.	NYOUMBI DLEUNA Christelle	Assistant	En poste

9- DÉPARTEMENT DE MICROBIOLOGIE (MIB) (24)			
1.	ESSIA NGANG Jean Justin	Professeur	*Chef de Département*
2.	NYEGUE Maximilienne Ascension	Professeur	*VICE-DOYEN / DSSE*
3	SADO KAMDEM Sylvain Leroy	Professeur	En poste
4.	ASSAM ASSAM Jean Paul	Maître de Conférences	En poste
5.	BOUGNOM Blaise Pascal	Maître de Conférences	En poste
6.	BOYOMO ONANA	Maître de Conférences	En poste
7	KOUITCHEU MABEKU Epse KOUAM Laure Brigitte	Maître de Conférences	En poste
8..	RIWOM Sara Honorine	Maître de Conférences	En poste
9	NJIKI BIKOÏ Jacky	Maître de Conférences	En poste
10	TCHIKOUA Roger	Maître de Conférences	En poste
11.	ESSONO Damien Marie	Chargé de Cours	En poste
12.	LAMYE Glory MOH	Chargé de Cours	En poste
13.	MEYIN A EBONG Solange	Chargé de Cours	En poste
14.	MONI NDEDI Esther Del Florence	Chargé de Cours	En poste
15.	NKOUDOU ZE Nardis	Chargé de Cours	En poste
16.	NKOUE TONG Abraham	Chargé de Cours	En poste
17	TAMATCHO KWEYANG Blandine Pulchérie	Chargé de Cours	En poste
18	TOBOLBAÏ Richard	Chargé de Cours	En poste
19.	SAKE NGANE Carole Stéphanie	Chargé de Cours	En poste
20.	EZO'O MENGO Fabrice Télésfor	Assistant	En poste
21.	EHETH Jean Samuel	Assistant	En poste
22.	MAYI Marie Paule Audrey	Assistant	En poste
23.	NGOUENAM Romial Joël	Assistant	En poste
24.	NJAPNDOUNKE Bilkissou	Assistant	En poste

10. DEPARTEMENT DE PYSIQUE(PHY) (42)			
1.	BEN- BOLIE Germain Hubert	Professeur	En poste
2	BIYA MOTTO Frédéric	Professeur	En poste
3.	DJUIDJE KENMOE épouse ALOYEM	Professeur	En poste
4.	EKOBENA FOUDA Henri Paul	Professeur	*Vice-Recteur. Uté Ngaoundéré*
5.	ESSIMBI ZOBO Bernard	Professeur	En poste
6.	EYEBE FOUDA Jean sire	Professeur	En poste
7.	HONA Jacques	Professeur	En poste

8.	NANA ENGO Serge Guy	Professeur	En poste
9.	NANA NBENDJO Blaise	Professeur	En poste
10.	NDJAKA Jean Marie Bienvenu	Professeur	*Chef de Département*
11.	NJANDJOCK NOUCK Philippe	Professeur	En poste
12.	NOUAYOU Robert	Professeur	En poste
13.	SAIDOU	Professeur	*Chef de centre/IRGM/MINRESI*
14.	SIMO Elie	Professeur	En poste
15.	TABOD Charles TABOD	Professeur	*Doyen FSUniv/Bda*
16.	TCHAWOUA Clément	Professeur	En poste
17.	WOAFO Paul	Professeur	En poste
18.	ZEKENG Serge Sylvain	Professeur	En poste
19.	ENYEGUE A NYAM épse BELINGA	Maître de Conférences	En poste
20.	FEWO Serge Ibraïd	Maître de Conférences	En poste
21.	FOUEJIO David	Maître de Conférences	*Chef Cell. MINADER*
22.	MBINACK Clément	Maître de Conférences	En poste
23.	MBONO SAMBA Yves Christian U.	Maître de Conférences	En poste
24.	MELI'I Joelle Larissa	Maître de Conférences	En poste
25.	MVOGO ALAIN	Maître de Conférences	En poste
26.	NDOP Joseph	Maître de Conférences	En poste
27.	SIEWE SIEWE Martin	Maître de Conférences	En poste
28.	VONDOU Derbetini Appolinaire	Maître de Conférences	En poste
29.	WAKATA née BEYA Annie Sylvie	Maître de Conférences	*Directeur/ENS/UYI*
30.	WOULACHE Rosalie Laure	Maître de Conférence	*En stage depuis février 2023*
31.	ABDOURAHIMI	Chargé de Cours	En poste
32.	AYISSI EYEBE Guy François Valérie	Chargé de Cours	En poste
33.	CHAMANI Roméo	Chargé de Cours	En poste
34.	DJIOTANG TCHOTCHOU Lucie Angennes	Chargée de Cours	En poste
35.	EDONGUE HERVAIS	Chargé de Cours	En poste
36.	KAMENI NEMATCHOUA Modeste	Chargé de Cours	En poste
37.	LAMARA Maurice	Chargé de Cours	En poste
38.	NGA ONGODO Dieudonné	Chargé de Cours	En poste
39.	OTTOU ABE Martin Thierry	Chargé de Cours	Directeur Unité de production des réactifs/IMPM
40.	TEYOU NGOUPO Ariel	Chargé de Cours	En poste
41.	WANDJI NYAMSI William	Chargé de Cours	En poste
42.	SOUFFO TAGUEU Merimé	Assistant	En poste

11- DÉPARTEMENT DE SCIENCES DE LA TERRE (ST) (42)			
1.	BITOM Dieudonné-Lucien	Professeur	*Doyen / FASA /UDs*
2.	EKOMANE Emile	Professeur	*Chef Div./Uté Ebolowa*
3.	GANNO Sylvestre	Professeur	En poste
4.	NDAM NGOUPAYOU Jules-Remy	Professeur	En poste
5.	NDJIGUI Paul-Désiré	Professeur	*Chef de Département*

6.	NGOS III Simon	Professeur	En poste
7.	NKOUMBOU Charles	Professeur	En poste
8.	NZENTI Jean-Paul	Professeur	En poste
9.	ONANA Vincent Laurent	Professeur	*Chef de Département/Uté. Eb.*
10.	YENE ATANGANA Joseph Q.	Professeur	*Chef Div. /MINTP*
11.	BISSO Dieudonné	Maître de Conférences	En poste
12.	Elisé SABABA	Maitre de Conférences	En poste
13.	EYONG John TAKEM	Maitre de Conférences	En poste
14.	FUH Calistus Gentry	Maître de Conférences	*Sec. d'Etat/MINMIDT*
15.	GHOGOMU Richard TANWI	Maître de Conférences	*Chef de Div. /Uté Bertoua*
16.	MBIDA YEM	Maitre de Conférences	En poste
17.	MBESSE Cécile Olive	Maitre de Conférences	En poste
18.	METANG Victor	Maitre de Conférences	En poste
19.	MOUNDI Amidou	Maître de Conférences	*CT/MINIMDT*
20.	NGO BIDJECK Louise Marie	Maître de Conférences	En poste
21.	NGUEUTCHOUA Gabriel	Maître de Conférences	*CEA/MINRESI*
22.	NJILAH Isaac KONFOR	Maître de Conférences	En poste
23.	NYECK Bruno	Maître de Conférences	En poste
24.	TCHAKOUNTE Jacqueline épse NUMBEM	Maître de Conférences	*Chef. Cell /MINRESI*
25.	TCHOUANKOUE Jean-Pierre	Maître de Conférences	En poste
26.	TEMGA Jean Pierre	Maître de Conférences	En poste
27.	ZO'O ZAME Philémon	Maître de Conférences	*DG/ART*
28.	ANABA ONANA Achille Basile	Chargé de Cours	En poste
29.	BEKOA Etienne	Chargé de Cours	En poste
30.	ESSONO Jean	Chargé de Cours	En poste
31.	MAMDEM TAMTO Lionelle Estelle, épouse BITOM	Chargée de Cours	En poste
32.	MINYEM Dieudonné	Chargé de Cours	*Chef Serv./Uté Maroua*
33.	NGO BELNOUN Rose Noël	Chargée de Cours	En poste
34.	NGO'O ZE ARNAUD	Chargée de Cours	En poste
35.	NOMO NEGUE Emmanuel	Chargé de Cours	En poste
36.	NTSAMA ATANGANA Jacqueline	Chargée de Cours	En poste
37.	TCHAPTCHET TCHATO De P.	Chargé de Cours	En poste
38.	TEHNA Nathanaël	Chargé de Cours	En poste
39.	FEUMBA Roger	Chargé de Cours	En poste
40.	MBANGA NYOBE Jules	Chargé de Cours	En poste
41.	KOAH NA LEBOGO Serge Parfait	Assistant	En poste
42.	TENE DJOUKAM Joëlle Flore, épouse KOUANKAP NONO	Assistante	En poste

Répartition chiffrée des Enseignants de la Faculté des Sciences de l'Université de Yaoundé I

NOMBRE D'ENSEIGNANTS					
DÉPARTEMENT	Professeurs	Maîtres de Conférences	Chargés de Cours	Assistants	Total
BCH	8 (01)	20 (12)	9 (04)	6 (05)	43 (22)
BPA	11 (01)	19 (09)	16 (05)	3 (02)	49 (17)
BPV	6 (01)	10 (02)	14 (08)	2 (00)	32 (11)
CI	7 (01)	14 (04)	5 (01)	1 (00)	27 (06)
CO	6 (01)	15 (05)	11 (05)	1 (00)	34 (11)
ER	1 (00)		/	/	1 (0)
IN	2 (00)	2 (00)	14 (01)	4 (00)	22 (01)
MAT	2 (00)	8 (00)	19 (02)	5 (01)	33 (03)
MIB	3 (01)	7 (03)	9 (05)	5 (02)	24 (11)
PHY	18 (01)	12 (04)	11 (01)	1 (00)	42 (06)
ST	10 (00)	17 (03)	13 (03)	3 (01)	43 (07)
Total	74 (07)	124 (42)	121 (35)	31 (11)	350(95)

Soit un total de----------350 (95) don't:
Professeurs ----------74(07)
Maîtres de Conférences----------124 (42)
Chargés de Cours----------121 (35)
Assistants----------31 (11)

( ) = Nombre de Femmes **95**

-

DEDICATION

This work is dedicated to my lovely kids, Meklit Abreham TADDELE and Bezawit Abreham TADDELE. Your unwavering love and support have been my guiding light throughout my journey, especially during the years I spent away pursuing my Ph.D. I deeply acknowledge that the time apart meant moments lost, moments I wish I could have shared with you, moments I could have shown my love by simply being there. This dedication is a small token of my immense gratitude for your patience, understanding, and encouragement.

ACKNOWLEDGEMENTS

I am deeply grateful to the institutions and individuals who have played pivotal roles in supporting me throughout my academic journey. Firstly, I offer my profound thanks to the Almighty God for His grace and blessings, which have made all of this possible. I extend my sincere appreciation to the European Union for granting me the esteemed opportunity through the GENES intra-Africa Academic Mobility scheme and to the University of Yaoundé I, Cameroon, for embracing me as a student. I am also indebted to the Department of Plant Science, College of Agriculture, Wolaita Sodo University, Ethiopia, for their unwavering support and understanding.

Special gratitude is reserved for Professor AMBANG Zachee, the head of the Department of Plant Biology, whose encouragement, advice, and reassurance have been a constant source of inspiration. I am deeply thankful to my main supervisor, Professor Joseph Martin BELL, for his invaluable guidance, creative insights, and profound kindness throughout my Ph.D. journey.

Likewise, I extend my profound appreciation to my advisor and host institution coordinator for the GENES project, Associate Professor Hermine Bille NGALLE, whose expertise, insightful comments, and high standards have significantly contributed to the development of my research and writing skills. I am particularly grateful to Dr. Joël Romaric NGUEPJOP for his academic guidance and valuable contributions to the development of plant materials. My sincere thanks also go to my home advisor, Dr. Mesfin Kebede GESSE, for his unwavering support and guidance at every stage of my doctoral research.

I want to thank Dr. Wosene Gebreselassie ABTEW, my in-country coordinator for the GENES Project, for his invaluable guidance. I am also thankful to the Genetics and Plant Breeding Unit (UGAP) of the Department of Plant Biology at the University of Yaoundé I for their support and availability. My appreciation extends to Doctors Achile NYOUMA, Luther MBO NKOULOU, Benoit LIKENG, and Godswill NTSOMBOH NTSEFONG for their support and guidance.

I am deeply appreciative of my classmates for their unwavering support and availability, with special mention to our class representative, Charly TCHAPDA, for his consistent assistance and encouragement. I would also like to express my heartfelt gratitude to my beloved children, Meklit Abrham TADDELE and Bezawit Abrham TADDELE, and my husband, Dr. Abrham Taddele TEREFE, for their unconditional love and patience throughout this journey.

I am also grateful to my sister, Yeshareg Chekole KASSIE, for her support in caring for my children during my study period. Warmest thanks to my sisters, brothers, and mother for their unwavering support and encouragement in my academic pursuits. Additionally, I appreciate my family on my husband's side for their encouragement.

I extend my gratitude to all the members of the jury who will dedicate their time to read and provide feedback to enhance the quality of this work. Finally, my sincere thanks to anyone who contributed to the success of my Ph.D. journey, even if their names are not mentioned here.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBRIVIATIONS

AB	:	Advanced Backcross
AB-QTL	:	Advanced backcross quantitative trait loci
AFLP	:	Amplified Fragment Length Polymorphism
AiAd	:	*Arachis ipaensis* x *Arachis duranensis*
ANOVA	:	Analysis of Variance
BeCA	:	Bioscience Eastern and Central Africa
CSSL	:	Chromosome Segment Substitution Line
CIM	:	Composite Interval Mapping
CTAB	:	Cetyltrimethylammonium Ammonium Bromide
DArT	:	Diversity Array Technology
DNA	:	Deoxyribonucleic Acid
FN	:	False Negative
FP	:	False Positive
ICIM	:	Inclusive Composite Interval Mapping
ICIMADD	:	Inclusive Composite Interval Mapping Additive
ICRISAT	:	International Crops Research Institute for the Semi-Arid Tropics
IM	:	Intervale Mapping
LG	:	Linkage Group
MABC	:	Marker Assisted Backcrossing
Mha	:	Million hectares
MARS	:	Marker Assisted Recurrent Selection

MAS : Markers Assisted-Selection

MAQP : Marker Assisted Quantitative trait loci Pyramiding

MAGIC : Multiple Advanced Generation Intercross

MIM : Multiple Interval Mapping

MT : Million Tons

NBPGR : National Bureau of Plant Genetic Resources

NCSU : North Carolina State University

NIAT : National Institute of Agricultural Technology

NIL : Near Isogenic Line

NIRs : Near Infrared Reflectance Spectroscopy

PCA : Principal Component Analysis

PCR : Polymerase Chain Reaction

PGRCU : Plant Genetic Resource Conservation Unit

PLS-DA : Partial Least Squares-Discriminant Analysis

PVE : Phenotypic Variance Explained

QTL : Quantitative Trait Loci

RFLP : Restriction Fragment Length Polymorphism

RIL : Recombinant Inbreed Line

SMA : Single Marker Analysis

SSR : Simple Sequence Repeat

SNP : Single Nucleotide Polymorphism

LIST OF APPENDICES

ABSTRACT

Peanut (*Arachis hypogea* L.) is globally known for its nutritional richness and culinary versatility. However, limited genetic diversity and underutilization of genetic resources are significant hurdles to peanut breeding progress. Interspecific hybridization and genetic diversity assessments using suitable characterization methods offer promising solutions to these challenges. This study aimed to explore both rapid germplasm screening for quality traits and identify favourable wild QTL linked to yield traits. Near-infrared spectroscopy (NIRS), coupled with chemometrics, was employed to assess germplasm variability in 680 samples. This included a core collection of 300 varieties and three sets of 133 (Fleur11 x ISATGR 278-18) genotypes from an interspecific population. Evaluation was conducted in Mbalmayo and Bafia, Cameroon, and Nioro, Senegal. NIR elemental spectra were gathered on six subsets of seeds in each sample after three rotation scans, with a spectral resolution of 16 cm-1 over the range of 867-2530 nm. Spectra were processed using principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA). As results, a huge variability was found between varieties and genotypes within and between environments at multiple wavelengths, particularly at 1723 nm. associated with oil content and fatty acid composition. PCA revealed substantial chemical diversity, clustering varieties and genotypes into four groups corresponding to sample sets. The core-collection displayed the highest genetic variation compared to interspecific genotypes within the environment. Environmental factors significantly impacted seed composition, with Bafia showing the greatest variation, followed by Mbalmayo and Nioro, using the same interspecific population. A PLS-DA model achieved 99.6% accuracy in classifying seed samples by environmental origin. Further exploration assessed phenotypic variability, estimated broad-sense heritability, and evaluated trait correlations for yield-related traits. Data collected in Cameroon at Marou, Mbalmayo, and Bafia utilised 133 interspecific genotypes along with a recurrent parent. Observations showed morphological variability in qualitative traits, including plant growth habit (semi-erect to prostrate to fully erect) and pod constriction and beak (ranging from slight to prominent and slight to deep, respectively). Conversely, based on mean values, moderate to high levels of phenotypic variation were detected for quantitative traits across different environments. Notably, genotype 11_28_10 consistently demonstrated superior performance and pooled data, followed by 11_28_20 for 100-pod and seed weight, pod and seed length, and width. Analysis of variance (ANOVA) confirms significant genotype variability among interspecific genotypes for all traits.

Broad-sense heritability estimates ranged from moderate to high, suggesting a strong genetic influence on the traits studied. Association analysis reveals several positive and significant trait correlations, highlighting potential avenues for trait improvement. Exploration into wild genomic regions (QTLs) associated with yield-related traits revealed insights. Utilising 133 BC2F4 lines, a genetic map comprising 1,450 loci across 20 linkage groups is constructed, spanning a total length of 1,358.02 cM, with an average distance of 2.21 cM between flanking markers. A total of 44 putative QTLs were detected on 17 linkage groups for 14 yield traits. Among these QTLs, four were newly identified loci that had not been previously mapped. Notably, 20 of the putative QTLs (45%) were associated with an increase in the phenotypic value of the trait and were linked to alleles from the wild relative. Thirteen out of the 44 QTLs were classified as major QTLs (>10% phenotypic variance explained), indicating their potential significance for marker-assisted selection (MAS) pending confirmation across diverse environments.

Keywords: DArT, intact-seed, NIRS, nutritional, QTL identification, Peanut, PLS-DA

RÉSUMÉ

La cacahuète (*Arachis hypogea* L.) est mondialement connue pour sa richesse nutritionnelle et sa polyvalence culinaire. Cependant, la diversité génétique limitée et la sous-utilisation des ressources génétiques constituent des obstacles significatifs au progrès de l'amélioration de la cacahuète. L'hybridation interspécifique et les évaluations de la diversité génétique à l'aide de méthodes de caractérisation adaptées offrent des solutions prometteuses à ces défis. Cette étude visait à explorer à la fois le criblage rapide du germoplasme pour les traits de qualité et à identifier les QTL sauvages favorables liés aux traits de rendement. La spectroscopie proche infrarouge (NIRS), associée à la chimiométrie, a été utilisée pour évaluer la variabilité du germoplasme dans 680 échantillons. Cela comprenait une collection de base de 300 variétés et trois ensembles de 133 génotypes (Fleur11 x ISATGR 278-18) d'une population interspécifique. L'évaluation a été réalisée à Mbalmayo et Bafia, au Cameroun, et à Nioro, au Sénégal. Les spectres élémentaires proches infrarouges ont été recueillis sur six sous-ensembles de graines dans chaque échantillon après trois rotations, avec une résolution spectrale de 16 cm-1 sur la plage de 867 à 2530 nm. Les spectres ont été traités à l'aide de l'analyse en composantes principales (PCA) et de l'analyse discriminante en moindres carrés partiels (PLS-DA). En résultat, une énorme variabilité a été trouvée entre les variétés et les génotypes à l'intérieur et entre les environnements à de multiples longueurs d'onde, particulièrement à 1723 nm, associée à la teneur en huile et à la composition en acides gras. L'ACP a révélé une diversité chimique substantielle, regroupant les variétés et les génotypes en quatre groupes correspondant aux ensembles d'échantillons. La collection de base a affiché la plus grande variation génétique par rapport aux génotypes interspécifiques dans l'environnement. Les facteurs environnementaux ont impacté significativement la composition des graines, Bafia montrant la plus grande variation, suivie de Mbalmayo et Nioro, en utilisant la même population interspécifique. Un modèle PLS-DA a atteint une précision de 99,6% dans la classification des échantillons de graines par origine environnementale. Une exploration supplémentaire a évalué la variabilité phénotypique, estimé l'héritabilité au sens large et évalué les corrélations entre les traits liés au rendement. Les données collectées au Cameroun à Marou, Mbalmayo et Bafia ont utilisé 133 génotypes interspécifiques ainsi qu'un parent récurrent. Les observations ont montré une variabilité morphologique dans les traits qualitatifs, notamment l'habitude de croissance des plantes (semi-érigée à prostrée à complètement érigée) et la constriction et le bec des gousses (allant de légère à proéminente et légère à profonde, respectivement).

En revanche, sur la base des valeurs moyennes, des niveaux de variation phénotypique modérés à élevés ont été détectés pour les traits quantitatifs dans différents environnements. Notamment, le génotype 11_28_10 a régulièrement démontré une performance supérieure et des données regroupées, suivi par 11_28_20 pour le poids de 100 gousses et de graines, la longueur et la largeur des gousses et des graines. L'analyse de variance (ANOVA) confirme une variabilité significative des génotypes parmi les génotypes interspécifiques pour tous les traits. Les estimations d'héritabilité au sens large ont varié de modérées à élevées, suggérant une forte influence génétique sur les traits étudiés. L'analyse d'association révèle plusieurs corrélations positives et significatives entre les traits, mettant en évidence des pistes potentielles pour l'amélioration des traits. L'exploration des régions génomiques sauvages (QTL) associées aux traits liés au rendement a révélé des insights. En utilisant 133 lignées BC2F4, une carte génétique comprenant 1 450 locus répartis sur 20 groupes de liaison est construite, couvrant une longueur totale de 1 358,02 cM, avec une distance moyenne de 2,21 cM entre les marqueurs de flanquement. Au total, 44 QTL putatifs ont été détectés sur 17 groupes de liaison pour 14 traits de rendement. Parmi ces QTL, quatre étaient des loci nouvellement identifiés qui n'avaient pas été précédemment cartographiés. Notamment, 20 des QTL putatifs (45%) étaient associés à une augmentation de la valeur phénotypique du trait et étaient liés à des allèles du parent sauvage. Treize des 44 QTL ont été classés comme des QTL majeurs (>10% de variance phénotypique expliquée), indiquant leur importance potentielle pour la sélection assistée par marqueurs (SAM) en attendant confirmation dans des environnements divers.

Mots-clés : Arachide, sélection, graines intactes, NIRS, nutrition, DArT, identification QTL.

INTRODUCTION

Peanut or groundnut (*Arachis hypogaea* L.) is an oilseed legume crop belonging to the *Leguminosae* family and *Arachis* genus (Krapovickas and Gregory, 1994). The genus contains about 81 species, mostly diploids (2n = 2x = 20), taxonomically divided into nine sections (Krapovickas and Gregory, 1994; Valls and Simpson, 2005). Peanut is allotetraploid (2n = 4x = 40) with an AABB genome, originating from a single hybridization event followed by chromosome doubling between *A. duranensis* (A-genome) and *A. ipaensis* (B-genome), about 3,500 years ago (Kochert *et al*., 1996; Seijo *et al*., 2004; Moretzsohn *et al*., 2013; Lu *et al*., 2018). Its genome size is approximately 2.7 billion base pairs (2.7 Gb) (Bertioli *et al*., 2019; Zhuang *et al*., 2019). This large genome size contributes to the complexity of the plant and the challenges faced in breeding and genetic engineering efforts.

The crop, originating in South America, has spread worldwide and is cultivated in over 100 countries across tropical and subtropical regions (Krapovickas and Gregory, 1994; Burow *et al*., 2009; Bertioli *et al*., 2011). It covers about 31 million hectares (Mha) area in 2022, with a global production of approximately 54 million tons (MT) and an average yield of about 1.8 tons per hectare (t/ha) (FAOSTAT, 2024). The major producing regions are Asia 32 MT and Africa 17 MT which together account for 29 Mha representing 95 % of the global peanut cultivated areas in the world. Together, Asia (58%) and Africa (32%) accounted for about 90% of the world's production, with China 18 MT, India 10 MT and Nigeria 4MT being the top three largest producing countries in 2022 (FAOSTAT, 2024). Peanut productivity significantly varies among regions, with Africa having the lowest mean yield of around 1 t/ha compared to Asia (2.6 t/ha) and America (3.7 t/ha) (FAOSTAT, 2024). The low peanut yields observed in many countries are related to rainfed and low-input growing conditions.

Peanut is a significant vegetable oilseed crop, with over 60 % of global peanut production being crushed for the extraction of oil for both edible and industrial uses (Janila *et al*., 2013, 2016). It was ranked fourth in oilseed production, following soybean, rapeseed, and sunflower seed. (Statista, 2023). The total vegetable oil production in this period was approximately 217 million tons (MT), with peanut ranking sixth at 6.5 MT. Palm oil held the top position at 76 MT, followed by soybean (60 MT), rapeseed (31 MT), sunflower seed (18.6 MT), and palm kernel (8.8 MT).

Nutritionally, peanut offer high-quality edible oil, protein, carbohydrates, vitamins, and minerals. The nutritional composition of peanuts varies, with the edible oil content ranging from (34-56%), protein (22-30%) and carbohydrates (10-25%) (Nigam, 2014; Nawade *et al*., 2018; Desmae *et al*., 2019). It has previously been observed that the high oleic acid content of peanut oil offers health benefits, including cholesterol reduction and combating inflammatory diseases (Pandey *et al.*, 2014b; Bonku and Yu, 2020). The other research suggests that peanut consumption may lower the risk of type 2 diabetes in women (Jiang, 2002). Additionally, peanut is economically important as a cash crop and plays a vital role in agriculture, aiding in soil fertility through nitrogen fixation in crop rotation systems (Janila *et al.*, 2013, 2016).

The demand-supply gap for food grains is continuously increasing due to the ever-growing global population, which is projected to expand to 9.6 billion by 2050 (Rajwade *et al*., 2015; Alhashim and Anandhi, 2022). This escalating population presents a serious challenge, as current trends in yield increases may not be sufficient to cope with the growing demand. It has been projected that global food production needs to increase by over 70% by 2050 in order to meet the anticipated demands. However, the productivity of peanut has not been able to be sufficiently enhanced due to various production constraints affecting the crop. These constraints include drought, pests, diseases, and environmental changes. Furthermore,the oil content of seeds, shelf life, aroma, flavour and cooking quality are all affected by these constraints (Bakal and Arioglu, 2019; Bakal, 2020; Parilli-Moser *et al*., 2022). In response to this pressing issue, peanut breeders have to enhance peanut productivity as well as seed composition-related traits such as oil and protein content and fatty acid composition to meet the projected demands.

The cultivated peanut has low genetic diversity due to its origin and reproductive isolation from its wild diploid relatives owing to ploidy differences (Kochert *et al.*, 1996; Bertioli *et al*., 2011; Pandey *et al*., 2014a; Stalker, 2017).This narrow genetic base, coupled with low utilization of genetic resources, has been a major limiting factor for peanut breeding globally. To address narrow genetic diversity, the strategy of interspecific hybridization has been adopted to introduce diverse genetic traits from wild species into the cultivated gene pool (Fonceka *et al*., 2009; Kumari *et a*l., 2014; Nguepjop *et al*., 2016). In addition, to resolve the issues of ploidy differences, synthetic compatible tetraploids have been developed (Simpson *et al*., 1993; Favero *et al*., 2006; Mallikarjuna et al., 2010) from diploid wild relatives that exhibited a high level of genetic variation (Barkley *et al*., 2007; Bechara *et al*., 2010) with potential trait variation that may be useful in peanut breeding through interspecific cross.

Additionally, various genetic approaches, such as Advance Backcross Quantitative Trait Loci (AB-QTL) and Chromosome Segment Substitution Lines (CSSL), are currently being applied to broaden the genetic base of cultivated peanut by introgressing chromatin from wild relatives into elite lines (Fonceka *et al*., 2012b; Tossim *et al*., 2020). Such wild polymorphisms have the potential to further improve peanut quantitative traits.

Likewise, genetic diversity assessment and the detection of promising genotypes are fundamental to germplasm utilization and management in breeding strategies to support food security. To facilitate the investigation of large germplasm, it is reasonable to begin by examining subsets of germplasm that embody appropriate diversity and of manageable size, such as core collections (Brown, 1989) or interspecific population (Gimode *et al*., 2020) derived from wild x elite crosses, using appropriate characterization procedures.

The improvement of quantitative traits largely depends on the magnitude of genetic variability and the extent to which its determining traits are heritable. Although significant efforts have been devoted to characterizing cultivars and germplasm collections for disease resistance and agronomic traits and for the most important agronomic traits (Upadhyaya, 2005; Upadhyaya *et al*., 2006, 2011; Mallikarjuna *et al*., 2012; Kumari *et al.*, 2014), less is known about quality traits (Grosso *et al*., 2000; Bianchi-Hall *et al*., 1993). This is primarily due to the phenotyping of these traits by chemical studies, which is expensive in terms of both direct monetary input and human labour, time - consuming and destructive (Nawade *et al*., 2018; Davis *et al*., 2021).

Efforts to improve the knowledge of seed attributes might be supported by rapid and non-destructive tools. These include modified refractive index, capacitance sensor (Kandala *et al*., 2008), hyperspectral imaging (Huang *et al*., 2014; Rabanera *et al*., 2021), and near infrared (NIR) (Govindarajan *et al*., 2009; Tao *et al*., 2019; Davis *et al*., 2021; Wang *et al.*, 2022; Panero *et al*., 2018, 2022). Among these, NIR-based methods are rapid, making it possible to analyse large number of samples. Some works previously described the feasibility of near infrared spectrometers to achieve some quick prediction of various peanut chemical compounds (Li *et al*., 2019 ; Yu *et al.*, 2020 ; Bilal *et al*., 2020; Liu *et al*., 2022).

Moreover**,** some scholars have already applied machine learning as promising statistical methods to assist humans in the modelling and analysis of complex spectral data (Song *et al*., 2018; Fordellone *et al*., 2020) in many research fields, including seed quality detection, genotyping of cultivars (Panero et al., 2018, 2022), varieties identification (Wang and Song,

2023; Xu *et al*., 2023). and classification (Sampaio *et al.*, 2021; Singh *et al*., 2023; Tian *et al*., 2023).

Apart from developing and characterizing interspecific genotypes, the identification of Quantitative Trait Loci (QTL) and valuable wild QTL alleles is vital for enhancing peanut yield traits, with synthetic tetraploids serving as a reservoir of beneficial alleles. This is particularly crucial, as direct phenotypic selection in plant breeding is labour-intensive, costly, and time-consuming (Watson *et al*., 2019). To overcome the constraints of phenotypic selection, several studies have used marker-assisted selection (MAS) in peanut as a potential tool to achieve desirable results in crops with the help of molecular markers (Chu *et al*., 2011; Huang *et al*., 2019; Nawade *et al*., 2019; Shasidhar *et al*., 2020).

Despite the limited genetic diversity within cultivated peanut, several studies have reported QTLs for yield component traits (Chen *et al*., 2016, 2017; Luo *et al.*, 2017, 2018; Liang *et al*., 2018), and seed quality traits (Shasidhar *et al.*, 2017; Liu *et al*., 2020; Sun *et al*., 2021; Guo *et al*., 2021). However, developing populations utilising wild relatives is crucial to broadening the genetic background of elite varieties.

To date, only six synthetic tetraploid peanuts have been used to develop mapping populations to dissect the genomic segments of wild relatives of peanut for various important quantitative traits, such as disease resistance (Burow *et al*., 2014; Khera *et al*., 2019; Kumari *et al*., 2020), oil content and fatty acid composition (Wilson *et al*., 2017), as well as yield component traits (Fonceka *et al*., 2012a, 2012b; Sambou *et al*., 2017). These studies have shown that about half of the QTL positive effects were associated with alleles of the wild parent, highlighting the synthetic tetraploid's potential as a reservoir of useful alleles for peanut breeding. This emphasizes the significance of the synthetic-derived population, particularly in introducing favourable alleles into cultivated parents. However, further comprehensive studies are necessary to fully understand and harness this potential. It is also crucial to consider the impact of environmental factors and genetic variations on complex traits. Therefore, it's important to note that QTLs identified at a specific location may not universally apply to different environmental conditions.

The current study emphasizes a non-destructive approach utilizing NIR spectroscopy to explore the environmental and genetic influences on germplasm variability using intact peanut seed spectra without the need for chemical.

Furthermore, we aimed to detect favourable exotic QTL alleles for the improvement of yield-traits using an advanced backcross (AB-QTL) approach population derived from the crosses of Fleur 11 and ISATGR 278-18.

Therefore, this study aimed to address the following research questions:

questions:

- ✓ How effective is NIRS in screening peanut core-collection and interspecific populations for seed composition?
- ✓ What is the extent of phenotypic variability observed for yield-related traits within an interspecific population, and are these traits heritable with identifiable correlations?
- ✓ What genomic regions are associated with yield-related traits in this population?

Identifying top-performing lines or varieties for desired traits can lead to new varieties or serve as a source of beneficial wild alleles for trait enhancement. Additionally, pinpointing genomic regions associated with these traits can aid in marker-assisted selection.

Three hypotheses can be formulated from this study:

- ✓ NIRS will demonstrate a high level of effectiveness in screening peanut core-collection and interspecific population for seed composition;
- ✓ There is significant phenotypic variability for yield-related traits within an interspecific population, and these traits exhibit heritability with identifiable trait correlations;
- ✓ Specific genomic regions (QTL) will be found to be significantly associated with various yield-related.

The general objective of the present study was to explore both, rapid germplasm screening for quality traits, and the favorable wild QTL linked to yield traits.

The specific objectives of this study are:

- ✓ To screen peanut core-collection and interspecific population for quality traits;
- ✓ To assess phenotypic variability, heritability and trait correlations for yield-related traits in an interspecific population;
- ✓ To identify wild genomic regions (QTLs) associated with yield-related traits.

CHAPTER I. LITREATURE REVIEW

I.1. Peanut crop

I.1.1. Origin, distribution and centres of diversity

The origin of cultivated peanut is believed to be in the region of southern Bolivia and northwestern Argentina. It's theorized that peanut originated as a hybrid species in these areas (Radhakrishnan *et al*., 2022b; Massa *et al*., 2024). The domestication of cultivated peanut is estimated to has occurred around 3,500 to 7,000 years ago in the region spanning southeastern Bolivia, northwestern Argentina, and southern Brazil (Raj et al., 2022; Pan et al., 2023). Archaeological evidence suggests that indigenous communities in these regions cultivated peanut and incorporated them into their diets and cultural practices (Ambika *et al*., 2022; Phung *et al*., 2023). After its domestication in South America (Ambika *et al*., 2022; Pan *et al*., 2023), peanut spread throughout the continent (Hancock, 2022). Spanish and Portuguese explorers played a crucial role in introducing peanut to Europe, Africa, and Asia during the 16th and 17th centuries (Hancock, 2022; Smith & Reeves, 2023). Peanut was introduced to Europe, Africa, and Asia through these explorations, becoming a significant crop in various parts of the world.

Peanut exhibit significant genetic diversity, which is evident in several identified gene centres across South America. Bolivia is recognized as the primary centre, with additional secondary centres located in the Guarani region (Paraguay-Parana), Goias, and Mina Gerais regions of Brazil (Tocantins, Sao Francisco), Rondona and northwest Mato Grosso in Brazil (south Amazon), Peru (upper Amazon and west coast), and northeast Brazil (Foncéka *et a*l., 2013; Singh and Nigam, 2016), further highlighting the plant's genetic diversity and adaptability. These centres contribute to the plant's genetic variability and adaptability, as documented by various. In Africa, peanut has become integral to local agriculture and diets, establishing the continent as a tertiary centre of diversity for the crop (Singh and Nigam, 2016). This widespread cultivation emphasizes its importance and adaptation across different African regions.

I.1.2. Taxonomy and gene pool

Peanut, scientifically known as *Arachis hypogaea* (L.), belongs to the genus Arachis, which is a member of the. legume family *Fabaceae* (*Leguminosae*)) (Zahran & Tawfeuk, 2019; Ogbole *et al*., 2023).

The genus comprises about 81 species, mostly diploids (2n = 2x = 20), classified into nine sections based on morphological characteristics, cytological study, geographic distribution, and cross-compatibility (Stalker,2017; Pandey *et al*., 2020b). These sections include *Trierectoides, Erectoides, Procumbentes, Rhizomatosae, Heteranthae, Coleorhizae, Extranervosae, Triseminatae,* and *Arachis*. Section *Arachis* contains cultivated peanut (*A. hypogaea*) and another 30 wild species (Stalker,2017). Most of these species are diploid (2n = 2x = 20) with metacentric chromosomes of similar size (genomes A, B, F. and K); one species (*A. glandulifera*) is diploid with an asymmetric karyotype (genome D), and three can be considered dysploidy (2n= 2x= 18) (Krapovickas and Gregory 1994; Valls and Simpson 2005). The single wild tetraploid species, *A. monticola*, is very closely related to *A. hypogaea* (Lu and Pickersgill, 1993), probably sharing the same origin, and it is considered *A. hypogaea*'s immediate tetraploid ancestor (Seijo *et al*., 2007).

The most frequent of the genome types among the species is the A genome, characterized by the presence of a chromosome pair of reduced size and strongly condensed centromeric bands (Seijo *et al*. 2004). The next most frequent genome type is B, lacking a small chromosome pair and with chromosomes showing a lower degree of centromeric DNA condensation. Genome types F and K were formerly considered B genome species; recent classification was based on rDNA loci and the presence of strongly condensed centromeric bands in most chromosomes (Robledo and Seijo, 2010). Phylogenies based on DNA sequence data strongly support the validity of these genome divisions (Moretzsohn *et al.,* 2004, 2013; Milla *et al.*, 2005).

Based on cross-compatibility, the genetic diversity of the genus *Arachis* is classified into four gene pools (Janila *et al*., 2013; Fonceka *et al*., 2013; Nigam, 2014; Abady *et al*., 2021a). These gene pools play a significant role in understanding the genetic diversity and potential for breeding new peanut varieties. The primary gene pool consists of cultivated peanut, the secondary gene pool includes closely related wild species, and the tertiary and quaternary gene pools encompass other species with varying degrees of cross-compatibility. The secondary gene pool of *A. hypogaea* includes its most closely related wild species, which can be utilised for peanut crop improvement.

Peanut is an allotetraploid with an AABB genomic constitution (2n = 4x = 40), and its genome size is approximately 2.7 billion base pairs (2.7 Gb) (Bertioli *et al*., 2019; Zhuang *et al.*, 2019). This genomic constitution likely originated from a single hybridization event between two diploid progenitors, *Arachis duranensis* and *Arachis ipaensis*, which donated the

A and B sub-genomes, respectively (Bertioli *et al*., 2019; Zhuang *et al.*, 2019). Genetic, cytogenetic, phylogeographic, and molecular evidence supports the notion that these two diploids are the most likely progenitors of cultivated peanut (Kochert *et al.*, 1996; Burow *et al*., 2009; Moretzsohn *et al*., 2004, 2013; Ramos *et al*., 2006; Seijo *et al*., 2004, 2007).

Morphological variation in cultivated peanut led to divide it into two subspecies, *hypogaea* and *fastigiata,* and six botanical varieties (Stalker, 2017; Al-Khayri *et al*., 2019). These subspecies are distinguished by patterns of reproductive and vegetative branching, as well as pod morphology. Subspecies *hypogaea* is further divided into two botanical varieties: *hypogaea* (virginia) and *hirsuta*. Subspecies fastigiata is also divided into four botanical varieties: *fastigiate* (Valencia), *vulgaris* (Spanish), *peruviana*, and *aequatoriana*.

Despite the high morphological diversity, different origins for the two subspecies were proposed, supported by partial reproductive isolation (Singh and Moss, 1982; Lu and Pickersgill, 1993). However, further investigation using molecular data contradicted this hypothesis. The genetic variability observed among commercial cultivars and landraces of peanut is relatively low, leading to the general acceptance that peanut is an allotetraploid of recent and single origin (Halward *et a*l., 1993; Kochert *et al*., 1996; Milla *et al*., 2005). This conclusion was reached after examining molecular data and understanding the genetic relationships between different peanut varieties.

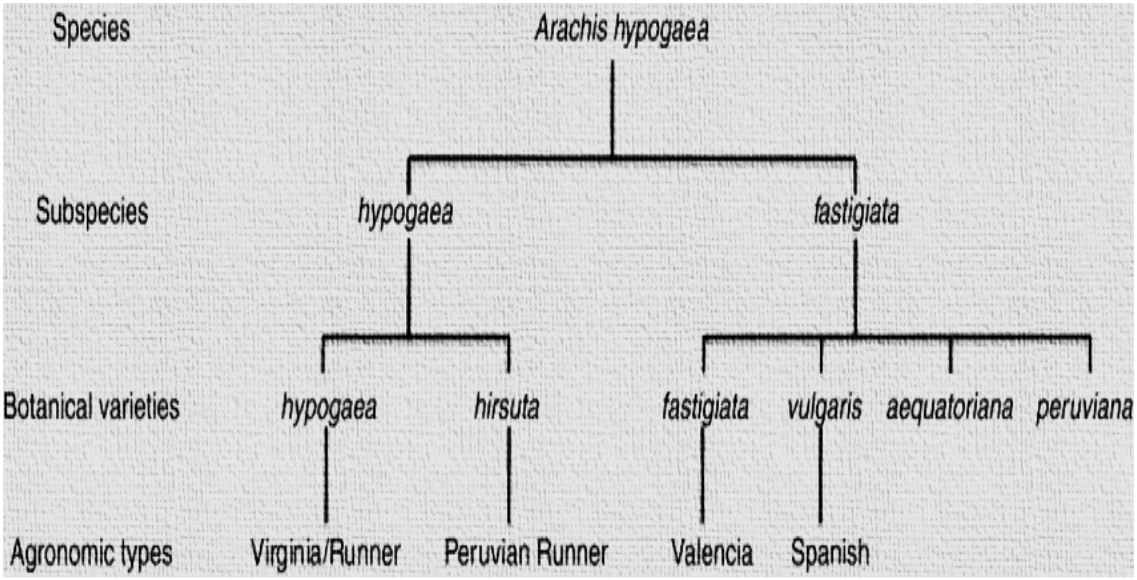


Fig. 1.The taxonomic arrangement of subspecies and botanical varieties (Stalker, 2017).

I.1.3. Botany and morphology

Understanding the botanical characteristics of peanuts is essential for advancing agricultural research and development. Peanut, known for its unique geocarpic reproductive habit, have fertilized ovaries that develop underground into pods (Harbau & Sanusi, 2023). The plant's pinnately compound leaves feature four leaflets symmetrically arranged on both the main stem and side branches (IBPGR and ICRISAT, 1992; Raj *et al*., 2022; Radhakrishnan *et al*., 2022a). These leaflets are typically ovate or elliptical, with smooth or slightly waxy margins and a cuticle that minimizes water loss.

Peanut growth habits, crucial for cultivation strategies, encompass various types: procumbent-1, procumbent-2, decumbent-1, decumbent-2, decumbent-3, and erect (Fig. 3) (IBPGR and ICRISAT, 1992; Upadhyaya & Gowda, 2009; Kayam *et al.*, 2017; Janila *et a*l., 2018; Kumari *et al*., 2020; Fang *et al*., 2023). Branching patterns, such as alternate, sequential, irregular with flowers on the main stem, and irregular without flowers on the main stem (Fig.4), further influence agricultural practices (IBPGR, 1992*;* Ntare *et al*., 2008; Upadhyaya & Gowda, 2009; Janila *et a*l., 2018*)*.

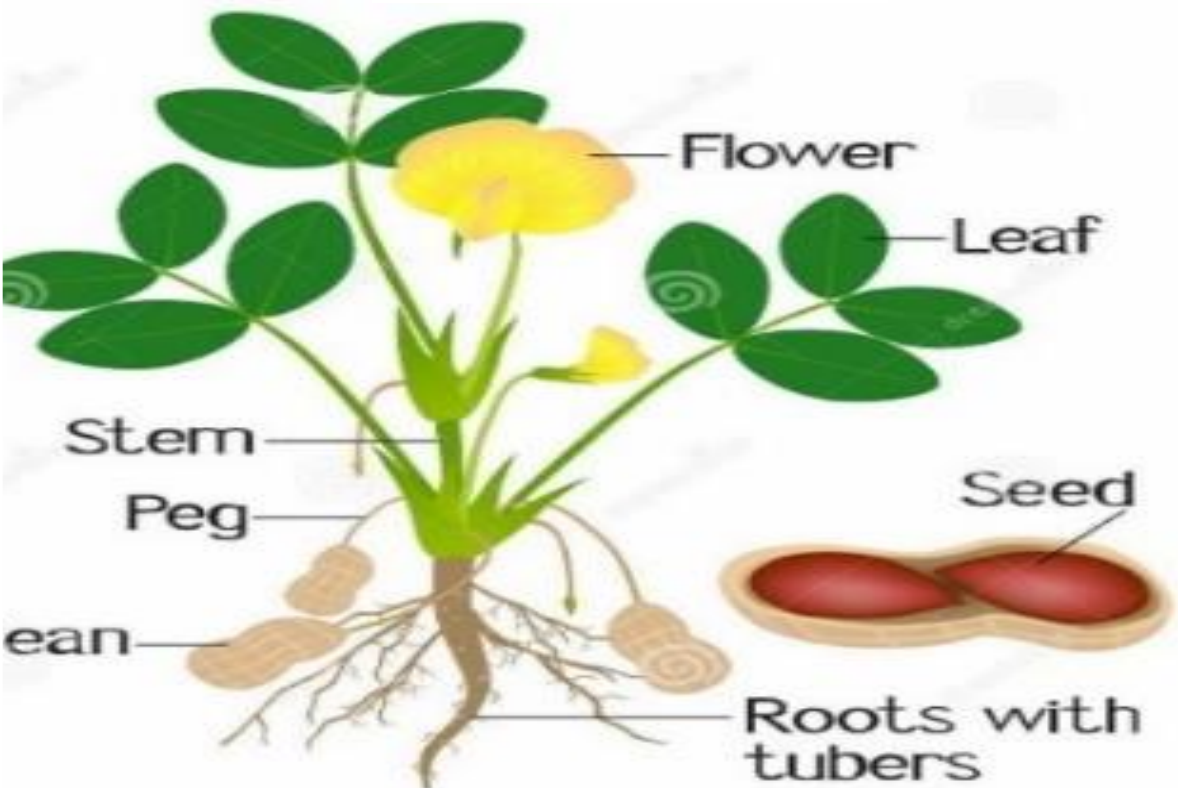


Fig. 2. Part of peanut (IBPGR and ICRISAT, 1992).

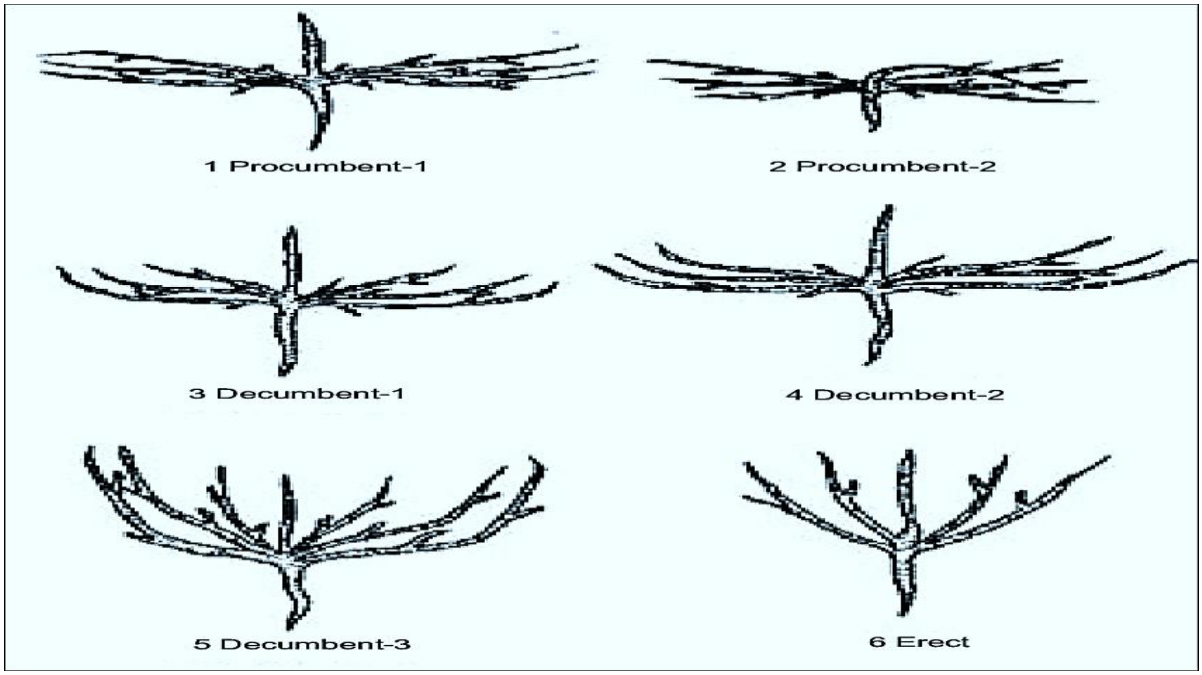


Fig. 3. Peanut growth habits (Upadhyaya & Gowda, 2009; Janila *et al*., 2018*)*.

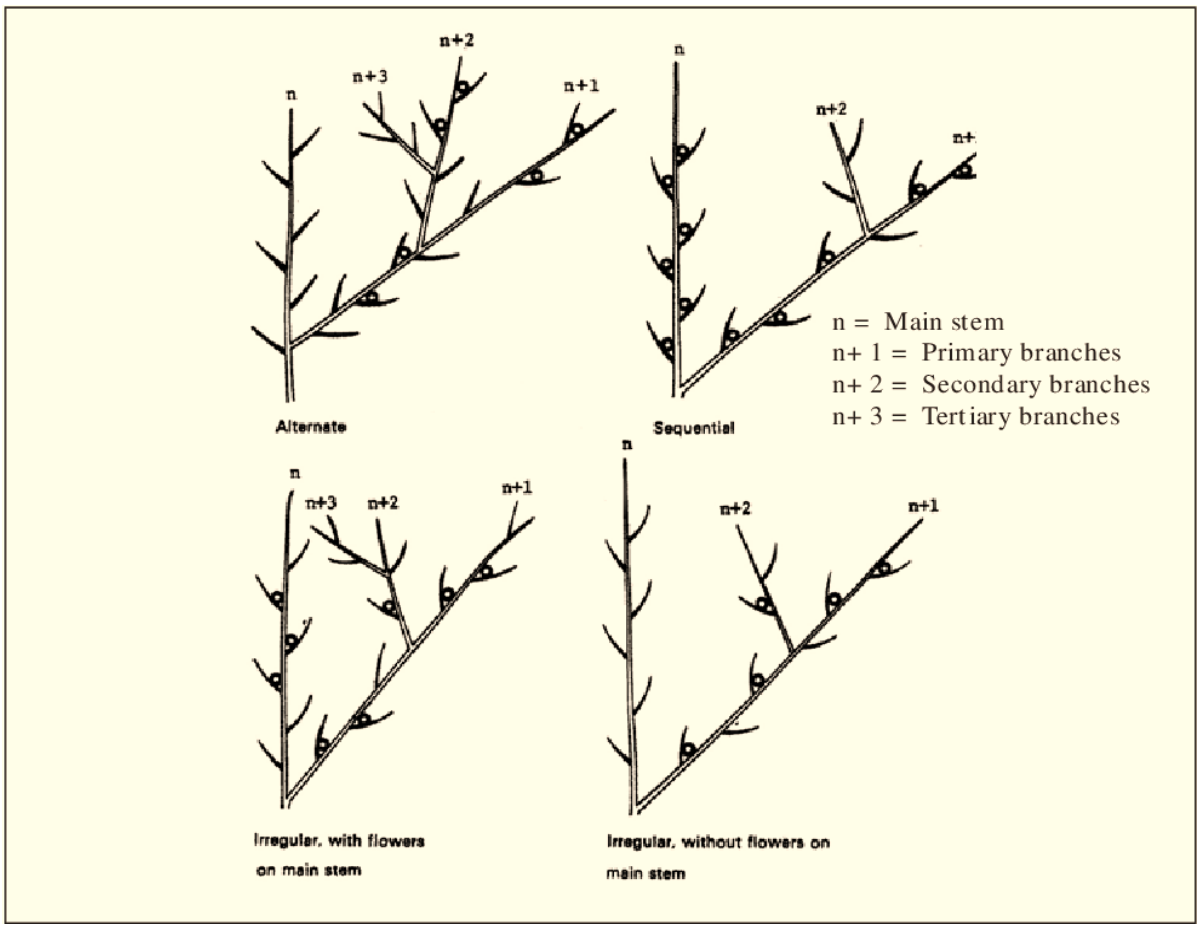


Fig. 4. Peanut branching patterns (*Ntare et al., 2008; Janila et al., 2018)*.

Peanut flowers are self-pollinating and arranged in inflorescences along the stem. They vary in colour from orange to yellow. After fertilization, the ovary elongates into a sturdy peg, which descends into the soil and develops into the mature pod (Nigam, 2014; Raj *et al*., 2022; Radhakrishnan *et al*., 2022a). Peanut plants have a well-developed taproot system that can extend up to 135 cm deep, accompanied by prostrate stems that can reach lengths of up to 60 cm (Harbau & Sanusi, 2023).

I.1.4. Production

Peanut is an important legume crop cultivated across more than 100 countries (Fig. 5), encompassing approximately 30 million hectares (Mha) of land, with global production reached around 54 million tons (MT) and an average yield of 1.8 tons per hectare (t/ha) (FAOSTAT, 2024). Asia stands as the primary peanut-producing region worldwide, with China and India being major contributors, yielding 18 MT and 10 MT, respectively. Africa follows closely, with Nigeria and Sudan as significant producers, yielding 4 MT and 2.5 MT, respectively (FAOSTAT, 2024). While Asia and Africa collectively dominate global peanut cultivation area and production, Africa faces challenges in achieving high yields compared to Asia and the Americas, with an average yield of 1 t/ha lagging behind Asia's 2.6 t/ha and America's 3.6 t/ha. Although Africa covers 56% of the global peanut-growing area, it contributes only 30% to global production. In contrast, Asia covers 39% of the global peanut-growing area but produces 59.6% of the global output, with China significantly impacting higher yields. The Americas, led by the United States, contribute 9.9% of global production from just 1.5% of the global peanut-growing area.

Over the past six decades, there has been a noticeable increase in global peanut productivity, with the global yield average rising from 800 kg/ha in 1965 to 1797 kg/ha in 2020. Africa has also shown improvement, with the average yield increasing from 609 kg/ha in 1983 to 1078 kg/ha in 2006. Countries like Cameroon and Ethiopia have experienced significant growth in both cultivated area and production. For instance, Cameroon observed a rise in peanut yield from 234 kg/ha in 1989 to a peak of 1747 kg/ha in 2006, while Ethiopia witnessed an increase from 456 kg/ha in 1998 to 1807 kg/ha in 2020.

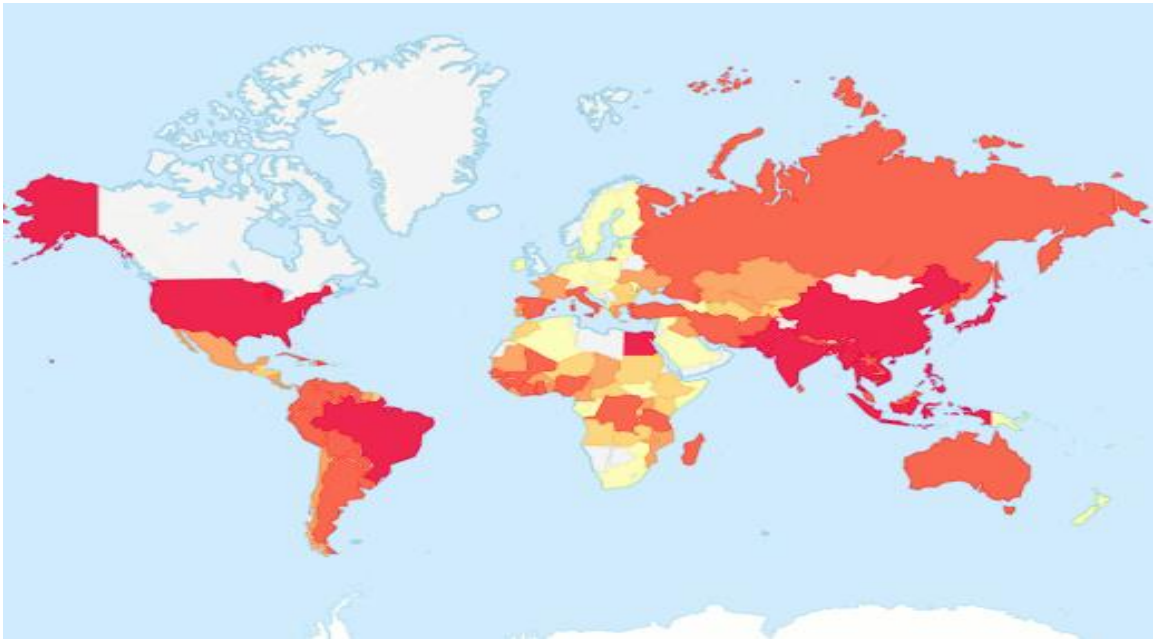


Fig. 5. Worldwide distribution of peanut production in 2022 (FAOSTAT, 2024).

As peanut ranks among the major vegetable oilseed crops, over 60% of global peanut production is crushed to extract oil for both edible and industrial uses, while 40% is used in food and other applications (Janila *et al*., 2013, 2016). According to Statista (2023), peanut oil production increased from 5.5 million tons in 2012-2013 to 6.5 million tons in 2022-2023, with an average production of 6.4 million tons during the latter period. In the same year, total vegetable oil production reached approximately 217 million tons (MT), with peanut oil ranking sixth in terms of production volume at 6.5 million tons. The leading vegetable oils by production were palm oil at 76 million tons, followed by soybean oil at 60 million tons, rapeseed oil at 31 million tons, sunflower seed oil at 18.6 million tons, and palm kernel oil at 8.8 million tons.

As the global population is projected to expand from 7.2 billion to 9.6 billion by 2050, the demand-supply gap for food grains continues to increase (Pandey *et al*., 2016). This challenge necessitates an increase in global food production by over 70% by 2050 to ensure a hunger-free society with nutritious food (Rajwade *et al*., 2015; Pandey *et al*., 2016; Alhashim and Anandhi, 2022). The projected rise in global demand for peanut and its related products underscores the need to enhance production and productivity. Meeting this demand requires maximizing efforts to develop improved high-yielding cultivars resistant to major stresses and possessing high seed quality traits through the integration of genomic tools with peanut improvement.

I.1.5. Importance

Peanut is valued for its versatility and widespread utility across different regions worldwide. Primarily grown for its vegetable oil, constitutes a significant source of nutrition and health benefits (Davis *et al*., 2016; Abady et al., 2021b). Its seed contains approximately 34 to 56% oil, 22 to 30% protein, and 10 to 25% carbohydrates. Moreover, it boasts a wealth of micronutrients vital for various bodily functions, including vitamin E, K, and B complex, folic acid, niacin, antioxidants, and biologically active polyphenolics such as flavonoids and isoflavones. Additionally, peanut offer an abundance of essential minerals like calcium, phosphorus, magnesium, zinc, and iron, all of which contribute to their nutritional value and potential health benefits (Harch *et al*., 1995; Janila et al., 2013, 2016; Desmae et al., 2019). With their high energy content from oil and protein, peanuts can serve as an alternative to calorie-dense foods like red meat, aiding in calorie reduction.

Peanut oil is rich in various fatty acids, comprising approximately 12 different types. Among these, oleic and linoleic acids dominate, collectively making up nearly 80% of the oil's composition. However, palmitic acid, a saturated fatty acid, also plays a significant role, contributing approximately 10% to the total fatty acid content (Davis *et al*., 2016; Nawade *et al*., 2018). Additionally, minor fatty acids such as stearic, arachidic, eicosenoic, behenic, lignoceric, and gadoleic acids are present, making up about 10% of the total fatty acid content (Davis *et al*., 2016; Nawade *et al*., 2018).

Furthermore, peanut oil with high levels of oleic acid exhibits nearly 10 times greater auto-oxidative stability compared to linoleic acid-rich oils (Nawade *et al*., 2018). This enhanced stability enables heating at high temperatures without smoking, facilitating faster cooking and reducing oil absorption during cooking. Additionally, oils high in oleic acid impart a pleasant aroma and contribute to the prolonged shelf life of peanut products. Consuming peanut products rich in oleic acid has been associated with several health benefits, including cholesterol reduction, tumor suppression, inflammation alleviation, blood pressure regulation, and improved lipid and glucose levels. (Pandey *et al*., 2014b; Nawade *et al*., 2018; Bonku and Yu, 2020). Studies also suggest a potential risk reduction of type 2 diabetes and cognitive enhancement in healthy young adults through regular peanut consumption (Jiang, 2002). In another study involving 63 healthy young adults, it was found that regular consumption of peanut and peanut butter may enhance memory function and stress response (Parilli-Moser *et al.*, 2022).

Beyond oil extraction, peanut is consumed in diverse forms (Fig. 6) such as raw, roasted, boiled, or processed into peanut butter, catering to varied culinary preferences and dietary needs. (Janila *et al*., 2013; Abady *et al*., 2021b). This versatility extends to their role in supporting small-scale producers by generating revenue and contributing to foreign currency earnings through exports. As legumes, peanut contribute to soil improvement by biologically fixing nitrogen, enhancing soil fertility sustainably. Additionally, peanut haulm serves as valuable livestock feed, offering essential nutrients like protein, lipids, minerals, and carbohydrates, particularly beneficial in regions with limited grazing lands (Janila *et al*., 2013, 2016).

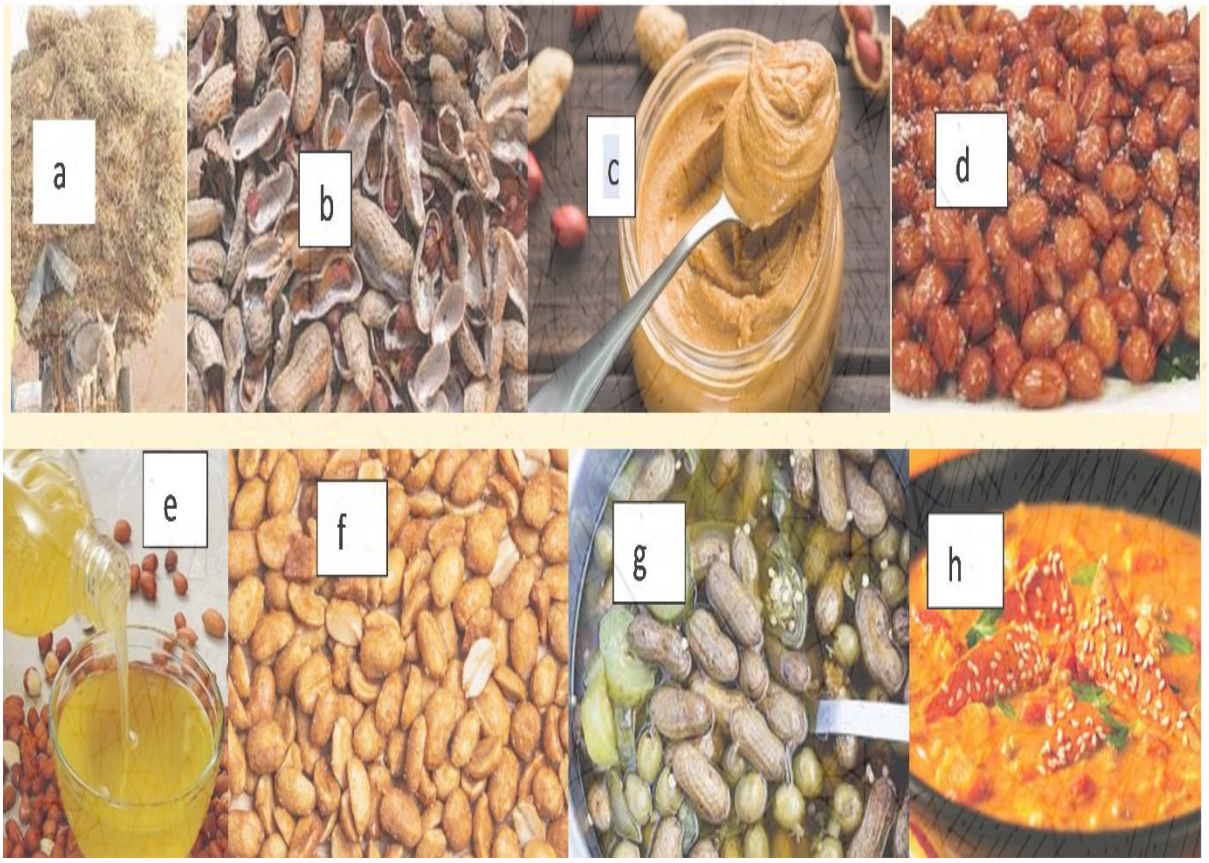


Fig. 6. Importance of peanut, a*)* peanut *halum, b). peanut shell, c)* peanut *butter, d) fried peanut, e) peanut oil, f) roasted peanut) boiled peanut and h) peanut source* (Janila *et al*., 2013; *Abady et al., 2021b*).

However, it is important to acknowledge the potential health risks associated with peanut consumption. One significant concern is the prevalence of peanut allergies, which can manifest as mild to severe symptoms, including anaphylaxis (Bonku and Yu, 2020). Another critical issue is aflatoxin contamination, a toxic compound produced by mold that poses serious health risks such as liver damage and cancer development (Yu *et al*., 2019).

I.2. Targeted traits for peanut improvement

In peanut cultivation, genetic improvement targets various traits to meet diverse regional needs (Fig.7). These include yield-related traits, resistance to diseases and pests, drought tolerance, and quality enhancements for consumers and industries (Janila *et al*., 2013, 2016; Vishwakarma *et al*., 2017). Factors such as growing seasons, producer requirements, consumer preferences, market demands, and industrial specifications influence the selection of these traits.

Yield-contributing traits such as pod yield per plant, number of pods per plant, shelling percentage, and 100-seed weight are crucial for maximizing productivity. Additionally, traits like peg strength, pod morphology (including reticulation, beak, and constriction), kernel shape and colour, fresh seed dormancy, and blanching ability are vital considerations to meet the needs of farmers, processors, and market demands. (Janila *et al.*, 2016). In rainfed regions, early maturity is pivotal to aligning the crop cycle with the duration of the rainy season, ensuring harvest before the onset of the dry season to mitigate drought-related risks.

In regions like Asia and Africa (Janila *et al.*, 2016), where peanut is primarily used for oil production, increasing oil content is a crucial target trait for advanced breeding programs. The specific intended use of peanut plays a significant role in determining the key traits that are essential for the breeding process. These traits include low oil and high protein contents for food consumption, high oil content for oil production, and a high oleic/linoleic fatty acid ratio for enhanced shelf-life of products derived from peanuts. (Teres *et al*., 2008; Vassiliou *et al*., 2009; Carrillo *et al*., 2012; Parilli-Moser *et al*., 2022). Breeding programs focused on improving shelf-life often prioritize achieving a high oleic/linoleic fatty acid ratio as the target trait. This emphasis is supported by research indicating that a higher oleic acid content and a lower linoleic acid content led to a more stable oil that is less prone to oxidation and rancidity.

For example, high-oleic acid content offers significantly greater auto-oxidative stability compared to linoleic acid, contributing to prolonged shelf life for peanut products and, consequently, enhancing human nutrition (Nawade *et al*., 2018; Desmae *et al*., 2019; Davis *et al*., 2021). By targeting traits such as high oleic/linoleic fatty acid ratios in peanut breeding programs, researchers and breeders aim to develop varieties that not only meet the specific needs of different markets but also offer improved product quality, extended shelf life, and enhanced nutritional benefits.

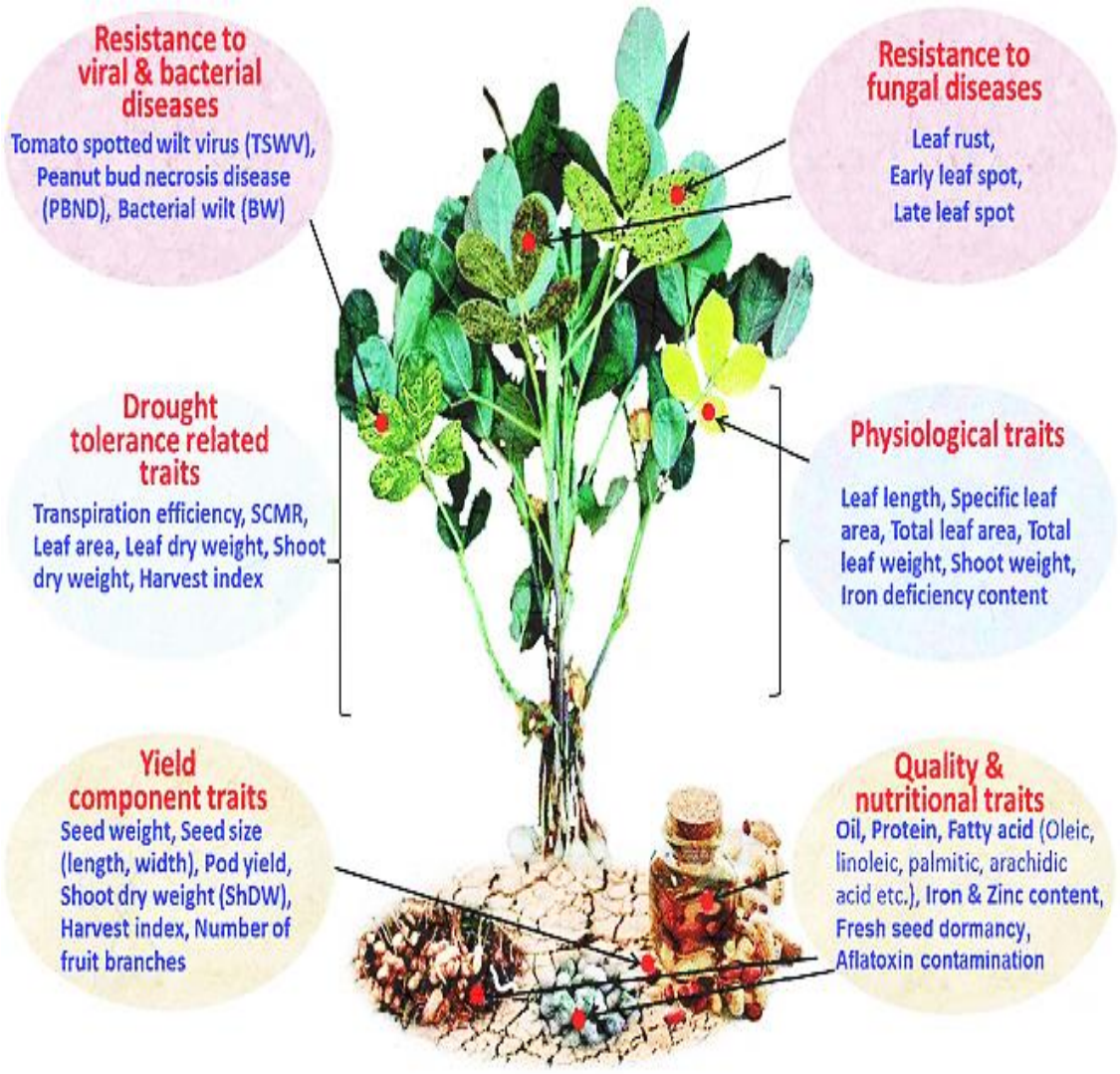


Fig. 7. Traits improved using genomic assisted breeding in peanut (Vishwakarma *et al*., 2017).

I.3. High-throughput phenotyping techniques in peanut seed quality traits

The two main factors restricting peanut breeding internationally are the limited genetic base of cultivated peanut and the inadequate exploitation of genetic resources. Thus, wide hybridization is a practical method for introducing potential diversity from wild species into the cultivated gene pool (Simpson *et al*., 1993; Fávero *et al.*, 2006; Mallikarjuna *et al*., 2010). Likewise, the use of germplasm in breeding techniques to improve food security depends critically on the assessment of genetic diversity and the identification of promising genotypes. To facilitate the investigation of large germplasm collections, it is reasonable to start by examining customized sets of germplasm that exhibit appropriate diversity and are of manageable size. Two common types of such customized sets are core collections and interspecific populations derived from wild × elite crosses.

Although significant efforts have been made to target traits in peanut for characterizing cultivars and germplasm collections for simple traits and for the most important agronomic traits (yield and resistance to pests and diseases) (Upadhyaya *et al*., 2002, 2005, 2006, 2011; Mallikarjuna *et al*., 2012; Kumari *et al*., 2014), less is known about various quality traits (Bianchi-Hall *et al*., 1993; Grosso *et al*., 2000)**.** This is mainly due to the fact that the phenotyping of these traits, regularly based on chemical surveys, is expensive in terms of both direct monetary input and human labour, time-consuming, complex, and irreversibly destructive. Another main factor limiting chemical studies are the difficulties to analyse many samples, each requiring many seeds (Nawade *et al*., 2018; Davis *et al*., 2021).

The currently employed chemical-based seed quality analysis methods include Soxhlet extractor (Shasidhar *et al*., 2017), nuclear magnetic resonance (NMR) (Pandey *et al*., 2014b; Liu *et al*., 2020; Guo *et al*., 2021), high-performance liquid chromatography (HPLC) and gas chromatography (GC) (Lin *et al*., 2016; Shasidhar *et al*., 2017). These methods may provide accurate results in laboratories, but the procedures are usually time-consuming, sample-destructive, expensive, and require skilled personnel to perform, which make them impossible for large-scale non-destructive screening detection. On the other hand, high-throughput phenotyping (HTP) techniques have revolutionized the study of plant traits by allowing researchers to collect large amounts of data in a short period. These include modified refractive index, capacitance sensor (Kandala *et al*., 2008), hyperspectral imaging (Huang *et al.*, 2014; Rabanera *et al.*, 2021) and near infrared spectroscopy (NIRS) (Govindarajan *et al*., 2009; Davis *et al*., 2021; Tao *et al.*, 2019). These techniques have been widely applied to various crops, including peanut, to assess seed quality traits.

Among HTP, NIR-based methods are the most rapid, make it possible to analyse large numbers of samples and for peanut breeders, an added benefit with NIR is that the method can be non-destructive. In the case of peanut, NIRS allowed for the efficient evaluation and selection of peanut germplasm based on various quality traits, such as oil content, protein content, fatty acid composition (Sarvamangala *et al*., 2011; Bansod *et al*., 2015; Lin *et al*., 2016; Shasidhar *et al*., 2017; Sun *et al*., 2021; Davis et al., 2021) , moisture content determination (Govindarajan *et al*., 2009) and to detect aflatoxin (Tao *et al.*, 2019). Moreover, a number of studies have also applied machine learning as promising statistical methods to assist humans in the modelling and analysis of complex spectral data (Wang and Song, 2023; Xu *et a*l., 2023) in many research fields including seed quality detection, varieties identification(Singh et al., 2023) and classification (Shang et al., 2023; Tian *et al*., 2023).

I.4. Peanut genetic resources

The genetic resources of the peanut plant play a crucial role in the improvement of peanut cultivars, as they provide valuable genetic variation for desirable traits. These resources are preserved in gene banks worldwide and serve as reservoirs of useful genes for current and future breeding programs. The peanut crop is blessed with large germplasm collections maintained in various institutions across the globe (see: Pandey *et al.*, 2012). At ICRISAT, India holds the largest collection, with 15,445 accessions from 93 countries. This is followed by the National Bureau of Plant Genetic Resources (NBPGR) with 14,585 accessions, and the Directorate of Groundnut Research (DGR) of the Indian Council of Agricultural Research (ICAR) with 9,024 accessions in India. Additionally, the Oil Crops Research Institute (OCRI) of the Chinese Academy of Agricultural Sciences (CAAS) holds 8083 accessions, and the Crops Research Institute of the Guangdong Academy of Agricultural Sciences maintains 4,210 accessions in China. In the United States, the Plant Genetic Resource Conservation Unit (PGRCU) at Griffin, U.S. Department of Agriculture (USDA) holds 9,024 accessions, while North Carolina State University (NCSU) maintains 1,146 accessions. Furthermore, EMBRAPA-CENARGEN holds 1,200 accessions, and the Instituto Agronomico de Campinas maintains 2,140 accessions in Brazil. Finally, National Institute of Agricultural Technology (NIAT) in Argentina holds 3,640 accessions, and Texas A&M University (TAMU) in the USA holds 1,200 accessions, along with the Northeast Botanical Institute (NEBOI) in Argentina maintaining 472 accessions (Pandey *et al.*, 2012).

The challenge of selecting appropriate lines from a large number of cultivated accessions can be addressed through the use of core collections, which are subsets typically containing 10% of the entire collection (Brown, 1989). For instance, core collections have been developed in various regions, such as the 1,704 core collection at ICRISAT, the U.S. core collection with 831 accessions (Holbrook *et al*., 1993), and the Chinese core collection with 576 accessions (Jiang *et al.*, 2008).

However, the size of these core collections can still be unwieldy for breeders to fully exploit. To address this, 'mini-core collections' have been developed, representing 10% of the core collections and 1% of the entire germplasm collection. These smaller collections consist of 184, 112, and 298 accessions at ICRISAT (Upadhyaya *et al*., 2002), USDA/ARS (Holbrook and Dong, 2005), and in China (Jiang *et al*., 2010), respectively.

Additionally, a global composite collection of 1,000 accessions was created, and a reference set of 300 genetically diverse accessions was further characterized using 21 SSR markers (Upadhyaya *et al*., 2006). In addition to germplasm collections that represent naturally occurring variation, amphiploids and autotetraploids (Mallikarjuna *et al*., 2010), over 3400 targeting-induced local lesions in genomes (TILLING) populations (Knoll *et a*l., 2011), multiparent advanced generation intercross (MAGIC) populations (Janila *et al*., 2013; Pandey *et al*., 2020b), and chromosome segment substitution (CSSL) lines (Tossim *et al*., 2020) have been developed and form important resources of groundnut breeding.

I.5. Use of wild relatives in peanut improvement

Utilising wild relatives in peanut improvement presents a promising avenue for addressing key production constraints and enhancing long-term sustainability in agriculture. Diversity studies employing molecular markers have revealed that cultivated peanut exhibits a low level of genetic diversity due to its origin and reproductive isolation from wild diploid species, primarily stemming from differences in ploidy level (Herselman *et al*., 2004). The limited genetic variability in cultivated germplasm poses challenges for peanut improvement to tackle key production constraints such as drought, environmental changes, and diseases. In contrast, wild peanut relatives demonstrate high genetic diversity and harbour valuable alleles for enhancing resistance, abiotic stress tolerance, yield potential, seed quality traits, and long-term sustainability amid climate change and evolving agricultural challenges (Upadhyaya *et al.*, 2011; Stalker, 2017). Consequently, peanut breeders have increasingly focused on incorporating new alleles from wild species into cultivated peanut to broaden its genetic base and unlock genetic potential. Two primary pathways have been described for introducing alleles from wild species into cultivated peanuts: the hexaploidy pathway and the tetraploid pathway (Simpson, 2001).

In the hexaploidy pathway, breeders directly cross a given diploid species with A. hypogaea, resulting in a triploid (3x = 30) sterile hybrid that can be doubled to the hexaploidy (6x = 60) level through colchicine treatment. Despite cytological instability, successful efforts have introgressed alleles from wild diploid species into cultivated peanut, enhancing resistance to diseases and pests. Notable successes include the release of cultivars such as 'Spancross' and 'Tamnut 74', developed from crosses between *A. hypogaea* and *A. monticola* (Fonceka *et al*., 2013).

Similarly, germplasm lines resistant to rust and late leaf spot have been released from crosses between *A. hypogaea* and *A. cardenasii* (Stalker, 2017; Motagi et al., 2022). Among these lines, GPBD 4, resistant to rust and late leaf spot, was released by crossing ICGV 86855, an interspecific derivative between A. hypogaea and A. cardenasii, with KRG1, an early maturing line from Argentina (Gowda et al., 2002).

The tetraploid pathway in peanut breeding has significantly advanced through the creation of amphidiploids, which are tetraploid plants developed from hybridizations between different *Arachis* diploid species. Autotetraploids (AAAA) or allotetraploids (BBBB) produced through colchicine treatment are cross-compatible with cultivated peanut and usable in introgression programs. Simpson *et al*. (1993) pioneered this pathway by creating the first amphidiploid from a three-way cross between A. cardenasii and A. diogoi. From this cross, commercial cultivars such as COAN, NemaTam, Tifguard, Webb, TifN/V OL, Georgia 14N, Tifguard and TifN/V OL have been relased (Simpson *et al*., 2013, Holbrook *et al.*, 2017; Motagi *et al.*, 2022). Additional amphidiploids, coded as AiAd, developed from hybridization between *A. ipaensis* and *A. duranensis*, have contributed to expand genetic diversity in the Arachis genus (Favero *et al*., 2006). Mallikarjuna *et al*. (2010) also developed 17 new amphidiploid genotypes through hybridization between different species of section Arachis. These amphidiploids have served as valuable resources for breeding programs, enabling the creation of cultivars with enhanced disease resistance and oil composition. Despite the promise of incorporating wild alleles, the process of introgressing useful traits presents challenges due to tight linkages between beneficial and undesirable traits, necessitating multiple backcrossing cycles to recover desirable traits.

Recent advancements, including the use of synthetic tetraploids and next-generation sequencing (NGS), have significantly improved the integration of peanut wild relatives in breeding programs. Synthetic amphidiploids have expanded genetic diversity in the *Arachis* genus and facilitated the creation of cultivars with improved disease resistance and oil composition. Synthetic amphidiploids such as TxAG-6, AiAd, ISATGR 1212, ISATGR 265, ISATGR 278-18, and ISATGR52B have been key in evaluating the genetic potential of wild species and mapping QTL regions (Fonceka *et al*., 2012a, b; Nguepjop *et al*., 2016; Wilson *et al*., 2017; Sambou , 2017; Khera *et al*., 2019; Kumari *et al*., 2020).

Molecular marker-based approaches such as Advance Backcross Quantitative Trait Loci (AB-QTL) and Chromosome Segment Substitution Lines (CSSL) are currently being utilised to broaden the genetic base of cultivated varieties, facilitate the introgression of desirable traits from wild relatives into elite lines and to map QTL (Fonceka *et al*., 2012a, 2012b; Tossim *et al*., 2020).

I.6. Quantitative trait loci (QTLs) mapping

A QTL is a genomic region that is responsible for the quantitative variation of a trait. A quantitative trait is a measurable attribute based on the combined activity of one or many genes and their interactions with the environment, which can vary between individuals over a given range to generate a continuous distribution of phenotypes (Collard *et al*., 2005). Quantitative trait loci mapping is a statistical method used for identifying responsible genes, understanding variation mechanisms, determining how many QTL contribute significantly to the trait, determining how much variation is due to additive, dominant. and epistatic effects, and determining the nature of the genetic correlation between different traits in a genomic region (Andersen and Torp, 2002). It helps breeders understand the genetic basis of complex traits by identifying the genetic variants (QTL) that contribute to the trait variation. The steps involved in biparental QT mapping are presented in (Fig. 8).

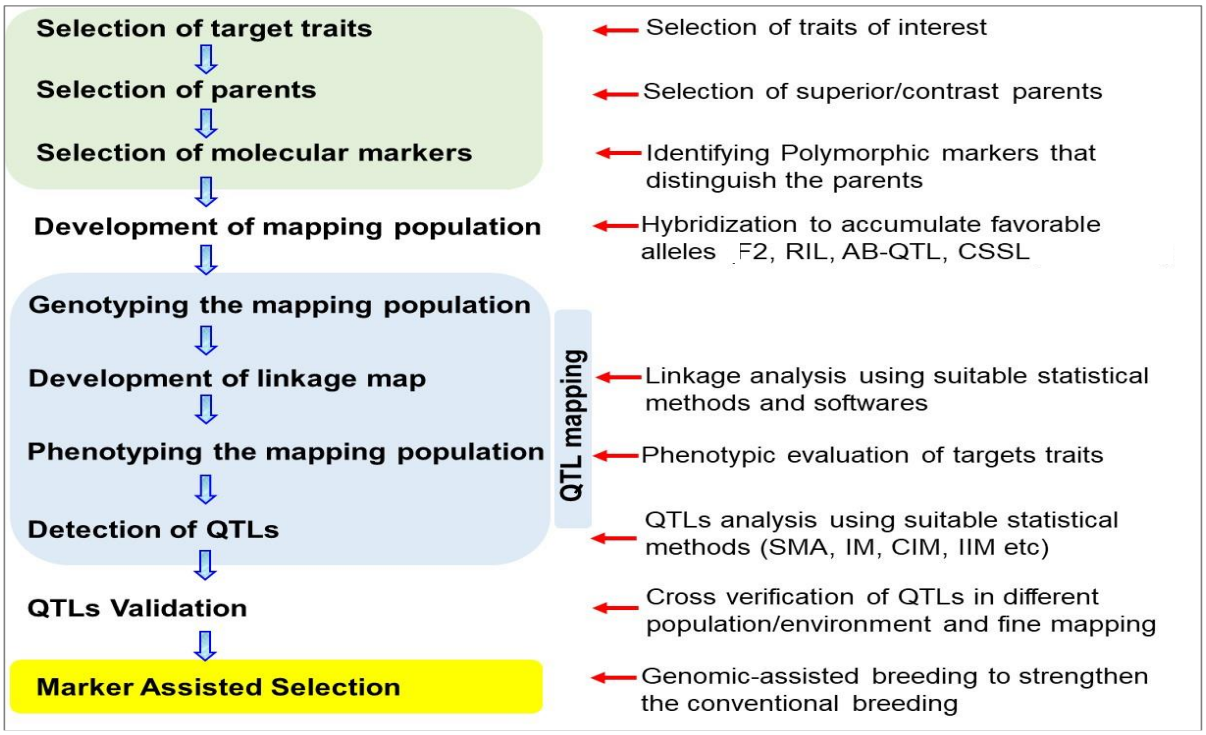


Fig. 8. Steps involved in biparental QTL mapping (Kassie *et al*., 2023).

I.6.1. Molecular markers

Molecular markers (DNA) are genes or DNA sequences with known chromosome locations that play crucial roles in plant breeding. They confirm hybrid identities, track introgressed chromosomal segments, and solve the problem of linkage drag (Singh and Singh, 2015; Nayak *et al*., 2017; Nadeem *et al*., 2018; Kumawat *et al*., 2020; Vishwakarma *et al*., 2022). These markers have been used in several genetic studies, including trait mapping and. most importantly. molecular marker-assisted breeding.

The early generation DNA markers Restriction Fragment Length Polymorphisms (RFLPs) (Halward *et al*., 1993; Burow *et a*l., 2001), Random Amplified Polymorphic DNAs (RAPDs) (Garcia *et al.*, 2005), and Amplified Fragment Length Polymorphisms (AFLPs) (Milla, 2003) were initially used for genetic mapping in peanut. However, due to their limited number and associated limitations, researchers have increasingly turned to microsatellite or simple sequence repeat (SSR) markers, which have been extensively utilized in genetic and QTL mapping studies (Varshney *et al*., 2009; Pandey *et al.*, 2012; Fonceka *et al*., 2013; Pandey *et al*., 2014a; Vishwakarma *et al*., 2017; Desmae *et al*., 2019). Moreover, single-nucleotide polymorphism (SNP) markers (Zhou *et al.*, 2014; Xiaojing *et al*., 2014; Bertioli *et al*., 2014; Liang *et al*., 2017; Liu *et al*., 2020; Sun *et al*., 2021) and diversity array technology (DArT) markers (Vishwakarma *et al*., 2016; Shashidhar *et al*., 2017; Khera *et al*., 2019) have been incorporated, significantly expanding the repertoire of available molecular markers for peanut genetic and breeding applications.

I.6.2. Genetic map

Genetic or linkage mapping involves arranging markers in order, indicating their relative distances, and assigning them to linkage groups based on recombination values from pairwise combinations (Andersen and Torp, 2002; Collard *et al*., 2005). This technique facilitates the identification of chromosomal locations harbouring genes and Quantitative Trait Loci (QTLs) linked to traits of interest. By determining the relative positions and distances of these markers, scientists gain insights into the genetic basis of traits and can enhance breeding strategies.

I.6.2.1. Diploid genetic map

I.6.2.1.1. Genetic maps for AA-genome

In the development of genetic maps for the AA-genome in peanut, various populations including F2, F5, F6, and BC1F1 have been utilised, alongside marker systems such as RFLP,

AFLP, RAPD, SSR, and SNP (Halward *et al.*, 1993; Milla, 2003; Moretzsohn *et al.*, 2005; Leal-Bertioli *et al*., 2009; Nagy et al., 2012). The initial map, based on RFLP markers, employed an F2 population from a cross between *A. stenosperma* and *A. cardenasii*, yielding 117 loci across 11 linkage groups covering 1,063 cM. Subsequent maps, utilizing AFLP and SSR markers, expanded marker diversity and map coverage (Halward *et al*., 1993). The first AFLP-based map was created using the F2 population developed from *A. kuhlmannii* x *A. diogoi* (Milla, 2003). This map consisted of 102 markers grouped into 12 linkage groups and spanned 1068.1 cM. The first peanut SSR-based map was created for an F2 population resulting from a hybrid between *A. duranensis* and *A. stenosperma* (Moretzsohn *et al*., 2005). A diploid backcross population derived from the same parents was also used to compute a linkage map (Garcia *et al*., 2005). One hundred and sixty-seven RAPD and 39 RFLP loci were mapped into 11 linkage groups, spanning 800 cM. The 39 RFLP markers were common to the F2-based map of Halward *et al*. (1993) and were used to establish correspondences between both maps.

The first peanut SSR-based map was constructed for an F2 population derived from a cross of two diploid species with *A genomes*, *A. duranensis* and *A. stenosperma* (Moretzsohn *et al*.2005). One hundred and seventy loci were mapped into 11 linkage groups covering 1,231 cM of total map distance. New markers were added to this map, resulting in 369 loci, including 188 microsatellites, 80 anchors and 35 resistance gene analogue (RGA) markers, mapped into ten linkage groups, as expected for diploid species of Arachis (Leal-Bertioli et al., 2009).

Another genetic map, created using an F2 population derived from the cross (*A. duranenis* x *A. duranensis*), employed a combination of markers such as 971 SSRs, 221 single-stranded DNA conformation polymorphism (SSCP) markers, and 1,127 SNPs mapped on 10 linkage groups (Nagy *et al*., 2012). Later, three more genetic maps were constructed using the F5 and F6 generation with the SSR, SNP, transposable element (TE), resistance gene analog (RGA) and anchor markers of the population generated from a cross between *A. duranensis* and *A. stenosperma* (Bertioli *et al*., 2014; Leal-Bertioli *et al*., 2016). These maps contain 597,384 and 502 markers on 544,705.10 and 1004.1 cM map distance respectively.

I.6.2.1.2. Genetic maps for BB-genome

For the BB-genome in peanut, only three genetic maps have been reported. The first map, comprising 149 SSR loci across 11 linkage groups and spanning 1,294 cM, was developed from an F2 population resulting from a cross between *A. ipaensis* (KG30076) and A. magna (KG30097) (Moretzsohn *et al*., 2009*).*

The other genetic map was constructed with 449 SSR loci using again a F2 population derived from the cross *A. batizocoi* (PI298639) x *A.* batizocoi (PI468327) (Guo *et al*., 2012). Later in the F6 generation a map was constructed on 10 LGs of 461 cM with 798 loci (Shirasawa *et al*., 2013).

I.6.2.2. Genetic maps for tetraploid (AABB) genome

Genetic maps for the AABB-genome in peanuts have been developed using various mapping populations, including F2, BC1F1, BC2F1, BC2F4, and recombinant inbred lines (RILs), and employing marker systems such as AFLP, SRAP, SSR, SNP, and DArT markers (Garcia *et al.*, 1995; Herselman *et al*., 2004; Varshney *et al*., 2009; Ravi *et al*., 2011; Zhou *et al*., 2014; Vishwakarma *et al*., 2016; Shashidhar *et al*., 2017). Initial efforts utilizing RAPD and RFLP markers resulted in the construction of genetic maps spanning 800 cM of genomic distance (Garcia *et al*., 1995). The next genetic map was constructed using AFLP markers, which resulted in the development of a partial map with only 12 AFLP marker loci (Herselman *et al*., 2004).

SSR markers have gained popularity in peanut genetic mapping, leading to the construction of several genetic maps. The first SSR-based map, developed from a Recombinant Inbred Line (RIL) population derived from TAG 24 x ICGV 86031, consisted of 135 SSR loci after screening 1,145 SSR markers (Varshney *et al*., 2009). Subsequently, this map was expanded to include 191 SSR loci across 20 linkage groups, covering 1,785 cM of the genome (Ravi *et a*l., 2011). Most genetic maps have been constructed using RIL populations (Khedikar *et a*l., 2010; Sarvamangala *et al*., 2011). Utilising the NGS-based ddRADseq technique, Zhou *et al*. (2014) provided a well-saturated map with 1685 marker loci, including 1621 SNPs and 64 SSR markers, spanning 1447 cM with an average distance of 0.9 cM. Additionally, the use of DArT and DArTseq genotyping resulted in the development of three genetic maps using F2 populations, comprising 854 loci (ICGV 07368 x ICGV 06420), 1152 loci (ICGV 00350 x ICGV 97045), and 1435 loci (ICGV 06420 x Sun Oleic 95R) (Vishwakarma *et al*., 2016; Shashidhar *et al*., 2017).

The first genetic map for the tetraploid genome of *Arachis* was established using a backcross population (BC1) with the amphidiploid TxAG-6 as the donor parent and *A. hypogaea* cv. Florunner as the recurrent parent (Burow *et al*., 2001). Another map was developed from a synthetic amphidiploid (*A. ipaënsis* × *A. duranensis*) crossed with Fleur 11, resulting in 88 BC1F1 individuals (Fonceka *et al*., 2009).

This population yielded an SSR-based linkage map with 298 markers across 21 linkage groups, covering a total distance of 1,843.7 cM. Additionally, a genetic map spanning 1792 cM was constructed using the recurrent parent Fleur 11 with ISATGR52B (Sambou , 2017). Recently, three more genetic maps were constructed, with two derived from crosses between ICGV 91114 × ISATGR 1212 and ICGV 87846 × ISATGR 265-5A, containing 258 loci (1415.7 cM map length with a map density of 5.5 cM/loci) and 1043 loci (1500.8 cM map length with a map density of 1.4 cM/loci), respectively (Khera *et al*., 2019). The third linkage map utilized a population from crosses between ICGS 76 x ISATGR 278-18, consisting of 114 loci spanning 746.15 cM, with an average inter-marker distance of 6.55 cM (Kumari *et al*., 2020). In summary, various genetic maps for diploid, tetraploid, and integrated genomes have been reviewed in previous studies (Pandey *et al*., 2012; Fonceka *et a*l., 2013; Pandey *et al*., 2014a; Vishwakarma *et al*., 2017; Desmae *et al*., 2019).

I.6.2.3. Integrated genetic maps

Dense genetic linkage maps offer numerous applications in genetics and breeding, including trait mapping, marker-assisted breeding, and map-based cloning (Shirasawa *et a*l., 2013; Pandey *et al*., 2014a). To maximize the mapping of marker loci, integrating data from multiple individual genetic maps into a consensus map is essential. Consensus maps offer several advantages, including mapping numerous marker loci onto a single map, determining marker stability across populations and genomes, and facilitating comparative genomic studies among related species.

Six integrated genetic maps have been developed for peanut, combining data from 2 to 16 mapping populations (Table 1) (Hong *et al*., 2010; Sujay *et al*., 2012.; Gautami *et al*., 2012a, b; Qin *et al*., 2012; Shirasawa *et al*., 2013). These maps provide comprehensive coverage of the peanut genome and offer insights into chromosomal rearrangements and gene duplication. For instance, the first integrated genetic map, based on three RIL populations, contained 175 marker loci across 22 linkage groups, covering 885.4 cM (Hong *et al*., 2010). Subsequent integrated maps expanded marker coverage, with one map comprising 225 SSR loci over a total distance of 1,152.9 cM (Sujay *et al*., 2012). Another integrated map incorporated data from three populations, featuring 293 marker loci distributed across 20 linkage groups, spanning a genome distance of 2,840.8 cM (Gautami *et al*., 2012b). The third integrated map was constructed by Qin *et al*. (2012) based on two mapping populations with 324 marker loci on 21 linkage groups covering a 1,352 cM genome distance.

Efforts to enhance marker density and number have led to the development of a reference consensus map, integrating data from 11 mapping populations, including one BC population and 10 RIL populations (Gautami *et al*., 2012a). This reference map comprises 897 marker loci, predominantly SSRs, mapped onto 20 linkage groups, covering a total distance of 3,607.97 cM with an average map density of 3.94 cM. Moreover, the INT map, constructed using 16 populations from diploid and tetraploid species, exhibits the highest resolution of 0.7 cM/locus and a genetic distance of 2651 cM with 3693 loci on 20 linkage groups (Shirasawa *et a*l., 2013).

Table 1.Summary of Integrated maps

Population	LGs	Total map distance (cM)	Marker loci	Marker	References
3 Population	22	885.4	175	SSR	(Hong *et al*., 2010)
2 Population	20	1,152.90	225	SSR	(Sujay *et al*., 2012).
3 Population	20	2,840.80	293	SSR	(Gautami et al., 2012b)
2 Population	21	1,352	324	SSR	(Qin *et al*., 2012)
11 Population	20	3,607.97	897	SSR, CAPS	(Gautami *et al*., 2012a)
16 Population	20	3693	2651	SSR, TE	(Shirasawa *et al*., 2013)

I.6.3. Statistical methods and limitation of QTL mapping

Statistical methods for family-based mapping include 1)single-marker analysis (SMA)(Sarvamangala *et al.*, 2011; Wilson *et al*., 2017) used for initial QTL mapping in biparental populations. It identifies QTL based on the difference between the average phenotypes of different genotype groups without using information about genetic distances in the linkage map. 2) Interval mapping (IM)(Lander and Botstein, 1989) is based on maximum-likelihood parameter estimation and regression. It efficiently estimates the effect and position of a QTL within two flanking markers.3) Composite interval mapping (CIM) is used to overcome the limitation of IM method, which is less accurate for analysing multiple QTL simultaneously. 4.Inclusive composite interval mapping (ICIM)(Singh and Singh, 2015; Xu *et al.*, 2017) (4) and multiple interval mapping (MIM) (5) are an extension of interval mapping to multiple QTL, tends to be more powerful and precise than CIM in identifying QTL and allows the simultaneous estimation of multiple QTL with epistasis**.**

A large number of software packages are available for parental mapping including PLABQTL(Sarvamangala *et al., 2011)*, Win QTLcartographer (2.5) (Pandey *et al*., 2014b; Huang *et al*., 2015; Chen *et al.*, 2016; Khedikar *et al*., 2017;Guo *et al*., 2021), IciMapping (Liu *et al*., 2020; Bomireddy *et al*., 2022), R/Qtl (Fonceka *et al*., 2012a) ,Mapchart(Pandey *et al*., 2020a), QTLNetwork-2.0 (Pandey *et al*., 2014b; Nian *et al.*, 2019).

The accuracy of any QTL mapping technique depends on several factors, including the statistical method's capacity to locate and estimate the genetic effect of the QTL, the type and size of the mapping population, the genetic and heritability of the trait, the number and contribution of each QTL to the total variance, their interactions, their distribution over the genome, the number and distance between consecutive markers, and the percentage of *cis*-regulatory elements (Asíns, 2002) and in such case, the ploidy and meiotic behaviour in peanut. Along with these accuracy criteria, QTL analysis has limitations like the other techniques. Some of these limitations include the inability to detect all loci, the number of QTL detected, their precise position, and their effects are subject to statistical error. Major QTL are often missed and epistatic effects and QTL environmental interactions (Würschum, 2012) are found in some cases. QTL mapping is often time-consuming, requires in-depth knowledge about the function and genomics of the trait of interest and has high cost of genotyping and phenotyping.

The development of the mapping population, the limited number of recombination events, large QTL size and low mapping resolution (>10 cM) are some of the challenges in biparental QTL mapping. In much cases, more experiments are needed to confirm the results of QTL mapped (Pascual *et al.*, 2014; Huang *et al*., 2015a).However, by using consistent QTL that have been mapped, it is expected that the next-generation crop varieties could be developed with enhanced quality traits, better yield and disease resistance.

I.6.4. Quantitative trait loci mapping in peanut

Quantitative trait loci (QTL) mapping has emerged as a crucial tool in peanut genetics and breeding, despite the narrow genetic diversity and the segmental tetraploid nature of cultivated peanut. A wide range of quantitative or metric traits in peanut, including those related to yield and yield components traits, flowering, seed dormancy, quality and nutritional traits, resistance to viral, bacterial, and fungal diseases and physiological traits have been subjected to QTL mapping (as reviewed in: Pandey *et al*., 2012; Vishwakarma *et al*., 2017; Desmae *et al*., 2019; Pandey *et al*., 2020b; Kassie *et al*., 2023).

In the following later sections, QTL for yield related traits and seed quality traits have been discussed in detail. and the QTL mapped to date have been reviewed by Kassie *et al*., (2023). In subsequent sections, the focus narrows to a detailed discussion of QTL associated with yield-related traits and seed quality traits.

I.6.4.1. Mapping QTL for seed quality traits

Important Key quality traits in peanut, such as oil, protein, and sugar content, along with fatty acid (FA), amino acid, and carbohydrate composition, can be assessed through biochemical analysis of the peanut kernel. Among these, the concentration of oleic acid is particularly significant due to its impact on the shelf life of peanut products and its health benefits (Nawade *et al*., 2018). In particular, the concentration of oleic acid is one of the most important quality traits because it can increase the shelf life of peanut products and is beneficial for human health (Vassiliou *et al.,* 2009; Pandey *et al*., 2014b).

Several studies have reported QTL mapping for traits related to oil and protein content, as well as fatty acid composition in peanuts, using biparental mapping populations such as F2, RIL, and BC segregating populations (Sarvamangala *et al*., 2011; Pandey *et al*., 2014b; Wang *et al*., 2015; Huang *et al.*, 2015b; Shashidhar *et al*., 2017; Wilson *et al*., 2017; Hu *et al*., 2018; Nian *et al*., 2019; Liu *et al*., 2020; Sun *et al*., 2021; Guo *et al*., 2021). For instance, a mapping population of 146 recombinant inbred lines (RILs) generated from a cross of TG26 x GPBD4 revealed QTL for protein, oil, oleic, and linoleic acid content, as well as for the oleic acid to linoleic acid ratio (Sarvamangala *et al*., 2011). As the authors have mentioned, GPBD4 has early maturity, high yield, high pod growth rate, desirable pod and kernel features, high oil and protein content, and an optimum oleic/linoleic acid (O/L) ratio, whereas TG26 is a semi-dwarf, erect cultivar with a high linoleic acid content. Although the genetic map has low coverage (45 SSR markers on 8 linkage groups), the authors reported 17 QTLs on 4 genomic regions, including 2 major QTLs for protein content.

Similarly, major QTLs for oil content, oleic acid, linoleic acid, and the ratio of oleic acid to linoleic acid were mapped using two genetic maps developed from RIL populations derived from the crosses between Sun Oleic 97R and NC94022 and between Tifrunner and GT-C20 (Pandey et al., 2014b). Two major QTLs for oil content on chromosomes A05 and A08 and 11 major QTL for oleic acid, linoleic acid, and the ratio of oleic acid to linoleic acid on the homoeologous chromosomes A09 and B09 were first mapped.

Furthermore, utilising the same mapping populations developed by Pandey *et al*. (2014b), 16 major QTLs were identified on B04 and A09/B09 for palmitic acid, stearic acid, arachidic acid, gadoleic acid, behenic acid, and lignoceric acid content (Wang *et al*., 2015). Additionally, 1 major QTL for oil content on chromosome B03 was detected, explaining 14.36% of the phenotypic variance (Huang *et al., 2015b)*.

Notably, 23 major QTL on 11 genomic regions were identified, explaining 10.4% to 41% of the phenotypic variance for oil content and fatty acid composition (Shasidhar *et al*., 2017). In a high-resolution genetic map from a cross between high and normal oleic cultivars, 29 major QTL were mapped for oleic and linoleic acid content, as well as the oleic to linoleic acid ratio, on chromosomes A03 and A09/B09 (Hu *et al*., 2018).

Recent studies have further elucidated QTL associated with oil content, revealing 14 major QTL on A05, A06, A08, B06, and B10, explaining up to 27.19% PEV, from the three mapping populations derived from Xuhual13 and Zhonghua6, Yuhual15 and W1205, and Zhonghua10 and ICG12625 (Liu *et al.*, 2020; Sun *et al*., 2021; Guo *et al*., 2021). Major QTL associated with protein stearic acid, behenic acid, and arachidic acid contents were mapped on chromosomes A05, A06, and A08 (Sun *et al*., 2021). The locations of the seed quality QTL on chromosomes were located in Fig. 9.

We performed a comparative QTL analysis using data from the studies above in order to gain more insight into the genome-wide distribution of kernel-quality QTL and to document the most consistent ones for future use in marker-assisted breeding. The map location of the QTL is presented in Fig. 9, and the detailed data of all the 413 quality-related QTL that have been mapped to date are found in (Kassie *et a*l., 2023). We found that QTLs for the quality trait are mainly clustered on chromosomes A05, A08, and A09 for the A genome, and B04, B08, and B09 for the B genome. For instance, many QTL for oil and protein content as well as fatty acid compositions (arachidic, arachidonic, behenic, stearic, palmitic, linoleic and oleic) colocalized on chromosome A05 and were consistent among environments in Fig. 9. Furthermore, QTLs for oleic acid, linoleic acid, and the oleic/linoleic ratio from different studies were found in common genomic regions on chromosomes A05, A08, B04 and B09. In chromosome B09, the common QTL are closely linked to markers ahFAD2B and SNP markers, Marker2575339 or Marker239598 (Pandey *et al.*, 2014b; Hu *et al*., 2018). The AhFAD2B QTL, on chromosome B09, explained up to 57 % of phenotypic variation of oleic acid or linoleic acid content.

Similarly, the AhFAD2A and Marker4391589 or Marker4463600 on chromosome A09, are common among studies and explained up to 29 % of phenotypic variation (Pandey *et al*., 2014b; Hu *et al*., 2018). Additionally, AhMXZ190701 was discovered to be tightly linked to a major and stable QTL on A08 for oil content (Pandey *et al*., 2014b; Liu *et al*., 2020). These consistent markers, AhMXZ190701, ahFAD2B, ahFAD2A, Marker2575339 or Marker2379598, have been used for QTL validation and MAS of quality traits (Zhao *et al*., 2016; Liu *et al*., 2020).

Likewise, several QTLs for arachidic, behenic, stearic, palmitic, linoleic and oleic acid and oil content, mapped in three studies, were linked to the marker RN34A10 on chromosome A7 (Fig. 9) (Pandey *et al.*, 2014b; Wang *et al.*, 2015). Furthermore, consistent QTL among traits and environments were also reported. A QTL mapping study on four environments, among the 110 QTL related to nine quality traits, 36 pleiotropic QTL were associated with two or more traits and showed consistent effects in more than one environment (Sun *et al*., 2021). As a conclusion, of the total main effect QTL that have been discovered to date for oil content, four stable QTL that were tightly linked to SSR and SNP markers were located on the 15.0- to 21.7-Mb, 6.4- to 10.9-Mb,, 99.15- to 108.29-Mb, and 6.3- to 7.8-Mb regions of A05 (Sarvamangala *et al*., 2011; Pandey *et al*., 2014b; Sun *et al.*, 2021; Guo *et al*., 2021), respectively. A stable QTL linked to SNP markers bin1572-bin1581, in the interval of 0.5 cM spanned by A05, corresponding to a 6.3–7.8 Mb physical region of chromosome A05 (Sun *et al.*, 2021), is the same as that reported (Pandey *et a*l., 2014b) for the flanked markers GM1878 and GM1890, which were mapped to the approximately 6.4-10.9 Mb region of A05.

The two stable and major QTL with 13.51-22.59% PVE linked with SSR markers (Ai06B29452), 136.42-137.05 Mb, and located on with 9.18-12.55% PVE, linked AGGS2133-1 were identified on B06 and B10 (Guo *et al*., 2021).The two stable QTL tightly linked to SNP markers AhMXZ190701- AhEXZ283046 (Liu *et al.*, 2020), and bin2782 and bin2787 (Sun *et al*., 2021), were closely located on a 39.9-43.8 Mb and a 37.0-38.2 Mb on chromosome A08, respectively. This suggests the reproducibility of the QTL in different studies. Generally, 4 stable and major QTL referring to oil content were located from 6.3 to 108.29 Mb on A05, 2 QTL37.0-39.9 Mb on A08 and the other 2 on A06, and A10. These consistent QTL within and across studies can be used in breeding special-purpose peanut cultivars. However, some QTL need to be validated with fine mapping considering their position on chromosomes that differed from different studies, probably due to the genetic material, large QTL intervals and statistic imprecisions.

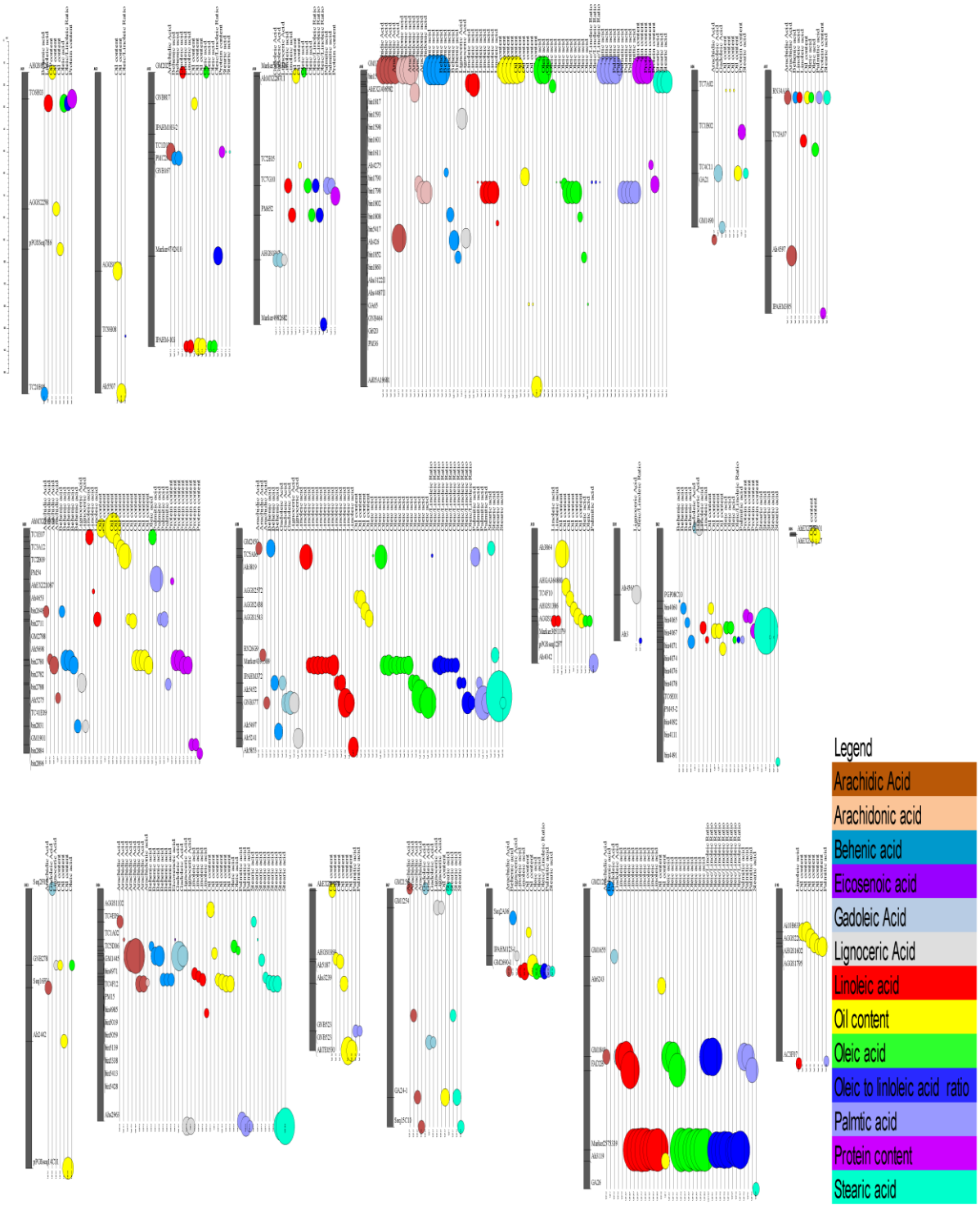


Fig. 9. Schematic map of known QTL related to quality traits in peanut (Sarvamangala *et al*., 2011; Pandey *et al*., 2014b; Wang *et al*., 2015; Wilson *et al.*, 2017; Shasidhar *et al*., 2017; Hu *et al*., 2018; Liu *et al.*, 2020; Sun *et al*., 2021; Guo *et al*., 2021).

I.6.4.2. QTL mapping for yield component traits

In the peanut breeding program, enhancing pod yield is important to meet the food demands of a growing global population. QTL mapping is a key tool utilised to achieve this objective. This section discusses the identification of QTL linked to yield component traits using SSR and SNP markers, primarily with F2 and recombinant inbred lines (RILs). Furthermore, the identification of QTL using DArT markers in addition to SSR and SNP markers using backcross populations will be discussed in detail in the following section.

A significant number of main effect QTL (approximately 292) (Table.2) have been mapped for traits such as pod weight, hundred-seed weight, pod and seed length, width, and pod number using F2 (Huang *et al*., 2015b; Chen *et al*., 2016) and RIL (Chen *et al*., 2017; Khedikar *et al*., 2017; Luo *et al*., 2017, 2018; Liang *et al*., 2018; Chavarro *et al*., 2020) mapping populations. For instance, 44 QTL were identified for 100 pod weight, explaining up to 38.15 % of the phenotypic variance on A05, A06, A07, A08, A09, B03, B04, B05, A07 , B08 and B10 (Huang *et al*., 2015b; Chen *et al*., 2017; Luo *et al*., 2017, 2018). Moreover, around 35 QTL were reported for 100 seed weight on A02, A03, A04, A05, A06, A07, A08, B02, B03, B04, B05, B06 and B08, explaining 5.68 to 35. 9% of phenotypic variance (Huang *et al*., 2015; Chen *et al*., 2017; Khedikar *et al*., 2017; Liang *et al*., 2018). Additionally, 6 SSR markers were found to be associated with major, consistent, and stable QTL for pod length, pod width, and 100 pod weight on chromosomes A05, A07, A09, and B05 (Luo *et al*., 2017, 2018).

Beyond seed and pod traits, QTL for flowering, plant height, and fresh seed dormancy were also identified. For flowering, 30 QTL were reported with varying phenotypic variance explained ranging from 1.15 to 21.82 % (Khedikar *et al*., 2017; Wang *et al*., 2020). In the case of plant height, 71 main affect QTL were identified (Faye *et al*., 2015; Huang *et al*., 2015; Chen *et al.*, 2017; Khedikar *et al.*, 2017; Lv *et al*., 2018), up to 26.27% of the phenotypic variance. Co-localization of QTL for various traits was observed, indicating potential pleiotropic effects. Several QTL associated with plant architecture, such as growth habit and plant height, co-localized with those associated with flowering (Lv *et al*., 2018; Li *et al*., 2019). As well, the QTL of yield components, such as hundred pod weight, pod weight, and pod length (Luo *et al*., 2017, 2018) were co-localized on chromosomes A05 and A07. Overall, QTL mapped related to agro-morphological traits include QTL related to plant architecture, flowering, fresh seed or seed dormancy, and yield component traits.

To conclude, for pod and seed-related traits such as pod width, pod length, and 100 pod weight, six stable, consistent and major genomic regions have been reported on A05, A07, A09 and B05 (Luo *et al*., 2017, 2018) with co-located genomic regions.

The co-localized interval on A07 was located on 5.7 cM (0.06-1.54 Mb) and harboured the major QTL for pod length, pod width, and 100-pod weight ranging 17.97-43.62 % of phenotypic variations by the flanking markers AhTE0025 and AHGS1836. On A05,the co-localized interval was located on 1.3 cM (99.50-99.78 Mb) explained 17.97-43.62 % of phenotypic variations by the flanking markers Ad05A20262 and AHGA160418 (Luo *et al.*, 2018). For these traits, three more major QTLs co-located in a about 2.47 Mb genomic region of the A05 with (13.75-26.68% PVE) by the flanking markers A05A1430-A05A1601 traits (Luo *et al*., 2017). Moreover, three major QTL commonly detected for pod length and seed length on A05 with up to 26.11% PEV (Chen *et al*., 2016). Many major and stable QTL detected on A05 in different studies for oil content and seed and pod related traits suggest it may harbour important genes controlling these traits, which can be used in marker assisted breeding.

As for the pod shape, 10 QTL for pod beak and 7 for pod constriction were reported, explaining 17.4% of the phenotypic variance on chromosomes A02, A07, A08, A09, A10, B01, B02, and B06 Fonceka *et al*. (2012a) using advanced population. In a subsequent study, Patil *et al*. (2018) employed a recombinant inbred line (RIL) population and identified a significant QTL for pod constriction on chromosomes B05 and B07. Mondal and Badigannavar (2019) further expanded this knowledge, pinpointing four QTLs for pod beak and two for pod constriction on chromosomes A08, B03, B07, and B08, with a cumulative phenotypic variation explained (PVE) of 8.89%. Recently, Zhang *et al*. (2023) reported on the genetic architecture of these traits, uncovering 10 QTLs for pod beak and 3 for pod constriction on chromosomes A02, A03, A05, and B06. These collective findings contribute to a deeper understanding of the genetic basis underlying pod morphology in peanut. These authors reported one stable and major QTL region with pleiotropic effects was mapped on A02 with 0-4.473 cM genetic distance. Overall, QTL mapping has revealed valuable genomic regions associated with agro-morphological traits, offering opportunities for marker-assisted breeding to enhance peanut yield related traits.

I.6.4.3. QTL Mapping involving synthetic tetraploid wild derivatives

The employment of synthetic tetraploid wild derivatives in peanut breeding has proven indispensable for the genetic mapping of various traits such as disease resistance, drought tolerance, agronomic characteristics, and oil quality. Advanced backcross (AB) mapping populations, including BC2F1, BC3F1, BC2F3, BC4F3, BC3F2, and BC3F6, have been instrumental in this endeavour, genotyped using markers like RFLP, DArT, and SSR. For instance, Burow *et al*. (2014) constructed a linkage map utilising a backcross population involving 'Florunner' and synthetic 'TxAG-6', identifying seven QTL associated with root-knot nematode resistance. Wilson *et al*. (2017) reported 29 QTL linked to oil content and fatty acid composition using BC3F6 generations of the same population, genotyped with SSR markers, with two major and stable QTL explaining up to 31% of phenotypic variance.

Furthermore, Fonceka *et al*. (2012a) developed an AB population from a cross between the cultivated parent 'Fleur 11' and an amphidiploid AiAd, leading to the identification of 82 QTL associated with traits related to plant architecture, domestication, and yield components under water-limited and well-watered conditions. Using the same resource, Fonceka *et al*. (2012b) identified 42 QTL for four morphological traits using 122 chromosome segment substitution lines. Using the same recurrent parent, Sambou (2017) also detected 38 QTL using the BC2F4 mapping population derived from the cross of the recurrent parent Fleur 11 with a different male parent (ISATR52B), underlying traits such as days to 50% flowering, plant architecture, yield-related characteristics, pod, and seed morphology. According to the aforementioned studies, it was observed that approximately half of the identified QTL with positive effects were attributed to alleles inherited from the wild parent. This underscores the significance of peanut wild relatives as a reservoir of beneficial alleles for breeding purposes.

Similarly, Khera *et al*. (2019) identified 15 QTL in ICGV 8764 x ISATGR 265-5A and 35 QTL in ICGV 91114 x ISATGR 1212, with favourable alleles derived from both wild and recurrent parents. Recently, Kumari *et al*. (2020) utilized an advanced backcross mapping population derived from a cross between ICGS 76 and synthetic amphidiploid ISATGR 278-18, identifying 24 QTL linked to plant height, shelling percentage, total pod weight, and disease resistance.

In summary, the use of synthetic tetraploid wild derivatives in peanut breeding has facilitated alien chromatin introgression, eased genetic and meiotic analyses, and enabled the detection of QTL for numerous economically important traits.

Despite challenges such as hybrid fertility and linkage drag, these efforts have resulted in the identification of QTL associated with critical traits such as drought tolerance, disease resistance, yield, and seed quality, thereby advancing peanut breeding efforts.

Table 2. Main effect QTL reviewed for yield related traits of peanut *(*Kassie *et al*., 2023)

Traits studied	QTL identified	Phenotypic variance explained (%)
Plant height	77	0.01–26.7
Hundred pod weight	48	3.33–38.15
Days to flowering	31	1.15–21.82
Pod weight	20	7.7–29.7
Pod length	52	1.25–26.46
Pod width	54	5.1–43.63
Seed length	32	3.03–20.8
Seed width	33	2.21–23.7
Harvest index	15	11.0–18.1
Hundred seed weight	42	5.68–35.9
Haulm weight	11	2.9–33.36
Total biomass	15	4.34–22.39
Growth habit	48	4.55–27.14

I.6.5. Marker assisted selection

Phenotypic selection in plant breeding is a traditional method that involves selecting plants based on their phenotypic traits (Watson *et al.*, 2019; Hasan *et al*., 2021. This process is labor-intensive, costly, and time-consuming due to the need for multiple cycles of backcrossing and the difficulty of introgressing several traits into a single parent. Furthermore, phenotypic selection is less efficient for quantitative traits that are often under selection. Marker-assisted selection (MAS) is an alternative approach that utilises molecular linked to genes or QTLs associated with target traits and use these markers to screen and select individuals in breeding populations (Vishwakarma *et al*., 2022; Kumari *et al*., 2024). The success of MAS depends on several factors, including the degree of association between the molecular marker and the target gene, the number of individuals that can be analysed, the genetic background into which the target gene has to be transferred, linkage disequilibrium in the plant population to be selected,

and a known linkage phase between the marker and the target gene (Collard and Mackill, 2008; Vishwakarma *et al*., 2022):. The efficiency of MAS can be increased by using markers flanking the target gene instead of a single linked marker (Song et al., 2023).

According to the research by Hospital *et al*. (1992) and Vishwakarma *et al*. (2022) suggested that the moderate distance between a marker and an interesting trait for efficient marker-assisted selection should be within 5 cM. Mohan *et al*. (1997) proposed that marker(s) should co-segregate with the desired trait or be closely linked, with a distance of 1 cM or less. The efficiency of MAS is enhanced and may be more efficient than traditional selection under the following circumstances, as defined by Vishwakarma *et al*. (2022): 1) the trait under selection has a low heritability; 2) the presence of tight linkage between QTL and marker (<5cM); 3) in earlier generations of selection prior to fixation of alleles at or near marker loci and recombinational erosion of marker QTL associations; 4) when large sample sizes for mapping and selecting QTL are used to improve estimates of QTL alleles. Markers very closely linked to the target genes or even located within the gene can greatly enhance the use of MAS in advanced generations, where the linkage disequilibrium becomes smaller. Different molecular approaches are used in MAS, including marker-assisted backcrossing (MABC), marker-assisted QTL or gene pyramiding (MAQP), and marker-assisted recurrent selection (MARS).

I.6.5.1. Marker-assisted backcrossing (MABC)

Marker-assisted backcrossing (MABC) is a pivotal technique in modern plant breeding aimed at integrating specific genes or quantitative trait loci (QTLs) from a donor parent into a recurrent parent (RP) (Singh and Singh, 2015; Kumawat *et al.*, 2020; Vishwakarma *et al*., 2022). It is invaluable when the recurrent parent lacks essential genes for traits like disease resistance, yield, or quality. By iteratively crossing the donor with the recurrent parent and selecting for target genes or QTLs associated with desired traits, it facilitates the introduction of desirable characteristics into elite cultivars. The primary objectives of MABC encompass several crucial aspects: first, the transfer of desired traits, such as disease resistance or high yield, from the donor plant to the recurrent parent through successive crosses and selection for the target gene or QTL. Second, the recovery of the genetic background of the recurrent parent to ensure progeny closely resemble it, except for carrying the desired trait. Third, the removal or minimization of unwanted donor genetic material, which may be linked to undesirable traits, is a phenomenon known as linkage drag (Singh and Singh, 2015; Kumawat *et al.*, 2020; Vishwakarma *et al*., 2022).

Molecular markers play a pivotal role in assisting selection during MABC to achieve these objectives. Foreground selection involves utilising molecular markers linked to the target gene or QTL for indirect selection, enabling breeders to identify plants carrying both the marker and the target gene/QTL without extensive phenotypic evaluation. Background selection employs codominant molecular markers distributed throughout the genome to track the proportion of recurrent parent alleles, aiding in the progressive recovery of the recurrent parent genome. Recombinant selection utilises codominant markers located on flanking regions of the target gene/QTL to identify recombinant individuals, thereby removing undesirable donor segments while retaining the desired trait.

Illustratively, utilising marker-assisted backcrossing, rust and leaf spot (LIS) resistant, and high oleic acid content peanut lines were successfully developed. In efforts to develop rust and leaf spot (LIS) resistance lines, crosses were conducted involving the variety GPBD-4 with three susceptible varieties: ICGV 9114, JL 24, and TAG 24 (Varshney *et al*., 2014). This breeding approach was guided by specific SSR markers linked to major QTLs, including IPAHM103, GM2079, GM1536, and GM2301, associated with rust and LIS resistance in the GPBD-4 cultivar. Kolekar *et al*. (2017 furthered this endeavour by crossing GPBD-4 with TMV 2, utilizing disease resistance markers such as GM2009, GM2079, GM2301, GM1839, and IPAHM103, resulting in the development of two rust and LIS resistance lines, TMG-29 and TMG-46. Additionally, Shasidhar *et al.* (2020) employed several SSR markers, including IPAHM103, GM1536, GM2301, GM2079, SEQ8D09, and GM1009, for the selection of rust and LIS resistance lines in populations derived from GJG-9, GG-20, and GJGHPS-1 with GPBD-4.

A variety of high oleic acid peanut lines were developed through strategic breeding approaches. Twenty-four high-oleic introgressed lines were achieved by crossing 'Yuhua 15', 'Yuanza 9102', 'Yuhua 9326', and 'Yuhua 9327' with Kainong 176 and KN 176, known for their high oleic acid content (Huang *et al*., 2019). Furthermore, 46 BC3F4 and 41 BC2F4 high oleic acid lines were reported, originating from crosses involving parent lines GJG 9, GG 29, and GJGHPS, with the high oleic acid variety 'Sun Oleic 95R' (Shasidhar *et al*., 2020). Notably, the recent development of the high-oleic-acid line 'YH61' stands as a testament to the continuous advancements in peanut breeding methodologies. 'YH61' was created through the crossing of 'huayu22' with the high-oleic-acid donor 'KN176', followed by four generations of backcrossing (Tang *et al*., 2022).

These efforts underscore the efficacy of marker-assisted breeding techniques in enhancing the oleic acid content of peanut varieties, thereby contributing to the diversification and improvement of peanut cultivars with enhanced nutritional profiles and commercial value.

I.6.5.2. Marker-assisted recurrent selection

Marker-assisted recurrent selection (MARS) is a breeding technique that utilises molecular markers to assist in the selection of plant genotypes carrying specific genes or quantitative trait loci (QTL) of interest (Collard *et al.*, 2005). This approach has gained popularity in plant breeding due to its ability to improve precision, reduce phenotyping costs, decrease cycle time, and generate superior genotypes within a population. The process of MARS begins by identifying molecular markers that are closely linked to the target genes or QTLs. Once the tightly linked markers have been identified, breeders can use them as diagnostic tools to identify plants carrying the desired QTLs. After identifying plants carrying the target QTLs, controlled pollination is performed to create progeny lines with a favourable combination of QTLs from both parents. By strategically crossing selected individuals, breeders aim to enhance the expression of desirable traits and accumulate multiple favourable QTLs within a single genotype (Gokidi *et al*., 2016).

I.6.5.3. Marker assisted QTL Pyramiding

Marker-assisted QTL pyramiding is a valuable technique in plant breeding, involving the simultaneous integration of multiple genes or quantitative trait loci (QTL) from different parental lines into a single genotype using molecular markers (Xu *et al*., 2012; Chukwu *et al*., 2019; Gautam *et al*., 2020). The process of marker-assisted gene pyramiding entails two essential steps: gene fixation and pyramiding. Gene fixation ensures that target genes are in a homozygous state, while pyramiding accumulates all target genes into a single genotype, enabling the development of lines with multiple beneficial traits (Servin *et al*., 2004).

This methodology has been successfully applied in peanut breeding to develop genotypes resistant to nematodes and rich in high oleic acid content (Chu *et al*., 2011). In this process, the nematode-resistant cultivar 'Tifguard' was crossed with two high-oleic parents, Georgia-02C and Florida-07. The resulting BC3F2 plant progenies underwent confirmation for homozygosity of the target alleles, ensuring the presence of both nematode resistance and high oleic acid content. Gene pyramiding has also been employed in peanut to enhance aflatoxin resistance and yield components (Jin *et al*., 2023).

CHAPTER II. MATERIAL AND METHODS

II.1. Materials

II.1.1. Study area

Experiments were conducted in four different locations across two countries: Maroua in 2019, Mbalmayo and Bafia in 2021 in Cameroon (Fig. 10), and Nioro in 2021 in Senegal, all under rain-fed conditions. The characteristics of each environment are detailed in Table 3. These locations were selected to capture environmental diversity, considering various criteria such as ecology (including climate and vegetation), the traditional practice of peanut cultivation, and economic factors.

Maroua is a tropical savanna environment with sandy clay vertisol soils (Kenga *et al*., 2005). This environment experiences an equatorial climate of the Sudano-Sahelian type, characterised by high temperatures averaging around 35°C and an annual rainfall of approximately 800 mm. Bafia is one of the main areas of peanut production in Cameroon. It is located in tropical savanna and has yellow vertisol soil (Temga *et al*., 2021) and an equatorial climate of the Sudano-Guinean type with an average temperature of 25.1°C and annual rainfall of 1500 mm. Mbalmayo is located in the tropical forest of Cameroon and has ocher vertisol soil (Temga *et al*., 2021) with a bimodal humid-forest rainfall climate with an average temperature of 26.5°C and rainfall of 2402.8 mm.

Nioro is located in the South of the Senegalese peanut basin and has a Sahelian semi-humid ecology with Deck Dior soil, a leached ferruginous tropical soil (Bogie *et al*., 2018). The annual rainfall is 758 mm with an average temperature of 30 °C. The fields at Bafia and Mbalmayo were one-year fallow land after maize cultivation by farmers and were cleared and plowed for the study. The previous crop at Nioro was millet. The experiments in Bafia and Mbalmayo were conducted during one of the two rainy seasons from April-July, while Maroua was conducted during rainy seasons from July- October, at the Research Institute for Agricultural Development. Whereas the Nioro experiment was done during the rainfall season between July and October, at the Research Station of the National Agricultural Research Centre.

Table 3. Characteristics of the field environments (Kenga *et al*., 2005; Bogie *et al*., 2018; Temga *et al*., 2021).

Characteristics	Maroua	Bafia	Mbalmayo	Nioro
Country	Cameron	Cameron	Cameron	Senegal
Location	Maroua	Bafia	Mbalmayo	Nioro
Ecology type	Tropical savanna	Tropical savanna	Tropical forest	Sahelian
Climate Type Name	Sudano-Sahelian equatorial	Sudano-Guinean equatorial	Humid-forest bimodal rainfall	Sahelian semi-humid
Temperature (°C)	35	25.1	26.5	30
Rainfall (mm)	800	1500	2402.8	758
Soil type	Vertisol	Yellow vertisol	Ocher vertisol	Deck Dior
Previous crop	Millet and sorghum	Maize	Maize	Millet
Experimental period	July - October	April-July	April-July	July - October

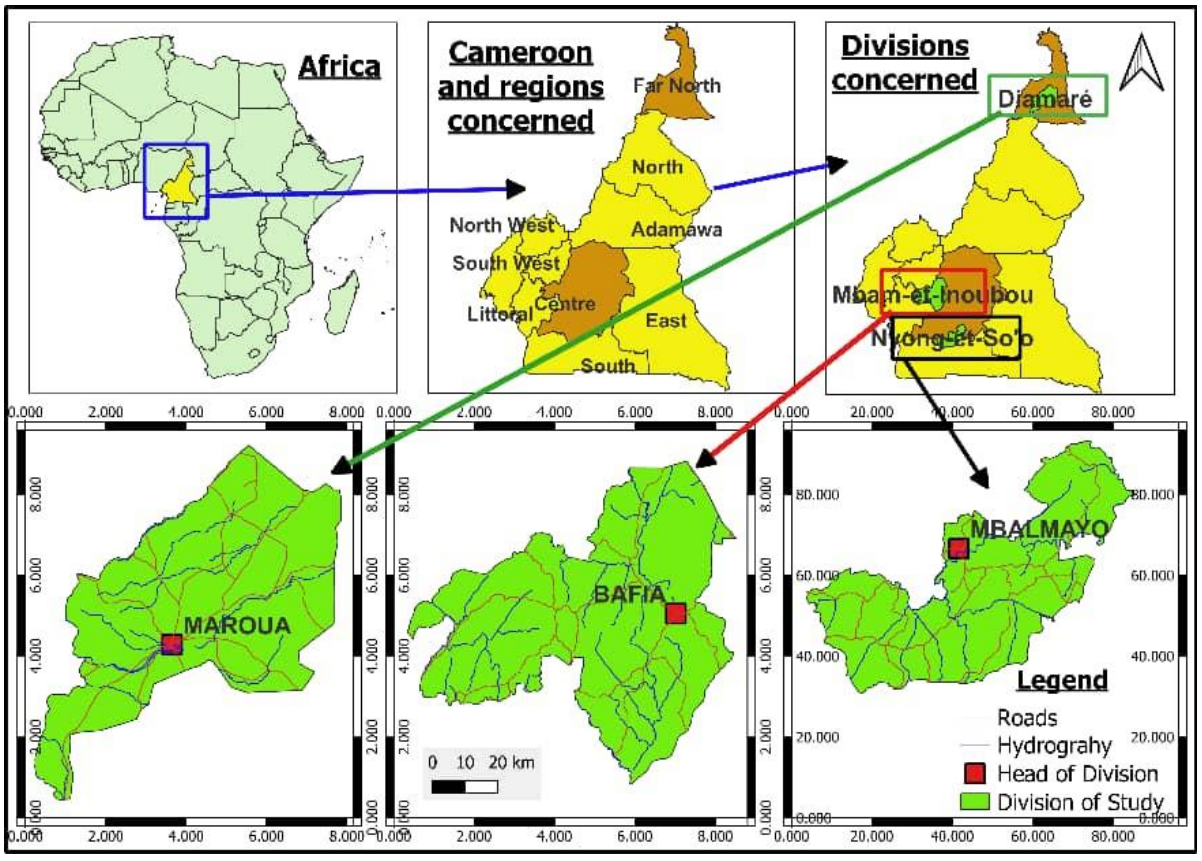


Fig.10. Location of study area in Cameroon (World Resource Institute, 2020).

II.1.2. Plant materials

Two distinct genetic materials were used in this study as in Supplementary Tables 1 and 2 of Kassie *et al*.(2024): an interspecific advanced backcross QTL (AB-QTL) population of 133 genotypes and a core collection of 300 African cultivars. The AB-QTL population of 133 BC_2F_4 derivatives was developed from the initial interspecific cross using Fleur11 as recurrent cultivated parent and the wild synthetic tetraploid 'ISATGR 278-18'(Nguepjop *et al*., 2016). The cultivated parent used, Fleur 11, is an elite Spanish-type variety, widely cultivated in West Africa. The wild parent, ISATGR 278-18 is derived from a cross between *A. batizocoi* ICG 13160 (GKBSPSc 30082, PI 468328) and *A. duranensis* ICG 8138 (GKP 10038, PI 262133) (Mallikarjuna *et al*., 2010). The CS16 variety and the cultivated parent Fleur11, also included. The African core collection of 300 cultivars was defined based on breeder's knowledge and on diversity data from a collection of 1050 accessions (breeding lines and landraces) held by 10 breeding programs in East, Southern and West Africa (Conde *et al*., 2023).

II.2. Methods

II.2.1. Field experimental design

The same experimental design, following common agricultural practices from sowing to harvest, was implemented in each of the four environments. An alpha-lattice design employed by Mohammed *et al.* (2019), Pankaj *et al*. (2022) and Bedru *et al*. (2024), was utilised with 3 replications and 10 blocks per replication (Fig. 11) . Each plot comprised rows spanning 3 meters, where 10 plants of the same genotype were sown per row. Plants were spaced 30 cm apart within the same row and 50 cm apar between adjacent rows or different genotypes according to (IBPGR and ICRISAT (1992), Upadhyaya & Gowda (2009) and Fonceka *et al*. (2012a). Prior to planting, the seeds were treated with benomyl 10% and carbofuran 20%) to safeguard against parasitic attacks, with manual sowing of one seed per hill at a depth of 4 cm.

Following standard cultural practices, a mineral fertilizer (6-20-10) was applied at a rate of 150 kg/ha 20 days after sowing (Fonceka *et al*., 2012a). Weed control was managed manually during vegetative development. Harvesting was conducted at 95 days post-sowing, followed by a one-month period of free-air drying as per groundnut descriptors IBPGR and ICRISAT (1992) and Upadhyaya & Gowda (2009) . Upon completion of the pod-drying stage, pods from each plant were separated from the haulms, stored, and subsequently dehulled.

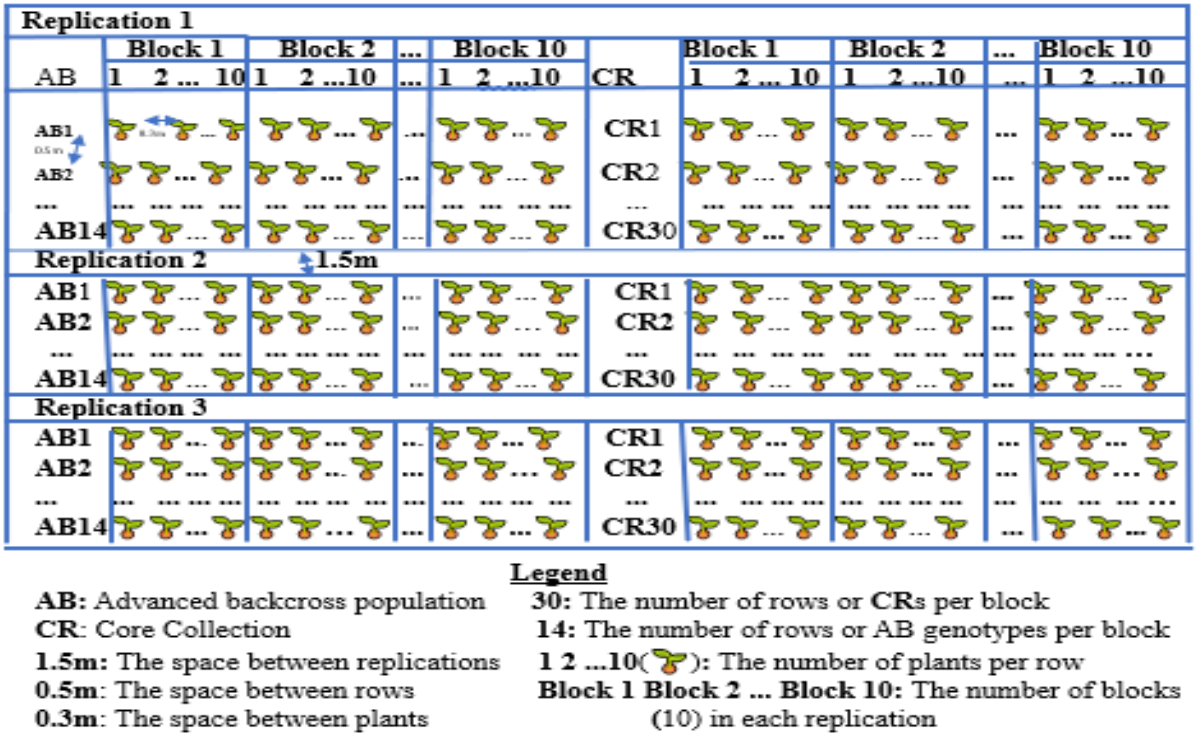


Fig. 11. Field layout of alpha Lattice experimental design for advanced backcross population (AB) and core collections (CR)

II.2.2. Screening peanut core-collection and interspecific population for quality traits using NIRS

II.2.2.1. Whole seed sample preparation

Whole seeds from the pods of the three agronomic replicates of each genotype were combined into specific samples, stored in biodegradable plastic bags, and labelled according to their respective names and environments. Therefore, intact seeds for each sample used in near-infrared (NIR) analysis were obtained from pods of 25-30 harvested plants of each genotype.

Out of the expected 699 samples, we discarded 21 that had fewer than 100 seeds: 3 from Bafia, 9 from Mbalmayo, and 9 from the core collection. Finally, a total of 680 samples of intact seeds were grouped into four sets based on genetics and environments: one set comprising 291 samples from African cultivars in Nioro, and three sets of interspecific genotypes (130 samples for Bafia, 124 samples for Mbalmayo, and 135 samples for Nioro, which included the 133 genotypes, the CS16 variety, and the cultivated parent Fleur11, commonly used as check varieties in Nioro). All samples sealed in hermetic plastic bags were conveyed to the laboratory and kept at ambient temperature prior to spectra acquisition.

II.2.2.2. NIR spectra acquisition

The NIR spectra acquisition process involved generating a reference database. Prior to recording the spectra, a gold reference was utilised. Spectra were then obtained from six subsets of each of the 680 samples, serving as replicates to reduce uncertainties arising from potential seed heterogeneity. Specifically, six random samples were taken from each sample to create biological and analytical replicates, ensuring coverage of the entire sample. The seeds from each subset were loaded into a ring cup with an internal diameter of 5 cm, and the six subsets of each sample were measured sequentially (Janila *et a*l., 2018). Spectra for each subset were collected after three rotation scans, with a spectral resolution of 16 cm-1, covering the range of 3952-11528 cm-1 (867-2530 nm), using the Tango spectrometer from Bruker (Manley & Baeten, 2018).

II.2.3. Phenotypic variability, heritability, and trait correlations in an interspecific population for yield-related traits

Phenotypic data collection involved recording three qualitative and twelve quantitative traits at one or more sites. Except plant growth habit, days to 50% flowering, and main stem height, all traits were recorded after harvest. The mean phenotypic value for each trait per genotype was calculated on a per-plant basis by dividing the total number of plants per genotype by the number of plants evaluated using descriptors for groundnut (IBPGR and ICRISAT, 1992). Details of each trait measurement are provided below:

Plant growth habit

- The plant growth habit (scale) was recorded at the podding stage on a 1-6 scale, where 1= procumbent 1, 2= procumbent 2, 3= decumbent 1, 4= decumbent 2, 5= decumbent 3, and 6= erect.

Pod beak and constriction

- Pod beak (scale) and constriction (scale) were evaluated on 30 pods using a scale of 0, 3, 5, 7, and 9 based on groundnut descriptors.

Days 50 % flowering

- The number of days from sowing and flowering of at least 50 % of plants was evaluated.

Plant height

- Main stem height **(**cm**)** was recorded 60 days after sowing, measured from plant collar to apex.

Pod morphology

- Pod length (mm) and width (mm) were measured on 30 pods using a calliper with digital a display.

Seed morphology

- Seed length (mm) and width (mm) were evaluated on 30 seeds using a calliper with a digital display.

Yield components

- Yield components were determined at all sites based on pod, haulm, harvest index, and seed dry mass. The process involved weighing the total biomass to determine the total biomass per plant (g). Pods were then removed and weighed to calculate the total pod weight per plant (g). Haulm weight per plant (g) was derived by subtracting the total pod weight from the total biomass. The process involved weighing the total biomass to determine the total biomass per plant (g). Pods were then removed and weighed to calculate the total pod weight per plant (g). Haulm weight per plant (g) was derived by subtracting the total pod weight from the total biomass. The harvest index (%) was computed as the percentage of pod weight to total biomass.
- Next, 100 pods were randomly sampled and weighed (g) to determine the weight of 100 pods. All seeds from these 100 pods were weighed, and mature seeds were separated and counted. The weight of 100 seeds (g) was calculated by dividing the weight of mature seeds by the number of mature seeds, then multiplying by 100.

II.2.4. Identification of QTL associated with yield trait

II.2.4.1. DNA extraction and library construction

For trait-marker discovery analysis, two-week-old leaves from 133 samples were sent to the Integrated Genotyping Service and Support (IGSS) platform, now known as SEQART AFRICA located at Biosciences Eastern and Central Africa (BecA-ILRI) Hub in Nairobi, for DNA extraction and Genotyping. DNA extraction was performed using the Nucleomag 96 plant genomic DNA extraction kit, employing the modified cetyltrimethylammonium bromide (CTAB) extraction method as outlined by Singh and Singh (2015).

The extracted genomic DNA ranged from 50 to 100 ng/µL. DNA quality and quantity were checked on 0.8% agarose. Subsequently, libraries were constructed according to Kilian *et al*.(2012) and Alam *et al*. (2018). DArTSeq complexity reduction method through digestion of genomic DNA and ligation of barcoded adapters followed by PCR amplification of adapter-ligated fragments.

II.2.4.2. Sequencing and genotyping

Libraries were sequenced using single-read sequencing runs for 77 bases on the HiSeq2500 platform, following the protocol outlined by Kilian *et al*. (2012). The IGSS platform utilises a genotyping by sequencing (GBS) DArTseq™ technology, which enables rapid, high-quality genome profiling, even from complex polyploid genomes. DArTseq markers scoring was achieved using DArTsoft14 which is an in-house marker scoring pipeline based on algorithms detailed in Kilian *et al*., 2012. Two types of DArTseq markers were scored, SilicoDArT and SNP markers, both assessed as binary values representing presence (1) or absence (0) of the respective marker sequence within the sample's genomic representation. These markers were aligned to a model reference genome (Bertioli *et al*., 2019) to identify their chromosomal locations.

II.2.5. Data analysis

II.2.5.1. Quality traits

II.2.5.1.1. Principal component analysis (PCA)

R software (R Core Team, 2021) with rchemo (Brandolini-Bunlon, *et al*., 2023) and rnirs packages (Lesnoff, 2021) were used to visualize raw spectra and perform data analysis. PCA was performed over the spectral range selected from 1000 to 2500 nm to describe variability across varieties and interspecific genotypes within and between environments.

PCA is a multivariate unsupervised statistical method able to project multivariate data and describe relevant trends in the analysed dataset (Manley & Baeten, 2018; Sampaio & Brites, 2021; Phuc *et al*., 2023). PCA can also reveal variables with loading that determine some inherent structure of the data, which can be interpreted in chemical terms. The reduction of the number of variables is achieved by making a linear combination of the original variables, which yields the so-called principal components (PC) that are decorrelated with each other. The PCA analysis was carried out on pre-treated spectra.

The full spectra underwent preprocessing to enhance the signal by reducing uncontrolled variations such as noise and baseline through Savitsky Golay (SavGol) and derivative techniques (Manley & Baeten, 2018; Li *et al*., 2019; Sampaio and Brites, 2021). In this study, the PCA results considered included (i) the score plot to visualize the sample projection on each PC and (ii) the loading plot to assess the influence of wavelength on each PC. PCA thus aids in emphasizing and interpreting variables, highlighting all relevant differences among genotypes within and between environments.

II.2.5.1.2. Mahalanobis distance Analysis

After conducting PCA, Mahalanobis distance was calculated to assess the distances among the 6 sample replicates for each sample. These distances were determined in units of standard deviations from the centre (mean) of the dataset. Initially, the Mahalanobis distances were computed individually for the 6 replicates of each sample. Subsequently, the distances were averaged for each sample, and Mahalanobis distances were recalculated based on these average values.

II.2.5.1.3. Sample classification using PLS-DA modelling on NIR spectra

PLS-DA was used for classifying varieties and interspecific genotypes by modelling and predicting genotype-specific spectra based on genetic and environmental origin. The data were split using the Duplex method (Snee, 1977) into a train set (N=541, 201, 108, 106, 126 respectively for Core population, AB-QTL Bafia, AB-QTL Mbalmayo, AB-QTL Nioro) and a test set (N=139, 42, 32, 31, 34 respectively for the previous populations), maintaining the same proportionality within each group. The train set was used to train the model, while the test set was used to evaluate its performance.

Prior to applying PLS-DA algorithms, the train set spectra were pre-processed by standard normal variate (SNV),Detrend, Savitsky Golay Filter and derivative and their combination (Manley & Baeten, 2018; Li *et al*., 2019; Sampaio and Brites, 2021). The best preprocessing was selected according to the error of classification by cross validation (2 K-fold group repeated 20 times) and the number of latent values was fixed. Then these parameters were used to build the PLS-DA model and applied on test set spectra. The resulting confusion matrix of a model was further evaluated to assess the model's performance using the following metrics for each group and for all (Recal & Demirel, 2021; Goodwin *et al*., 2022; Phuc *et al*., 2023).

- Recall (sensitivity): Proportion of samples of a specific class predicted by the model as belonging to that class (Recall=TP/(FN+TP)).
- Specificity: Number of samples predicted correctly to be in the negative class out of all the samples in the dataset that actually belong to the negative class (Specificity=TN/(FP+TN)).
- Precision: proportion of correct predictions among all predictions for a particular class (Precision=TP/(TP+FP)).
- Accuracy: number of samples correctly classified out of all samples present in the test set (Accuracy=(TP+TN)/(TP+FN+FP+TN)).
- False-negative rate (FNR): proportion of false negatives (FNR=FN/(TP+FN)).
- False-positive rate (FPR): proportion of false positives (FPR=FP/(TN+FP)).
- F1-score: Harmonic mean of precision and recall (F_1 score= (2 x Recall x Precision)/(Recall+Precision)).

True positive (TP) refers to a sample belonging to the positive class being classified correctly. True negative (TN) refers to a sample belonging to the negative class being classified correctly. False positive (FP) refers to a sample belonging to the negative class but being classified wrongly as belonging to the positive class. False negative (FN) refers to a sample belonging to the positive class but being classified wrongly as belonging to the negative class. Model performances were evaluated by their classification accuracy, which was calculated as the ratio of the number of correctly classified samples to the total number of samples.

II.2.5.2. Phenotypic data analysis

II.2.5.2.1. Analysis of variance

Phenotypic data analyses for each environment were conducted using the R statistical programming language (R Core Team, 2021). Basic statistical analyses, such as mean and range calculations, were performed for each trait. The data collected were subjected to analysis of variance (ANOVA) following the method of Kuznetsova *et al*. (2017) to estimate the genetic and replication effects on each trait within each environment. This was done following a standard linear model with genotype, replication, block, and interaction effects:

$$Yijk = \mu + Gi + rj + bjk + eijk$$

Where: Yijk = observed value for a given trait, μ = mean of the population, Gi = genotype effect, rj = replication effect, bjk = block within replication effect, and eijk = residual error.

II.2.5.2.2. Broad sense heritability

To assess and quantify the genetic variability among the interspecific lines, heritability in broad sense (H_2) was computed from the ANOVA following the method as outlined in Essandoh *et al*. (2022) as follow:

$$H_2 = ((MSG-MSE)/r)/ ((MSG-MSE)/r + (MSE/r))$$

Where:

MSG= mean square of genotype

MSE= mean square of residual or environment

r= the number of replications.

II.2.5.2.3. Phenotypic correlation

Phenotypic correlation analysis was conducted using Pearson's method to assess the relationship between traits across different environments, following the methodology suggested by Miller *et al.* (1958). Best linear unbiased predictors (BLUP) were extracted from the alpha lattice model for each genotype and trait at lme4 R package (Bates *et al*., 2015) and used for trait corelation assessment and QTL analysis.

II.2.5.3. Molecular data analysis

II.2.5.3.1. Construction of a genetic linkage map

Genotyping data obtained from the BC2F4 population was used for carrying out genetic linkage analysis using Joinmap software (Van Ooijen, 2006). Genotyping data for a total of 16,279 SNPs has been generated using DArT markers. These DArT SNP markers were renamed with the prefix 'Ah_' and chromosome name where 'Ah' stands for *Arachis hypogaea*, followed by physical position such as Ah01_492746 refereed to the research reported by Khera *et al*. (2019) and Pandey *et al*. (2020). Initially, the raw SNPs were filtered based on MAF =5.0 and missing calls of 30 %. After filtration, a total of 12,230 SNPs were retained for further analysis. The Chi-square (X^2) test was applied to the 12,230 SNPs to remove the distorted SNP markers. Finally, a set of 1,450 SNP markers was used for the construction of the genetic map. The identical markers were removed using the function remove identical. The individual linkage group for each chromosome of a tetraploid peanut was constructed at LOD threshold 3.0. The unmapped loci that do not show any linkage with SNP markers in a particular linkage group were considered as distorted and not mapped forcefully in the resulting linkage group. The intermarket distance was calculated by dividing the map length by the total number of SNP

loci mapped. In each linkage group, the intermarket distance was set to be less than 50 cM to avoid large genome intervals during QTL analysis. Kosambi mapping function was used to estimate the genetic distances and to convert the recombination frequencies into map distances in centimorgans (cM) (Kosambi, 2016).

II.2.5.3.2. QTL analysis

The QTL mapping was performed using the BLUPs value of each trait in each environment, Maroua, Mbalmayo and Bafia, together with a refined genetic map and genotypic data. Thus, marker-trait association analyses were conducted to investigate QTLs that are associated with the 15 traits in 123 BC2F4 lines. The inclusive composite interval mapping (ICIMADD) method (Meng *et al*., 2015) implemented in QTL IciMapping software v4.1.0.0 was used to detect QTL and estimate their phenotypic effects. QTLs with a positive or negative additive effect for a specified trait imply that the increase in the phenotypic value of the trait is contributed by the alleles from the recurrent parent, Fleur11, and wild parent, ISATGR 278-18, respectively. In order to indicate the presence of a significant QTL effect for each trait and get more information about QTL, the threshold method, LOD ≥ 2.5, value at type I error 5 %, was determined using 1000 permutation times. The specific parameters were set as: for detecting additive QTL, Step was 1.0 cM and PIN was 0.001. QTL were declared major if the phenotypic variance explained was >10% and minor when the variance explained was less than 10 % (Collard et al., 2005).

The final high-resolution linkage map for the major effect QTL was generated using the linkage map view package in R software (Ouellette *et al.*, 2018). QTL for different traits were considered co-located, when their positions with significant LOD scores were located in the same marker intervals. The name of QTL was designated with an initial letter "q" (abbreviation of QTL), followed by the capital letters to designate the respective trait and, then, the chromosome number. This QTL nomenclature was referenced in the research report by Luo and Chen *et al*. (2017). Similarly, if more than one QTL for the same trait were identified, we added another number on the basis of the relative position of QTLs on the chromosome.

CHAPTER III. RESULTS AND DISCUSSION

III.1. RESULTS

III.1.1. Screening peanut core-collection and interspecific population for quality traits using NIRS

III.1.1.1. Spectra profiles and quality control

From the raw spectra, eleven relevant absorbance peaks were observed around the wavelengths of 929, 1033, 1465, 1763, 2306, 2350 and 2510 nm, with four wide spectral peaks appearing close to 1210, 1723, 1932 and 2140 nm (Appendix. II.). Only 2 of 4080 spectra (0.04%), were identified as an outlier and were discarded for analyses. PCA was performed to check the effect of date on spectra acquisition, and no cluster related to date was found (Appendix III). With few exceptions, the Mahalanobis distance among the six subsets of each sample was consistent among samples (Fig. 12). Thus, the spectra graph is presented in Fig. 13 as the average absorption of each sample from the six replicated spectra.

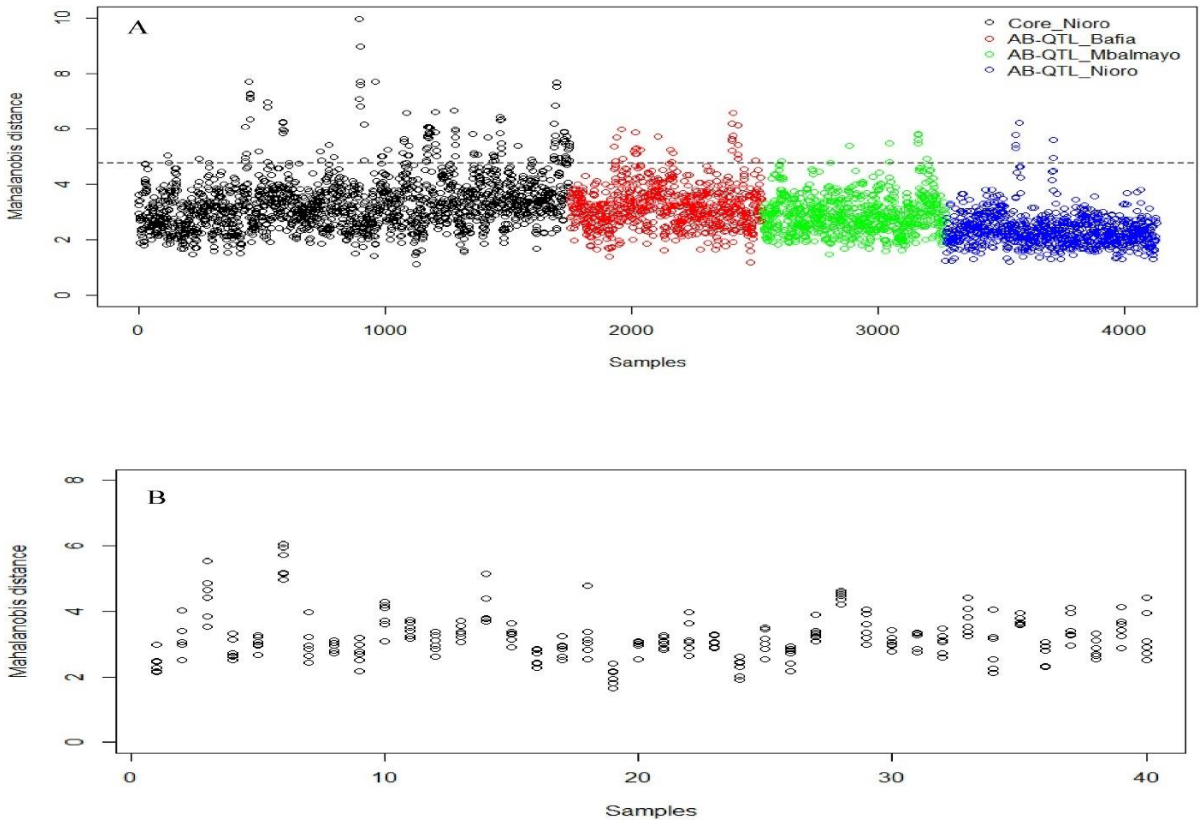


Fig. 12. Plot showing Mahalanobis distance among the six subsets of each sample of four populations**.** Each dot represents one spectrum. MD Details of 40 samples (B) is figured from the 4080 spectra (A) for a better MD visualization among the 6 spectra of each sample.

III.1.1.2. Genetic variability and environmental impact on intact-seed composition

The mean absorbance spectra of varieties and interspecific genotypes, according to their environment, are presented in Fig. 13. A huge variation of absorbance along the spectra was observed among varieties and interspecific genotypes within and between environments. Four absorbance strata, superimposed on each other, were observed for all wavelengths from 1000-2500 nm (Fig. 13A).

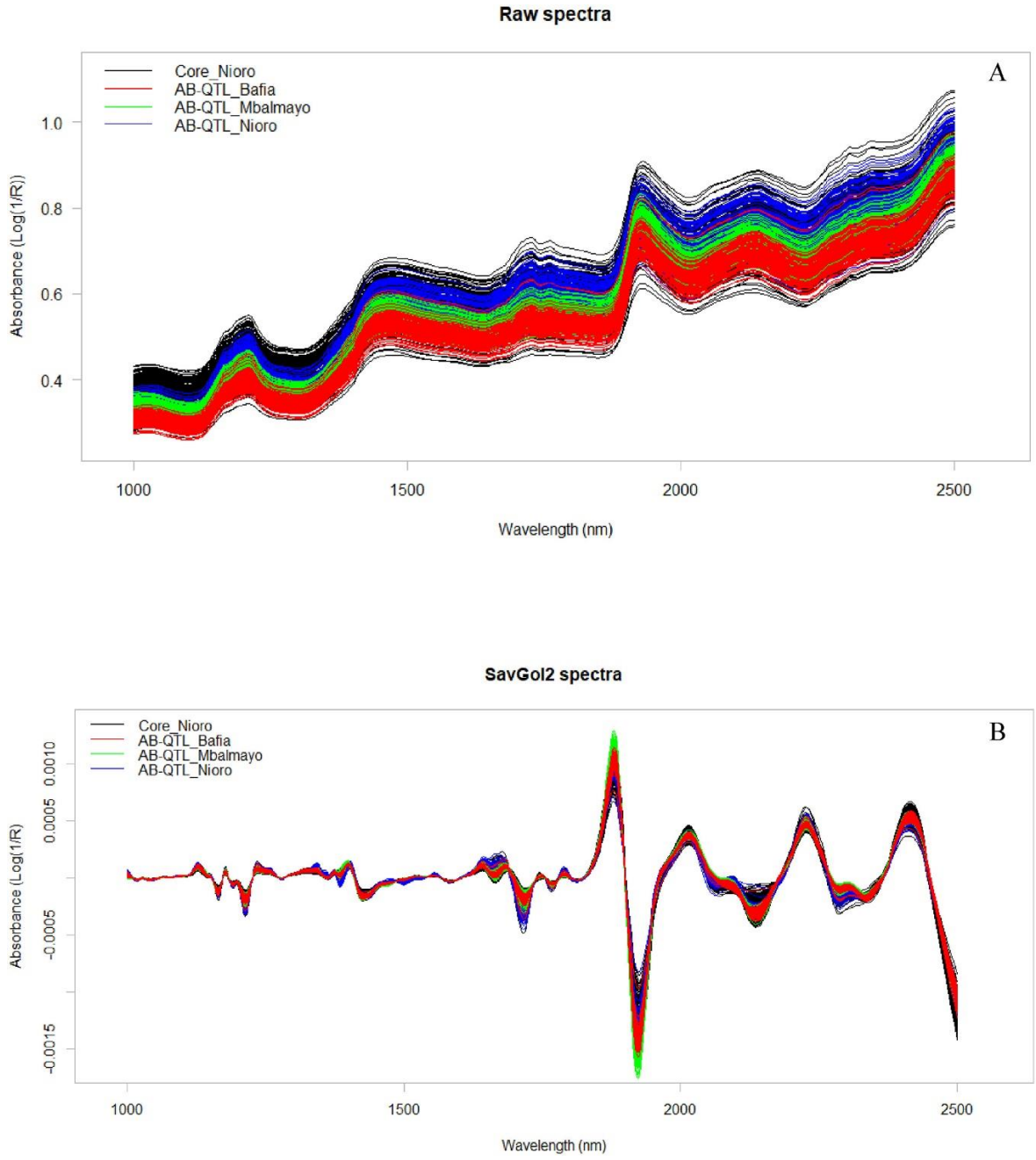


Fig. 13. NIRS spectra of intact-seed according to genetic and environment origin of samples without treatment (A) and after Stavisky Golay filter with derivative 2 pre-processing (B).

Each spectra group corresponds to each of the four studied sets. The widest stratum corresponded to the set of the core collection while the three other strata were each specific to the three sets of the interspecific population, each from one of the three studied environments, Bafia, Mbalmayo and Nioro (Fig. 13A). The absorbance range of interspecific population was highest in Bafia followed by Mbalmayo and Nioro, pointing out the effect of environmental factors on chemical composition of seeds.

III.1.1.3. Pretreatment effects on spectra

The absorbance spectra pre-treated by Savitzky-Golay filter with a window width of 15 points and the first derivative are shown in Fig. 13B. In contrast to raw data, the absorbance range of the interspecific population, particularly from Mbalmayo, was highest, at the relevant peak of 1723 and 1932 nm wavelengths (Fig. 13B). As expected, this indicates that the pretreatments eliminated physical effects due to seed dimension, surface of seed, etc., with consequences on light diffusion. In contrast to raw data, the absorbance range of the interspecific population was highest at the 1723 and 1932 nm wavelengths, particularly at Mbalmayo (Fig. 13B). Furthermore, a huge MD, from 1 to 8, was found among varieties and genotypes in one hand and among the 3 environments in the second hand, with the highest value from Mbalmayo compared to other environments (Fig. 14).

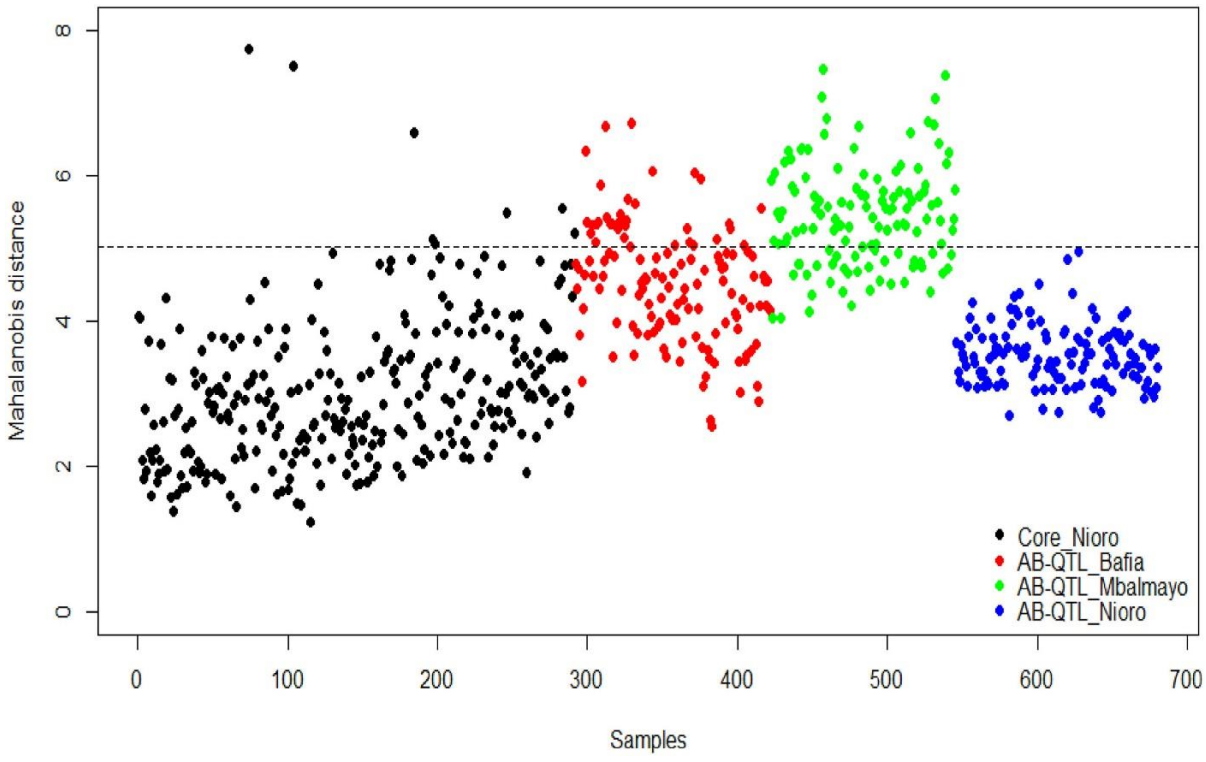


Fig. 14. Plot showing Mahalanobis distance among varieties and interspecific genotype.

III.1.1.4. Principal component analysis

The first 5 PC represent 98 % of the total variability with the values 60.5, 17.0, 15.5, 3.6 and 1.6, respectively. The 4 PC were presented as PC1/PC2 and PC3/PC4 score plots in (Fig. 16). As expected, these figures show greater variability in the core collection and less variability in the other groups. The PC3/PC4 plot makes it easy to distinguish the 4 seed lots. These plots showed that samples from different genetic and environmental origins are able to be well clustered and that they have great potential to be correctly identified. Loading plots showing how each variable correlates with PC are shown in (Fig. 15). The first loading indicates that the regions around 1900 and 2150 nm have a higher influence on PC1.Likewise, regions around 1210, 1720 and 2300 nm are more related to PC2. For PC3, the region around 2400 nm seems to be more important. And PC4 is more related to 1400, 1800, 1950 and 2150 nm region.

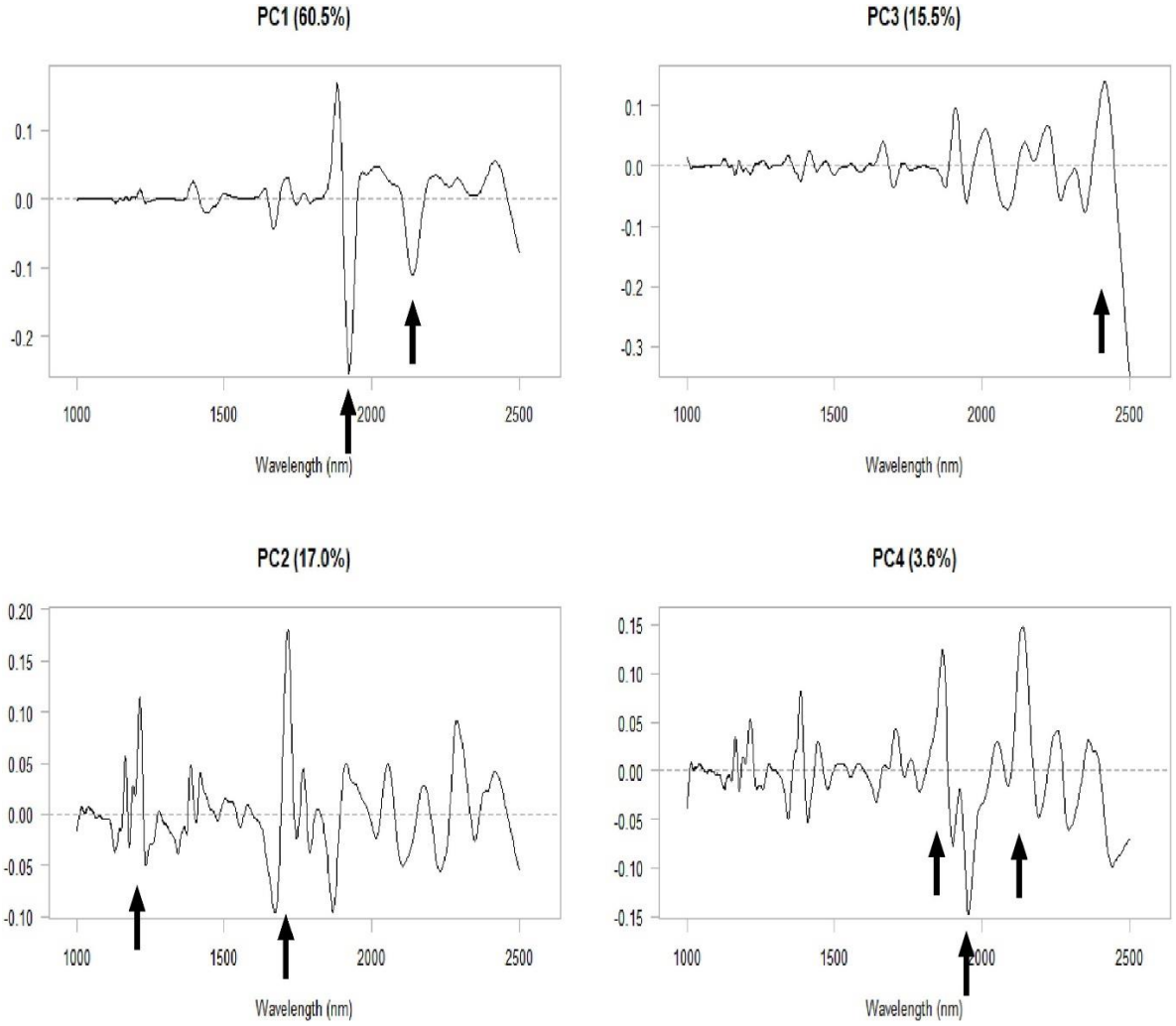


Fig.15. PCA loading plots for the fourth first PCs showing how each variable correlate to each PC for wavelength.

III.1.1.5. Discrimination of genetically related interspecific genotypes among environments

The score plots show that data could be grouped into four clusters, with overlapping main clusters at the margin, some interspecific genotypes and varieties superimposed, particularly, at the Nioro environment-set cluster (Fig. 16). The two most separated environments in the plane, determined by plot scores, were Mbalmayo and Bafia. Variation of seed constituents was widest in Bafia, followed by Mbalmayo and Nioro. With few exceptions, all interspecific genotypes from Mbalmayo have high positive values at the PC3 compared to the other environments. This suggests that Mbalmayo environment positively increases the seed traits associated with PC3. Finally, the African varieties studied in one environment added genetic variability to the environmental variability, resulting in a wide range of differences.

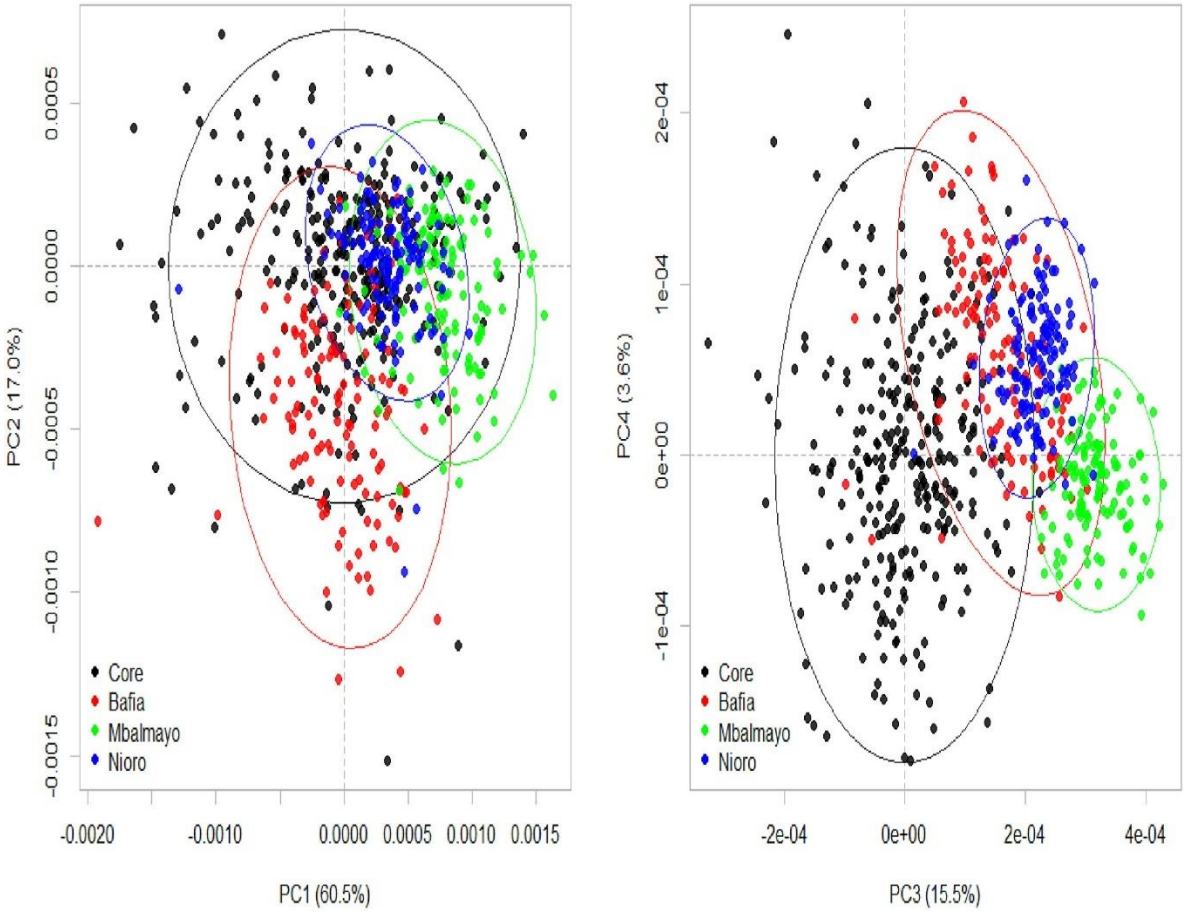


Fig. 16. PCA visualization of core varieties and interspecific genotypes among environments. PCA 2-dimensional score plots of PC2 and PC1 (A) and PC3 and PC4 (B) using NIRS spectra. African varieties are labelled in black. interspecific AB-QTL genotypes from Bafia, Mbalmayo and Nioro, environments are labelled in red, green and blue, respectively.

III.1.1.6. Classification based on intact-seed spectra

A PLS-DA model was developed, and the classification results of the model are shown in Table 4. The classification accuracy on the test set was 99.6% with correctly classified instances of the 4 samples sets, i.e., African varieties in one environment and the interspecific genotypes from the 3 environments (Table 4). Interestingly, the confusion matrix achieved for the two sets, Bafia and Nioro shows 100% of instances classified correctly with 100% at both sensitivity and specificity. These two sets do not show incorrect instances, even in the model generated when all other sets are considered, thus confirming that their intact-seed composition is very different from each other and from those of the other intact-seed samples. These results suggest that NIRS combined with machine vision is a very promising tool for the classification of peanut genotypes, depending on each combination of the genetic and environmental origins, that determine plant nutritional availability.

Table 4. Confusion matrix showing classification performance of PLS-DA model applied to test set sample (N=139, Class 1: Core, Class 2: AB-QTL Bafia, Class 3: AB-QTL Mbalmayo, Class 4: AB-QTL Nioro).

		Predicted								
		1	2	3	4	Actual	Accuracy	Precison	Recall	F1-score
Actual	1	41	0	0	0	42	0.993	0.976	1.000	0.988
	2	0	32	0	0	32	1.000	1.000	1.000	1.000
	3	1	0	31	0	31	0.993	1.000	0.969	0.984
	4	0	0	0	34	34	1.000	1.000	1.000	1.000
	Pred	41	32	32	34	139				

Accuracy	0.996
Specificity	0.998
Recall	0.993
Precision	0.993
Proportion of false-negatives	0.007
Proportion of false-positives	0.007

III.1.2. Phenotypic variability, heritability, and trait correlations in an interspecific population for yield-related traits

III.1.2.1. Phenotypic variability

Phenotypic variability across environments was observed, with all yield-related data combined showing a continuous normal or almost-normal distribution (Fig. 17). The phenotypic range of variation was moderate to high for all quantitative traits, except for days to 50% flowering in each environment (Appendix. II). Plant growth habit (GH) exhibited morphological variation ranging from semi-erect to the ground to totally erect. A similar range of variation was observed for pod beak (PB) and pod constriction (PC), ranging from slight to prominent and from slight to deep, respectively. Except for growth habit (GH), pod beak (PB), and pod constriction (PC), the mean population values for all other quantitative traits tended to be skewed toward the phenotypic value of the recurrent parent, Fleur11, for pooled data (Fig. 17), across three environments. These mean ranges were observed to be: plant height: 10.71 to 18.25, total biomass: 38 to 65.5, pod weight: 7.22 to 13.84, haulm weight: 31.31 to 52.13, harvest index: 16.63 to 27.67, hundred pod weight: 65.47 to 80.13, hundred seed weight: 34.08 to 40.4, pod length: 23.27 to 26.74, pod width: 10.6 to 11.69, seed length: 11.57 to 12.8, seed width: 6.9 to 7.76. The range of variation in each environment is presented in Appendix IV.

When comparing the phenotypic performance of each genotype across all three environments and combined data, an interspecific line, 11_28_10, consistently outperformed, followed by 11_28_20, for hundred pod weight, hundred seed weight, pod length, pod width, seed length, and seed width. Despite many genotypes showing high mean values outside of the cultivated parent's mean value, they were not consistently performing for many traits across all environments. Almost all studied traits showed moderate to high phenotypic variation, indicating transgressive segregation for these traits. Analysis of variance (ANOVA) showed significant differences ($P < 0.001$) among BC2F4 lines in each environment (Table 5) for all traits except for hundred pod weight (HPW) and pod width (PW) ($P<0.05$) at Mbalmayo and Bafia, respectively.

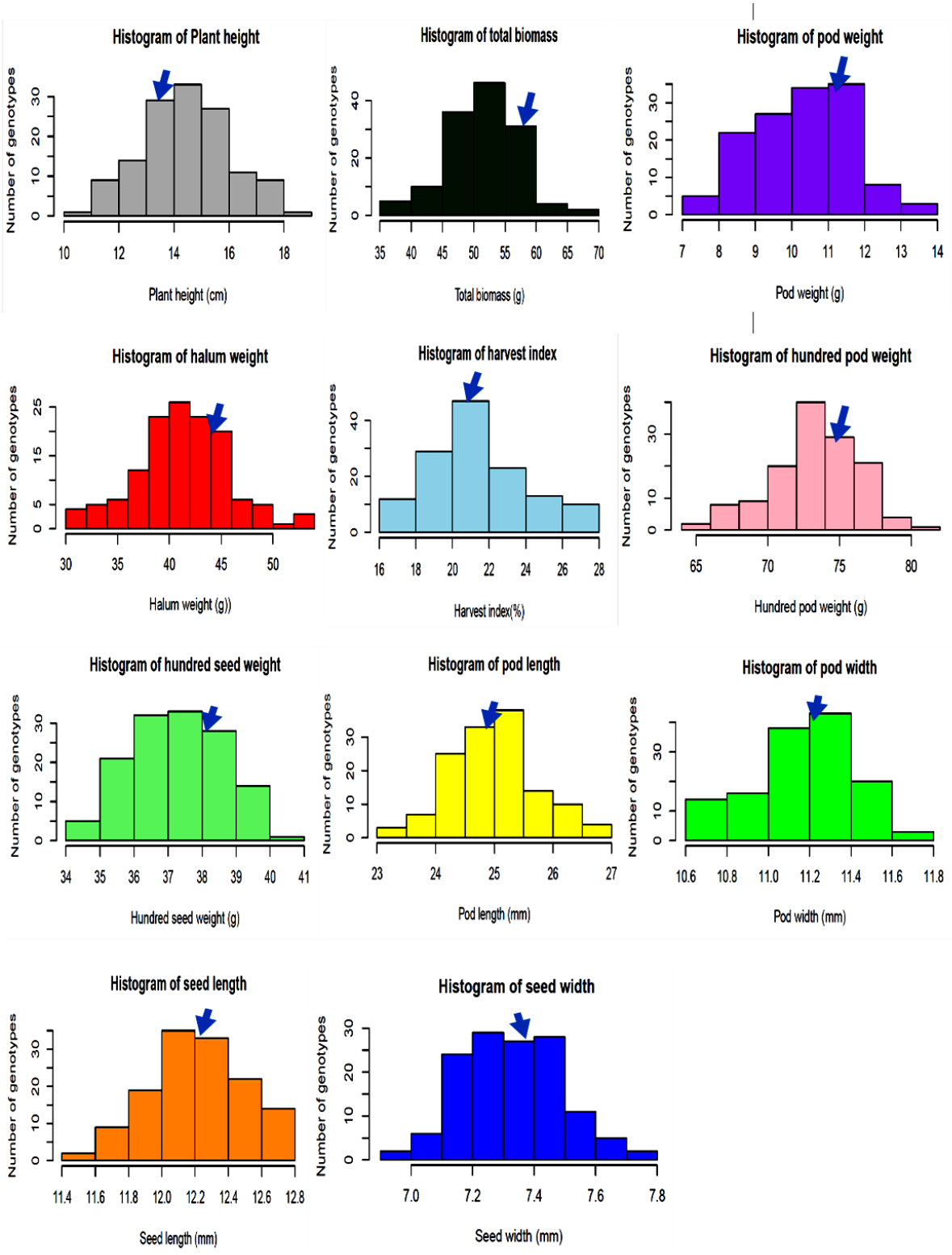


Fig. 17. The frequency distribution of yield-related traits among interspecific genotypes and the recurrent parent, Fleur11. The presence of an arrow sign indicates the position of the recurrent parent within the distribution.

III.1.2.2. Broad sense heritability

Moderate (025) to high (0. 99) broad sense heritability estimates were observed for all studied traits. The heritability was relatively high for most traits at Maroua, possibly due to the smaller phenotypic variance and reduced environmental influence on trait expression (Table 5). Briefly, the heritability estimates were observed for days to 50% flowering (D50%F) (0.58), plant height (PH) (0.59-0.69), total biomass (TB) (0.62-0.64), pod weight (PWT) (0.59-0.69), halum weight (HaW) (0.51-0.63), harvest index (HI) (0.54-0.77), 100 pod weight (HPW) (0.25-0.63), 100 seed weight (HSW) (0.56-0.61), pod length (PL) (0.54-0.84), pod width (PW) (0.39-0.89), seed length (SL) (0.52-0.8), seed width (SW) (0.44-0.8), pod constriction (PC) (0.71-0.85) and pod beak (PB) (0.66-0.79).

III.1.2.3. Phenotypic correlations

The correlation between traits is shown in Table 6. Of these significant associations, 9 associations were negative, and 30 significant associations were positive. Among yield contributing traits, TB showed strong positive correlations with PL (r=0. 29***), PWT (r=0. 57***), HaW (r=0. 93***), and positive with HPW (r=0. 25**), HSW (r=0. 26**) and PW (r=0. 2**), however, it was negatively correlated with HI (r=-0. 39***). HI was also negatively correlated with HaW (r=-0. 57***). Pod weight (PWT) positive corelated with PL (r=0. 41***), HaW (r=0. 38***), HI (r=0. 33***), HSW (r=0. 24**) and PW (r=0. 21*). The four pod and seed traits, PL, SL, HPW and HSW, had significant positive associations between each other. These, PL with SL (r=0. 41***), HSW (r=0. 4***) and HPW (r=0. 28***), SL with HSW (r=0. 32***) and HPW (r=0. 22**), and HPW with (r=0. 5***) suggest that these traits tend to vary together. Similarly, positive correlations were also revealed between each pair of the following traits: PW with SW (r=0.41***), HSW (r=0.48***) and HPW (r=0.57***), SW with HPW(r=0.3***) and HSW(r=0.37***) and HSW with HSW (r=0.57***), suggesting possible co-regulation or interdependence between these traits. PC was positively correlated with PB (r=0.49***), and SL (r=0.29***). Whereas, PB with SW (r=-0.33***), and PW (r=-0.5***) were negatively correlated. PL and HSW were positively correlated with all traits except PL with PW and SW, and HSW with PC and HI, indicate their strong influence on overall plant productivity and seed quality.

Table 5. Summary statistics of traits in the three environments

.	Maroua				Mbalmayo				Bafia			
Traits	Mean	F	Pr	H_2	Mean	F	Pr	H_2	Mean	F	Pr	H_2
GH					4.66	3.41	<0.001***	0.71				
D50%F					27.00	2.30	<0.001***	0.58				
PH	12.32	234	<0.001***	0.99	16.41	2.45	<0.001***	0.59				
TB	67.24	2.62	<0.001***	0.62	39.21	2.70	<0.001***	0.63	50.25	2.78	<0.001***	0.64
PWT	16.34	2.45	<0.001***	0.59	9.22	2.51	<0.001***	0.60	7.920	3.28	<0.001***	0.69
HaW	51.20	2.70	<0.001***	0.63	30.12	2.04	<0.001***	0.51	43.90	2.18	<0.001***	0.54
HI	27.12	3.80	<0.001***	0.74	24.23	2.17	<0.001***	0.54	15.27	3.00	<0.001***	0.67
HPW	76.94	2.70	<0.001***	0.63	69.59	1.33	0.027*	0.25	72.59	2.20	<0.001***	0.55
HSW	34.86	2.30	<0.001***	0.57	35.52	2.27	<0.001***	0.56	39.45	2.53	<0.001***	0.61
PL	25.00	6.08	<0.001***	0.84	25.29	2.60	<0.001***	0.73	24.67	2.18	<0.001***	0.54
PW	11.49	7.79	<0.001***	0.89	10.98	2.44	<0.001***	0.59	11.09	1.67	0.002**	0.39
SL	12.34	6.83	<0.001***	0.86	11.88	2.88	<0.001***	0.65	12.02	2.12	<0.001***	0.52
SW	6.95	4.66	<0.001***	0.80	7.71	1.75	<0.001***	0.44	7.120	2.11	<0.001***	0.50
PC	3.85	6.63	<0.001***	0.85	4.60	4.96	<0.001***	0.80	4.590	3.44	<0.001***	0.71
PB	4.87	4.76	<0.001***	0.79	5.33	4.23	<0.001***	0.77	5.290	2.97	<0.001***	0.66

D50%F: days to 50% flowering, PH: plant height, GH: growth habit, PB: pod beak, PC: pod constriction, PL: pod length, PW: pod width, SL: seed length, SW: seed width, TB: total biomass, PWT: pod weight, HaW: haulm weight, HPW: hundred pod weight, HSW: hundred seed weight, F: f calculated, Pr: probability value ($p<0.001$=***, $p<0.01$=** and $p<0.05$=*), Env: environment and H_2=broad sense heritability.

Table 6. Pearson correlations for yield component traits evaluated over 3 environments.

										0.49 ***	PC
									0.29* **	0.16	**SL**
								0.12	-0.05	0.08	**HI**
							0.17*	**0.41*** **	**0.32*** **	0.24 **	**PL**
						0.41* **	**0.33*** **	0.14	-0.02	-0.07	**PWT**
					0.38***	0.19*	**-0.57*** **	0.09	0.01	-0.2*	**HaW**
				0.97* **	**0.57*** **	**0.29*** **	**-0.39*** **	0.1	0.00	-0.21*	**TB**
			0.2**	0.15	0.21**	0.13	0.02	0.12	-0.23**	**-0.5*** **	**PW**
		0.57* **	0.25**	0.2**	0.16	**0.28*** **	**0.00**	0.22**	-0.19*	-0.26**	**HPW**
	0.5* **	**0.48*** **	0.26**	0.19*	0.24**	**0.4*** **	0.02	**0.32*** **	-0.02	0.18*	**HSW**
0.37***	**0.3*** **	**0.41*** **	0.12	0.16	0.15	0.18	0.09	0.14	-0.08	**-0.33*** **	**SW**
HSW	**HPW**	**PW**	**TB**	**HaW**	**PWT**	**PL**	**HI**	**SL**	**PC**	**PB**	

Values with *, ** and *** implies significant at p = 0.05, p < 0.01 and p < 0.001 respectively. HPW: hundred pod weight, HSW: hundred seed weight, PW: pod width, TB: total biomass, HaW: halum weight, PWT: pod weight, PL: pod length, SL: seed length, HI: harvest index, PC: pod constriction, PB: pod beak and SW: seed width.

III.1.3. Identification of QTL associated with yield traits

III.1.3.1. Linkage map construction

A genetic linkage map was constructed (Table 7) containing 1,450 DArT SNP loci. spans a total length of 1,358.02 cM, with an average distance of 2.21 cM between adjacent markers on 20 linkage groups (LGs). The number of markers per linkage group varies from 5 to 254, averaging 72.5 markers per group. The LGs ranged from 36.07 to 112.47 cM in length, with average inter-marker distances of 0.37 to 8.73 cM, and 5 LGs contained over 100 marker loci. Among the linkage group, Ah18 was the shortest, with 49 loci spanning 36.07 cM, while Ah02 was the longest group, with 59 loci spanning 112.47 cM. A Chromosome, Ah16 contained the fewest markers with 5 loci, whereas Ah08 had the highest density with 254 loci. The highest marker density of 2.68 SNPs/cM was observed on Ah04, while Ah13 displayed the lowest marker density of 0.11 SNPs/cM.

Table 7 . Summary of the genetic map constructed.

Linkage group	Number of SNPs Mapped	Map distance(cM)	Inter -marker distance (cM)
Ah01	211	87.07	0.41
Ah02	59	112.47	1.91
Ah03	63	85.69	1.36
Ah04	112	41.78	0.37
Ah05	26	37.37	1.44
Ah06	117	59.32	0.51
Ah07	16	32.58	2.04
Ah08	254	109.39	0.43
Ah09	203	85.55	0.42
Ah10	91	99.49	1.09
Ah11	33	79.46	2.41
Ah12	34	68.61	2.02
Ah13	9	78.59	8.73
Ah14	57	90.89	1.59
Ah15	14	42.58	3.04
Ah16	5	37.64	7.53
Ah17	43	45.8	1.07
Ah18	49	36.07	0.74
Ah19	44	73.56	1.67
Ah20	10	54.11	5.41
Total	1450	1,358.02	2.21

III.1.3.2. QTL identification

The QTL analysis revealed that at least one QTL was detected for each of the analysed traits, except for pod beak. A total of 44 main effects QTLs were identified for 14 yield-traits, scattered across 17 different linkage groups (LGs), as summarised in Table 8 and Appendix V. Notably, no QTLs were detected on Ah04, Ah12 and Ah16.

Days to 50% flowering: -

- Four QTLs were identified for the days to 50% flowering (D50%F). These QTLs namely qD50%FAh01, qD50%FAh10.1, qD50%FAh10.2 and qD50%FAh14 were detected on Ah01, Ah10 and Ah14, individually explaining 4.33-7.05% of the phenotypic variance. In all cases, the alleles associated with flowering precocity belonged to the cultivated parent, Fleur11.

Plant architecture: -

- Plant height (PH) and growth habits (GH) are the most important plant architecture traits. In total, six QTLs were identified on 5 LGs. Among these, two QTLs, qGHAh03 and qGHAh06 on Ah03 and Ah06 respectively, were detected for GH. Both QTL exhibited phenotypic values associated with alleles from the recurrent parent. Additionally, four QTLs, namely qPHAh01.1, qPHAh01.2, qPHAh15, and qPHAh17 located on Ah01, Ah15, and Ah17, were identified for PH. Except for qPHAh01.2, the positive contribution was linked to alleles of ISATGR 278-18. These three QTLs collectively explained 28.27% of the phenotypic variance.

Yield related traits: -

- Ten QTLs were identified for the most important pod and seed traits, including pod weight (PWT), hundred pod weight (HPW) and hundred seed weight (HSW). Among these, four QTLs for HPW (qHPWAh01, qHPWAh05, qHPWAh08, and qHPWAh13) were detected on Ah01, Ah05, Ah08, and Ah13, explaining 7.49% to 15.39% of the phenotypic variance. Notably, qHPWAh05 showed increased HPW associated with ISATGR 278-18 alleles, while the other three QTLs were linked to Fleur11 alleles, collectively contributing 38.4% to the phenotypic variance. For HSW, five QTLs were identified, all showing positive contributions from the recurrent parent, Fleur11. Additionally, a single QTL, qPWTAh01, was found for PWT on Ah01, with the PWT increase linked to ISATGR 278-18 alleles.

It was noted that three QTLs (qHPWAh01, qHSWAh11 and qHSWAh17) appear to be novel and have not been previously reported.

Table 8. Summary of detected QTL

Traits	Evt	LG	C1	PS	C2	PEV%	LOD	ADD	Parent
D50%F	MB	1	56.50	57	57.50	7.05	2.64	-0.07	ISATGR 278-18
D50%F	MB	10	40.50	41	41.50	4.33	2.63	-0.22	ISATGR 278-18
D50%F	MB	10	45.50	46	54.50	5.86	3.14	-0.27	ISATGR 278-18
D50%F	MB	14	61.50	62	65.50	6.26	2.78	-0.33	ISATGR 278-18
GH	MB	3	74.50	75	75.50	9.10	2.62	0.30	Fleur 11
GH	MB	6	46.50	50	51.50	6.56	2.57	0.29	Fleur 11
PH	MR	1	72.5	73	73.5	6.10	2.79	0.16	Fleur 11
PH	MB	1	20.50	21	21.50	9.42	3.70	-0.64	ISATGR 278-18
PH	MB	15	5.50	7	9.50	8.77	3.33	-1.11	ISATGR 278-18
PH	MB	17	31.50	32	32.50	10.08	2.72	-0.53	ISATGR 278-18
TB	MB	7	22.50	23	25.50	9.86	2.73	2.80	Fleur 11
PWT	BF	1	82.5	86	86	12.73	3.6	-1.14	ISATGR 278-18
HaW	MR	19	59.5	60	60.5	7.56	2.72	2.48	Fleur 11
HaW	MB	7	22.50	24	25.50	10.06	2.76	0.90	Fleur 11
HaW	BF	1	2.5	4	4.5	3.67	3.47	-0.07	ISATGR 278-18
HaW	BF	2	41.5	42	42.5	10.19	7. 96	4.26	Fleur 11
HaW	BF	18	0	2	2.5	2.86	2.7	-0.37	ISATGR 278-18
HaW	BF	19	60.5	61	61.5	7.96	4.8	-5.12	ISATGR 278-18
HI	MR	2	82.5	84	85.5	9.13	2.97	-4.75	ISATGR 278-18
HI	MB	18	31.50	32	33.50	9.37	2.57	1.86	Fleur 11
HI	BF	8	88.5	89	91.5	17.56	7.5	-2.22	ISATGR 278-18
HI	BF	14	15.5	18	20.5	5.13	2.7	-1.71	ISATGR 278-18
HI	BF	14	46.5	48	48.5	4.71	2.5	-1.92	ISATGR 278-18
HPW	MR	13	6.5	11	16.5	12.32	2.51	3.95	Fleur 11
HPW	MB	1	21.50	22	22.50	10.69	2.84	0.70	Fleur 11
HPW	MB	8	56.50	57	57.50	15.39	3.36	0.63	Fleur 11
HPW	BF	5	4.5	8	8.5	7.49	2.66	-1.09	ISATGR 278-18

Traits	Env.	LG	C1	PS.	C2	PVE%	LOD	ADD	Parents
HSW	MR	3	56.5	57	57.5	8.73	2.66	0.04	Fleur11
HSW	MR	11	20.5	22	23.5	10.7	3.66	0.01	Fleur11
HSW	BF	8	88.5	89	91.5	10.17	2.99	0.27	Fleur11
HSW	BF	13	8.5	12	16.5	6.95	4.08	2.45	Fleur11
HSW	BF	17	0	1	3.5	8.52	2.92	1.15	Fleur11
PL	MB	14	5.50	6	8.50	6.07	2.82	1.05	Fleur11
PW	BF	11	42.5	44	44.5	8.58	5.36	-0.14	ISATGR 278-18
SL	MB	14	42.50	44	47.50	11.19	3.07	0.33	Fleur11
SL	BF	1	21.5	22	22.5	7.74	5.12	-0.01	ISATGR 278-18
SL	BF	6	26.5	27	27.5	7.24	4.63	0.26	Fleur11
SL	BF	8	28.5	29	29.5	14.53	6.7	0.03	Fleur11
SL	BF	9	80.5	81	81.5	4.80	2.8	0.13	Fleur11
SL	BF	18	3.5	4	4.5	7.66	2.74	0.19	Fleur11
SW	MR	19	0	2	6.5	9.16	2.54	-0.25	ISATGR 278-18
SW	MR	20	50.5	54	54	8.84	3.05	-0.20	ISATGR 278-18
SW	BF	1	2.5	4	4.5	7.04	2.67	0.09	Fleur11
PC	BF	17	6.5	7	8.5	11.49	4.12	0.48	Fleur11

D5%F: Days to 5% flowering, GH: Growth habit, PH: Plant height, TB: Total biomass, PWT: Pod weight, HaW: Haulm weight, HI: Harvest index, HPW: Hundred pod weight, HSW: Hundred seed weight, PW: Pod width, PL: Pod length, SL: Seed length, PC: Pod constriction, SW: Seed width, LGs: Linkage groups, CI: Confidence interval, PS: Position, PVE: Phenotypic variance explained and ADD: Additive.

- Moreover, a total of 12 QTLs explaining 2.86% to 17.56% of the phenotypic variance were detected for three other yield traits: total biomass (TB), haulm weight (HaW), and harvest index (HI) mapped across 8 LGs. Specifically, qTBAh07 for TB, qHaWAh01, qHaWAh02, qHaWAh07, qHaWAh18, qHaWAh19.1and qHaWAh19.2 were identified for HaW, while qHIAh02, qHIAh08, qHIAh14.1, qHIAh14.2 and qHIAh18 were associated with HI mapped on Ah01, Ah02, Ah07, Ah08, Ah14, Ah18 and Ah19. Among these QTLs, qTBAh07, qHaWAh02, qHaWAh07, qHaWAh19.1, and qHIAh18 showed increased phenotypic values for their respective traits linked to the recurrent alleles.

Pod morphology: -

- A total of 3 QTLs were detected for traits related to pod constriction (PC), pod length (PL) and pod width (PW) spread over 3 different LGs (Ah11, Ah14, and Ah17). These QTLs individually explained 6.07% to 11.49% of the phenotypic variance (Table 8). For a QTL, qPWAh11 the amphidiploid alleles contributed with the increase pod width.

Seed morphology: -

- The QTL analysis for seed morphology revealed six QTL for seed length (SL) and three QTL for seed width (SW) mapped on 8 LGs. These QTLs, qSLAh01, qSLAh06, qSLAh08, qSLAh09, qSLAh14, qSLAh18, qSWAh01, qSWAh19 and qSWAh20, explained between 4.8% to 14.53% of the observed phenotypic variance. For qSLAh01, qSWAh19 and qSWAh20 the favourable alleles were associated with the wild parent alleles. Out of 44 detected QTLs linked with yield-traits, 13 were major (PVE >10%) linked with PH, PWT, HaW, HI, HPW, HSW and SL(Fig.18). Among the 13 major QTLs, qPWTAh01, qPHAh17 and qHIAh08, favourable alleles were from ISATGR 278-18. The stability of the QTL is an important parameter determining the utility of QTLs. Among the 13, QTL for HaW, HPW, HSW and SL were detected stable across two environments.

III.1.3.3. Co-localization of QTL

The co-localization of QTLs for different traits in the study provided valuable insights into the genetic control of key agronomic traits in peanut. Five genomic regions were identified to contain co-localized QTLs for different traits, with two QTLs co-mapped on Ah01, Ah07, Ah08, and Ah13 showing pleiotropic effects or tight linkage. On Ah01, a 2.5-4.5 cM region (Ah01_200647 - Ah01_1334363) carried co-localized QTLs for haulm weight (HaW) and seed width (SW) at Bafia (BF). Another co-mapped QTL on the same LG, between 21.5-22.5 cM (Ah01_4928888-Ah01_4929547), clustered for seed length (SL) and hundred pod weight (HPW) at Bafia (BF) and Mbalmayo (MB), explaining 7.74 and 10.69% PVE, respectively.

On Ah07, a co-mapped QTL for total biomass (TB) and HaW was identified at 22.5-25.5 cM (Ah07_79222318 - Ah07_9653368) at MB. On Ah08, a major QTL was found at 88.5-91.5 cM (Ah08_48931192 - Ah08_49571607) exhibiting pleiotropic effects on harvest index (HI)

and hundred seed weight (HSW), explaining 17.56 and 10.17% PVE, respectively. Additionally, a QTL on Ah13 flanked by markers Ah13_6340341 - Ah13_6420673 mapped at position 6.5-16.5 cM was co-localized for HSW (BF) and HPW at Maroua (MR) with 6.95–12.32% PVE respectively.

Out of 10 co-localized QTLs, qTBAh07 and qHaWAh07, and qHSWAh13 and qHPWAh13 were expected given the strong positive correlations between these co-localized traits (Table 8), indicating the existence of pleiotropic effects of a single gene or tight linkage. However, qHaWAh01 and qSWAh01, and qHIAh08 and HSWAh08 were not phenotypically correlated, as well as the allele derived from ISATGR 278-18 increased the value of haulm weight and harvest index but decreased the value of hundred seed weight and seed width

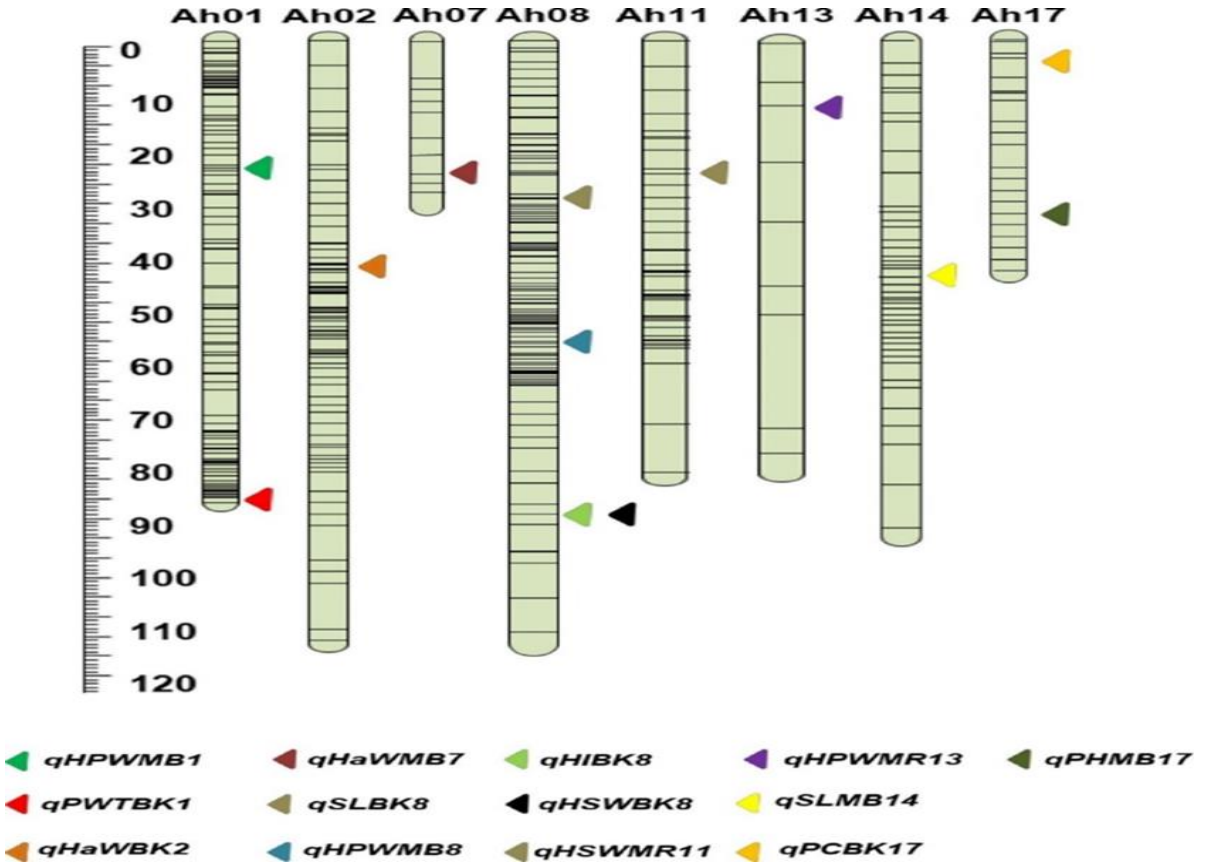


Fig.18. Genomic location of major QTLs. q: QTL, HPW: hundred pod weight, HaW: halum weight, HI: harvest index, PH: plant height, PWT: Pod weight, SL: seed length, HSW: hundred seed weight PC: pod constriction, BK: Bokito/Bafia, MB: Mbalmayo, MR: Maroua and 1, 2. 7, 8, 11, 13,14 and 17 related to Ah01-Ah17.

III.2. DISCUSSION

This study aimed to employ a non-destructive extraction method to screen peanut germplasms for seed quality traits, while simultaneously assessing phenotypic variability and identifying genomic regions or QTL associated with yield-related traits.

III.2.1. Screening peanut core-collection and interspecific population for quality traits

This part of the study aimed to assess the potential of near infrared spectroscopy (NIRS) for rapid screening of germplasm variability using intact seeds from 300 cultivars of an African core-collection and 133 genotypes of an interspecific population. These samples were field-evaluated in three environments across two countries. We conducted PCA analysis for genotype discrimination and developed a PLS-DA based prediction model to classify varieties and genotypes. Subsequently, we discussed the efficacy of NIR spectroscopy as a tool for fast and non-destructive characterization of large germplasm in various environments, within both intra and interspecific breeding contexts.

III.2.1.1. NIR as tool for rapid and non-destructive large samples assessment in peanut

Peanut is an important oilseed and the need to characterize peanut germplasm is essential as the demand for peanut is increasing continuously in various end product applications. According to the rapid and non-destructive attributes of the NIR, a total of 6 days was required to obtain all spectra of the six subsets of intact-seed of the 680 samples. The low level (0.04%) of outlier spectra on the global data set was considered as a good foundation for analysis. From raw spectra, eleven major peaks were observed around the wavelengths 929, 1033, 1210, 1465, 1723, 1763, 1932, 2140, 2306, 2350 and 2510 nm. The region around 1210, 1763, 2306 and 2350 nm could all be assigned to fatty acids or oil content, of which the spectral position in agreement with earlier studies by Govindarajan *et al*. (2009), Sundaram *et al*. (2010) and Tao *et al*. (2019).

The spectral peak around 2140 nm would likely result from the absorbance of proteins. The absorbance peak around 1465 nm might be related to the O-H overtone bond. The sharp peak around 1932 nm was due to the strong absorption of water contained in peanut kernels (Govindarajan *et al*.,2009; Sundaram *et al*., 2010; Tao *et al*.,2019).

III.2.1.2. Germplasm variability and environment impact on seed chemical composition

A wide genetic variation was found among varieties and interspecific genotypes within environments.

The magnitude of the genetic influence among varieties and genotypes suggested that nutritional related traits were amenable to improvement through intra and interspecific breeding (Huang *et al.*, 2019; Nawade *et al*., 2019; Shasidhar *et al*., 2020). An environmental effect on seed compounds was highlighted by using the same interspecific population, thorough 3 environments. The largest variation was found in Bafia, followed by Mbalmayo and Nioro. Bafia in savanna and Mbalmayo in forest, grown under yellow and ocher vertisol, respectively in Cameroun while Nioro in Sahel in Senegal exhibited leached ferruginous soil. The interaction between all agroecological scenarios (climate, vegetation and soil) and spatial factors create a complex system of environments that affect peanut plant growth and development, leading to a discrimination among genotypes within and between environments.

As previously reported by chemical studies, seed composition is influenced by environment but also has a strong genetic component. The variation of oil composition has been related to temperature (Harris and James, 1969**)**, planting date (Andersen and Gorbet, 2002), location and soil moisture (Holaday and Pearson, 1974**;** Young *et al*., 1974**),** photoperiod (Dwivedi *et al*.,2000), market grade (Mozingo *et al.,* 1988**)** and genotype (Harris and James, 1969; Worthington and Hammons, 1977; Holaday and Pearson, 1974; Norden *et al*., 1987**;** Mozingo *et al*., 1988; Harch *et al*., 1995**;** Gimode *et al*.,2020). However, with multiple environmental factors mentioned above, it is difficult to decipher factors underlining variation in this study. Likewise, identifying suitable peanut genotypes for global ecological zones remains a challenging task due to the significant genotype variability across environments. Finally, the African varieties studied in one environment added genetic variability to the environmental one, resulting in a wide range of variability.

III.2.1.3. Pattern of genetic variability of interspecific population in comparison to core-collection

According to the spectra profiles and PCA plot, the genetic pattern of interspecific population covers, remarkably half of the spectrum of the core-collection, that turned out to be largest, as we expected. Interestingly, we found specific genetic variation among interspecific genotypes that was not subtle cover by the core-collection at the common Nioro environment. Interspecific genotypes with positive value on the main PCA axis were recorded as promising genotypes for quality traits. As early reported, three introgression lines with elevated Oleic/Linoleic profiles were found using chemical survey of 77 Peanut interspecific lines (Gimode *et al*.,2020), showing the results of wild genes. The genotypes recorded in study could be recommended for further breeding for developing suitable varieties.

In this respect, evaluation of the segregating interspecific population could further ease the discovery of QTL and valuable wild genes that contribute to improved seed quality.

III.2.1.4. Classification of varieties and interspecific genotypes using PLS-DA modelling

A PLS-DA model was successful developed from seed spectra to classify varieties and genotypes according to their genetic and environmental origin. A robust prediction accuracy of 99.6% was achieved. The confusion matrix achieved for the two environments, Nioro and Bafia shows 100% of instances classified correctly with 100% at both sensitivity and specificity. This confirm that their seed chemical composition was very different from each other and from those of the other seed samples. These results suggested that PLS-DA model could be used to classify peanut genotypes depending on the combination of the genetic and environment origins of seeds, which influence plant nutritional properties. In further studies, the current model would be confronted to wide others breeding populations in different environment to predict genetic and environment origin and nutritional content of whole seeds.

III.2.2. Phenotypic variability, heritability, and trait correlations in an interspecific population for yield-related traits

One of the aims of this study was to assess phenotypic variability among 133 BC2F4 interspecific genotypes, with a particular emphasis on estimating heritability and determining trait associations for yield traits.

III.2.2.1. Phenotypic variability

The study revealed significant genetic differences among the 133 genotypes of the interspecific BC2F4 population, indicating rich phenotypic diversity. This diversity, demonstrated across all quantitative traits except days to 50% flowering, aligns with an approximately normal distribution, indicative of a polygenic inheritance model. This model is particularly suited for QTL analysis, as highlighted by previous studies (Singh & Singh, 2015; Getahun, 2021). Moreover, qualitative traits displayed morphological variation, suggesting control by multiple genes with minor effects. For instance, the qualitative trait of plant growth habit (GH) exhibited a spectrum of morphological variations, ranging from semi-erect to the ground to completely erect. Similar variability was observed for pod beak (PB) and pod constriction (PC), ranging from slight to prominent and slight to deep, respectively. These findings align with previous research by Fonceka *et al*. (2012a) and Sambou *et al*. (2017), further highlighting the polygenic nature of these traits.

The presence of transgressive segregation underscores the potential of introgression breeding to broaden the genetic base of cultivated crops. Certain lines within the studied population surpassed the original parent in performance, indicating promising targets for breeding endeavours. In this study, across the three environments studied, significant phenotypic variability persisted, emphasizing the strong influence of both environmental factors and genotype responses. Hence, yield-related traits of peanut are strongly determined by environmental factors and different genotypes, which may react differently to environmental factors.

Introgression breeding, involving the transfer of desirable genes from wild germplasm into cultivated parental backgrounds, presents a unique avenue for enhancing genetic diversity and improving crop traits. Previous studies, such as that by Sharma *et al*. (2013), have underscored the role of hybridization and introgression in generating phenotypic variability and developing improved crop varieties. In this study, high rates of segregation were observed based on mean values of yield traits, with certain advanced backcross lines exhibiting superior traits compared to the cultivated parent. Consistently outperforming in all environments and pooled data, 11_28_10 emerged as a notable line, followed by 11_28_20, across various yield-related traits such as hundred pod weight, hundred seed weight, pod length, pod width, seed length, and seed width. Transgressive segregation was evident, with mean values of certain traits falling beyond the ranges defined by the parent line Fleur11. Despite some inconsistencies across environments and traits, most traits exhibited significant degrees of transgressive segregation among genotypes. This finding was also reported by Faye *et al*. (2015), Chen *et al*. (2017), Khedikar *et al*. (2017), Liang *et al*. (2018), Huang *et al*. (2015b) and Chen *et al*. (2016), using RIL population, Fonceka *et al*. (2012a), and Sambou (2017) with BC population, who reported transgressive segregant yield component traits in various populations of peanut.

ANOVA analysis revealed highly significant phenotypic variations ($P<0.001$) among the 133 BC2F4 lines for most traits, further indicating the substantial contribution of genomic regions or QTL from the wild relative of cultivated peanut to yield-related traits. This underscores the potential of harnessing genetic traits from wild relatives in breeding programs aimed at improving peanut yield.

III.2.2.2. Broad sense heritability

Broad-sense heritability estimation offers insight into the proportion of transmissible genetic variation relative to total variation under specific environmental conditions, providing valuable information for selection decisions (Falconer and Mackay, 1996). Traits exhibiting higher broad-sense heritability across diverse environments or seasons are more reliably selectable in breeding programs, as they are primarily governed by genetic factors. Elevated heritability values indicate a substantial influence of additive genetic effects on the target traits.

In this study, a range of heritability values was observed across different traits and experimental environments, suggesting varying degrees of genetic and environmental influence on trait expression. Notably, moderate to high broad-sense heritability (0.39-0.89) was observed for most traits, with exceptions such as relatively low heritability (0.25) for hundred pod weight at Mbalmayo and very high heritability (0.99) for plant height at Maroua. These findings are consistent with previous research by Fonceka *et al*. (2012a), Huang *et al*. (2015), Chen *et al*. (2017, 2018), Luo *et al*. (2017, 2018), who reported high heritability for traits such as hundred pod and seed weight, pod and seed length, and width (0.63-0.95).

The high broad-sense heritability values (up to 0.84) for these traits, including in this study, indicate that genetic factors predominantly govern the observed variation, with minimal influence from environmental factors. This suggests the potential for progress through selection strategies. Across the three environments, broad-sense heritability values exceeding 0.57 in Maroua were notable, likely due in part to the uniform environmental conditions at the onset of the experiment.

Similarly, moderate to high heritability estimates (0.51-0.74) were observed in this study for traits such as haulm weight, harvest index, pod weight, and total biomass, consistent with findings by (Fonceka *et al*., 2012a; Faye *et al*., 2015). Plant height, a trait closely linked to plant architecture, lodging resistance, biomass, yield, and mechanized harvesting adaptability in crops like peanut (Lv *et al*., 2018), exhibited moderate to high broad-sense heritability (0.43-0.89), as reported by previous studies (Fonceka *et al*., 2012a; Huang *et al*., 2015b; Khedikar *et al*., 2017; Lv *et al*., 2018). These results underscore the complexity of trait heritability, which fluctuates with varying environmental conditions. Consequently, breeding lines selected based on performance in one location may not necessarily perform optimally in another environment.

III.2.2.3. Phenotypic correlation

Correlation analysis provides insight into the indirect selection of quantitative traits by assessing the degree and direction of association between them. In this study, 39 significant associations were identified, with 9 being negative and 30 being positive. For instance, total biomass and haulm weight exhibited a highly positive correlation (0.97***), consistent with findings by Fonceka *et al*. (2012a) (0.95***) and Sambou (2017) (0.93***). Additionally, significant positive associations were observed between total biomass and pod weight (0.57***), hundred seed weight and hundred pod weight (0.57***), pod width and hundred pod weight (0.5***), hundred pod weight and hundred seed weight (0.5***), pod constriction and pod beak (0.49***), and pod width and seed width (0.48), among others. Conversely, negative significant associations were noted between harvest index and haulm weight (-0.57***), harvest index and total biomass (-0.39***), pod beak and pod width (-0.5***), and pod beak and seed width (-0.33***), consistent with previous studies by Fonceka *et al*. (2012a) and Faye *et al*. (2015), who reported negative correlations between harvest index with total biomass and halum weight.

Although significant positive associations were observed between traits such as hundred pod and seed weight, and pod and seed length, no significant correlations were found between pod length and pod width, and seed length and seed width. However, previous studies have reported positive correlations among these traits (Fonceka *et al*., 2012a; Huang *et al*., 2015b; Chen *et al*., 2016, 2017; Luo *et al*., 2017, 2018). Notably, pod length and hundred seed weight exhibited significant correlations with most studied traits, suggesting their regulatory roles in other yield traits. Significant positive correlation was also reported between pod constriction with pod length and pod beak in the current study by (Fonceka *et al*., 2012a; Zhang *et al*., 2023), as this study confirms that pod constriction significantly positively correlated with pod beak, pod length and seed length, and, pod beak positively correlated with pod length whereas significant negatively correlated with pod width and seed width.

Strong positive correlations between yield-related traits suggest the involvement of common genes or pathways at the molecular level, potentially sharing genomic regions. Conversely, negative correlations imply independent inheritance of these traits. Utilizing these strong positive correlations can aid in simultaneous trait improvement through modern breeding approaches.

III.2.3. Identify genomic regions (QTLs) associated with yield- traits

This study aims to identify genomic regions (QTLs) associated with yield-contributing traits in peanut, which are complex and influenced by multiple genes, modifiers, and environmental factors. Understanding the genetic control behind these traits is crucial for marker-assisted selection (MAS), which has proven successful in enhancing traits such as rust and late leaf spot resistance and increasing oleic acid content (Chu *et al*., 2011; Kolekar *et al*., 2017; Shaidhar *et al*., 2020).

Despite challenges posed by limited genetic diversity within cultivated peanut germplasm due to its monophyletic origin and polyploid nature, wild peanut species harbour valuable traits that could enrich cultivated varieties (Upadhyaya *et al*., 2011). However, transferring genes from wild relatives faces hurdles like ploidy differences and the risk of unwanted traits (Burow *et al*., 2001; Kumari *et al*., 2014). Synthetic tetraploids, merging wild relatives with cultivated peanuts, have overcome these challenges, expanding the gene pool (Simpson *et al*., 1993; Favero *et al*., 2006; Mallikarjuna *et al*., 2010).

Amphidiploids have enabled exploration of wild peanut genetics, revealing many positive QTL effects from wild alleles (Fonceka *et al*., 2012a, 2012b; Sambou , 2017). Therefore, this study utilizes an interspecific backcross population (BC2F4) derived from the cross between recurrent parent Fleur11 and the amphidiploid ISATGR 278-18. Through genotyping with Diversity Array Technology (DArT) and DArTseq, QTLs associated with yield-related traits were mapped across three distinct peanut cultivation environments in Cameroon, aiming to identify chromosomal regions linked to these traits and tap into beneficial alleles from wild species.

III.2.3.1. Linkage map

In genetics research, creating linkage maps is crucial for understanding gene inheritance and its connection to specific traits. In our study, we constructed a linkage map using 1,450 DArT markers, covering 1,358.02 cM with an average distance of 2.21 cM between adjacent markers across 20 linkage groups.

Previous peanut AB-QTL studies have utilized various populations and crosses to effectively map genetic loci. For instance, the first tetraploid genetic map in peanut, involving the amphidiploid (cultivated ‘Florunner’ x synthetic ‘TxAG-6’), incorporated 370 RFLP loci across 23 linkage groups (Burow *et al*., 2001).

Similarly, the second map, based on an AB-QTL population derived from a cross between cultivated tetraploid and synthetic amphidiploid [(Fleur 11 x (*A. duranensis* V14167 x *A. ipaensis* KG30076)] x 4), mapped 147 loci (Fonceka *et al*., 2009). Additionally, the synthetic amphidiploid ISATGR 278–18 (*A. duranensis* ICG 8138 x A. batizocoi ICG 13160) crossed with the recurrent parent ICGS 76 produced a linkage map with 114 microsatellite SSR markers spanning 746.15 cM (Kumari *et al*., 2014). More recent studies have constructed two AB-populations, (ICGV 91114 x ISATGR 1212 and ICGV 87846 x ISATGR 265-5A), with 258 loci (1,415.7 cM map length and map density of 5.5 cM/loci) and 1,043 loci (1,500.8 cM map length with map density of 1.4 cM/loci), respectively (Khera *et al*., 2019). These populations derived from wide crosses exhibit high polymorphism rates compared to those from within cultivated genotypes, indicating the introgression of novel alleles from the wild parent into the cultivated parent. This introgression results in a more diverse and potentially adaptive genetic background.

In our study, both the number of loci and the density of the genetic map surpass those of previous genetic maps based on SSR and DArT markers in peanuts, as reported by several researchers (Foncéka *et al*., 2009; Sambou , 2017; Wilson *et al*., 2017; Khera *et al*., 2019; Kumari *et al*., 2020). In BC populations, marker loci ranged from 114 to 330, with a density of 5.5 to 6.6 cM. Other studies utilizing F2 and RIL populations (Pandey *et al*., 2014b; Huang *et al*., 2015b; Chen *et al*., 2016, 2017; Shaidar *et al*., 2017) reported marker loci ranging from 206 to 854, with marker density between 4 and 9 cM. The increased marker density in our genetic map provides more detailed insights into the studied traits, enabling more accurate QTL mapping and identification of closely linked markers for marker-assisted selection.

While the density of our map aligns relatively closely with those published by Shirasawa *et al*. (2013), Chen *et al*. (2017), Luo *et al*. (2018), Lv *et al*. (2018), Hu *et al*. (2018), Chavarro *et al*. (2020), and Pandey *et al*. (2020a), it is less dense compared to high-density genetic maps used for QTL identification of yield, oil quality traits, and disease resistance (Luo *et al*., 2017; Shaidar *et al*., 2017; Han *et al*., 2018; Li *et al*., 2019; Khera *et a*l., 2019; Liu *et al*., 2020; Sun *et al*., 2021; Guo *et al*., 2020), with marker density ranging from 0.45 to 1.6 cM/locus.

When comparing our map to previously published maps of cultivated peanut, the length was smaller, except for relatively similar studies (Chen *et al*., 2017; Wilson *et al*., 2017; Luo *et al*., 2018; Lv *et al*., 2019; Li *et al*., 2019). This suggests a higher level of genetic linkage between markers, indicating a more efficient and informative genetic map for QTL mapping and gene discovery.

These high-density genetic maps, including ours, facilitate anchoring QTLs to the physical map, enable the identification of QTLs in narrow physical intervals, closely linked flanking markers, provide a better understanding of genetic architecture and candidate gene information, and facilitate marker-assisted selection.

III.2.3.2. QTL identification

Our study identified a total of 44 QTLs associated with 14 yield component traits across three environments. These findings offer invaluable insights into the genetic underpinnings of peanut yield and lay the groundwork for discussing the implications of the identified QTLs for peanut breeding programs. Notably, approximately 45% of the detected QTL positive effects were attributed to alleles from the wild relative, ISATGR278-18, highlighting the potential of integrating wild germplasm in peanut breeding efforts to enhance yield-related traits. These QTLs positively influenced several valuable yield traits, including plant height (8.77-10.08%), total pod weight (12.73%), halum weight (2.86-9.13%), harvest index (4.71-17.56%), 100 pod weight (7.49%), pod width (8.58%), seed length (7.74%) and seed width (8.84 and 9.16%).

In this study, three QTLs associated with qualitative traits such as plant growth habit (GH) and pod constriction (PC) were identified on chromosomes Ah03, Ah06, and Ah17. The recurrent parent contributed favourable alleles that increased the expression of these traits, resulting in moderate pod constriction and an erect growth habit. This observation aligns with findings reported by Foncéka *et a*l. (2012a) and Sambou (2017), who also utilized the same recurrent parent, indicating a consistent genetic basis for these traits across different studies. The collective contribution of favourable alleles to the recurrent parent in our study and above authors underscores the potential role of specific genomic regions in shaping plant growth habit and pod constriction in peanut. On the other hand, the identification of 4 QTLs for days to 50% flowering on Ah01, Ah10, and Ah14, predominantly contributed by the wild parent leading to longer days of flowering. A QQTL mapped on Ah10in this study confirmed by Khedikar *et al*. (2017).

A total 13 major QTLwere detected, however, no consistency for QTLs across environments it may be due to presence of high environmental influence in in the experimental sites. Exception for harvest index, 100-seed weight, and seed length, which had six QTLs mapped on LGs Ah08/Ah18, Ah03/Ah13, and Ah08/Ah18, respectively, all remaining traits were identified on non-homeologous chromosomes.

Despite an equal distribution of QTLs between peanut A and B genomes (24 and 20 QTLs, respectively), the presence of QTLs on non-homologous chromosomes for the same traits suggests genetic complexity and potential epistatic interactions influencing trait expression. This phenomenon indicates the involvement of multiple genetic loci and pathways in controlling the expression of these traits, leading to the localization of QTLs on different chromosomes. Similarly, previous studies (Fonceka *et al*., 2012a; Huang *et al*., 2015b; Sambou , 2017) have reported that more than 96% of QTLs are mapped in non-homologous regions. Non-homeologous QTL locations may result from the absence of segregating alleles in one genome compared to the other or from natural and/or human-driven selection of different genes in the two subgenomes contributing to variation in the same trait (Fonceka *et al*., 2012a). These results may also be explained by the differential control of gene expression in subgenomes and/or by the movement of genes resulting in the disruption of collinearity, as a consequence of interspecific hybridization.

In this study, no QTLs were identified on LGs Ah04, Ah12, and Ah16. This absence could be attributed to several factors, including the lack of wild chromosome introgressions on these particular LGs due to strong disequilibrium within the population. Additionally, it's possible that there is a lack of polymorphic genes located in the cited LGs that are associated with the variations of the studied traits. This suggests that these genomic regions may not harbour genetic variation relevant to the traits under investigation in the specific population studied.

Comparing the QTLs detected in the present study to those identified in other studies for the same trait, it's evident that QTLs for hundred-seed weight are distributed across multiple linkage groups in the peanut genome. Previous studies (Fonceka *et al*., 2012a; Huang *et al*., 2015b; Faye *et a*l., 2015; Khedikar *et al*., 2017; Sambou, 2017; Chen *et al*., 2017; Liang *et al*., 2018; Khera *et al*., 2019; Pandey *et al*., 2020) have reported QTLs for hundred-seed weight on LGs Ah02, Ah03, Ah04, Ah05, Ah06, Ah07, Ah08, Ah09, Ah12, Ah13, Ah14, Ah15, Ah16, Ah18, and Ah19. In our study, five QTLs were identified for hundred-seed weight, located on chromosomes Ah03, Ah08, Ah11, Ah13, and Ah17. Notably, the QTLs identified on Ah11 and Ah17 were novel discoveries. Interestingly, a QTL for this trait was reported in close genomic proximity on chromosome Ah13 by Khedikar *et al*. (2017), linked by the SSR marker 'ahFAD2A-TC5D06. Additionally, Huang *et al*. (2015), Sambou (2017), and Khera *et al*. (2019) also reported the presence of five QTLs on Ah13, further emphasizing the importance of this genomic region and suggesting the potential presence of important genes influencing hundred-seed weight.

The clustering of QTLs for hundred-seed weight on certain linkage groups, particularly Ah13, across multiple studies suggests that this region may harbour crucial genes or genetic elements regulating this trait. Further investigation of these genomic regions could lead to the identification of candidate genes underlying hundred-seed weight and facilitate their utilization in peanut breeding programs aimed at enhancing seed yield and quality.

Similarly, among the four QTLs mapped for hundred-pod weight (HPW) on chromosomes Ah01, Ah05, Ah08, and Ah13, one QTL (qHPWAh01) was not previously detected and may represent a novel QTL. The presence of a QTL on chromosome Ah05 is noteworthy, as it has been associated with important genes for hundred-pod weight, with one QTL from our study and 21 from earlier studies (Sambou , 2017; Luo *et al*., 2017, 2018; Chen *et al*., 2017) reported in this region. For seed width, three QTLs were detected in the present study, mapped on chromosomes Ah01, Ah19, and Ah20. In contrast, previous studies (Fonceka *et al*., 2012a; Huang *et al*., 2015b; Chen *et al*., 2016, 2017; Sambou, 2017) have reported 33 QTLs located on chromosomes Ah01, Ah02, Ah03, Ah05-Ah13, Ah15, Ah16, and Ah19. Therefore, the identification of a QTL on Ah20 in our study represents a novel finding for this trait. Notably, for newly mapped QTL of seed width (HSW), an increase in the phenotypic value of the trait was associated with alleles from the wild parent. This suggests that wild germplasm may harbour favourable alleles for this trait and highlights the potential utility of wild relatives in peanut breeding programs aimed at improving seed width.

In this study, QTL for other important yield traits, including plant height, total biomass, pod weight, harvest index, and haulm weight, were mapped on Ah01, Ah02, and Ah07. These findings align closely with consistent QTL reported by Ravi *et al*. (2011), Gautami *et al*. (2012a), Fonceka *et al*. (2012a), and Faye *et al*. (2015). Overall, our study provides comprehensive insights into the genetic architecture of peanut yield and highlights the potential utility of wild germplasm in peanut breeding efforts.

III.2.3.3. Co-localization of QTL

Our QTL analysis unveiled significant co-localization of QTLs within five genomic regions: Ah01, Ah07, Ah08, and Ah13, each hosting two QTLs. These findings offer valuable insights into understanding yield components and accelerate Marker-Assisted Selection (MAS) breeding in peanuts. The observed co-localizations underscore the interdependence among yield traits, with pleiotropic QTLs emerging as pivotal genetic determinants shaping these traits within our BC2F4 population.

This highlights the presence of pleiotropic QTLs containing multiple tightly linked, trait-specific genes or genes influencing multiple traits, as discussed by Chen *et al*. (2017) and Chavarro *et al*. (2020).

For instance, within the 2.5-4.5 cM region on Ah01, two co-localized QTLs governing haulm weight and seed width were identified, prevalent at the Bafia (BF) location. Similarly, on Ah01, the QTL flanked by markers Ah01_4928888-Ah01_4929547 in LG Ah01 (21.5-22.5 cM) was found to be co-located for seed length and hundred-pod weight at both BF and Mbalmayo (MB), explaining 7.74% and 10.69% PVE, respectively. Previous study has reported similar findings for seed length and hundred-pod weight on Ah13 between 91-135 cM (Sambou, 2017). Likewise, another co-mapped QTL for total biomass and haulm weight was detected on Ah07 at 22.5-25.5 cM, observed primarily at MB, which explained 9.86% and 10.06% PVE, respectively. Consistent results were reported by Fonceka *et al*. (2012a), linking haulm weight and total biomass with the same marker, gi-0385_A, on Ah07 within 0-31 cM distance. Notably, a major QTL mapped on Ah08 (88.5-91.5 cM) exhibited pleiotropic effects for harvest index and 100-seed weight, elucidating 17.56% and 10.17% PVE, respectively. Furthermore, in our study, 100-pod and seed weight were co-located on Ah13, flanked by Ah13_6340341 - Ah13_4226296, spanning a 10 cM distance, supported by findings from Huang et al. (2015) and Sambou (2017), with SSR markers HAS0738_AHGS1788 and TC1E06_B-RN12H01_B, respectively. Additionally, the colocation of 100-pod and seed weight was detected on Ah12 by Fonceka *et al*. (2012a), associated with TC1B02_B SSR marker within the 0-10.9 cM distance.

Traits demonstrating strong positive correlation, sharing the same chromosomal location, and exhibiting the same additive effect due to genetic linkage suggest their inheritance as a unit. This implies that the observed strong positive correlation between these traits likely stems from their close physical proximity on the same chromosome, being passed down together due to genetic linkage.

CHAPTER IV. CONCLUSION, RECOMMENDATIONS AND PERSPECTIVEES

IV.1. CONCLUSION

In this study, we aimed to explore the potential of NIR combined with chemometric techniques for the rapid assessment of peanut germplasm. We also assessed phenotypic variability, estimated broad-sense heritability, and examined phenotypic correlations among interspecific genotypes for yield-related traits. Additionally, our objectives included the identification of beneficial genomic regions, or QTLs, associated with these yield traits in wild peanut specie.

To assess the potential of NIR for rapid germplasm screening for seed quality traits, we employed whole seeds from both a core-collection and an interspecific population, evaluated in three different environments. NIR spectroscopy proved to be an efficient and non-destructive tool for characterizing large germplasm collections across multiple environments. The rapid acquisition of spectra, requiring only six days for 680 samples, underscores its practicality for high-throughput analysis.

Our analysis of spectra and multivariate techniques revealed wide genetic variation among core-collection and interspecific genotypes, both within and between environments. While the core-collection exhibited the widest genetic diversity, specific variations were also observed among interspecific genotypes not covered by the core-collection. This suggests the potential for discovering new sources of diverse nutritional polymorphisms from wild derivatives. Furthermore, our study highlighted significant environmental impacts on seed composition within the interspecific population, with the largest variation observed in Bafia, followed by Mbalmayo and Nioro. We developed a robust PLS-DA model that accurately classified whole seed samples according to their environments.

Overall, NIR coupled with chemometric techniques proved useful for accurately assessing and distinguishing whole seeds within different environments. This approach not only facilitates rapid and non-destructive discrimination of peanut germplasms but also avoids the use of reagents and minimizes harmful residue generation, thereby ensuring environmental preservation. Based on our findings, this capability could facilitate predictions of intact-seed nutritional content and enhance the utilization of germplasm in breeding programs.

For the assessment of phenotypic variability, estimation broad sense heritability and trait correlation for yield traits, the study elucidated significant phenotypic variability among 133 genotypes, with almost all quantitative traits exhibiting an approximately normal distribution, indicative of a polygenic inheritance model suitable for QTL analysis. The persistence of significant phenotypic variability across three environments emphasized the combined impact of environmental factors and genotype responses on peanut yield-related traits. ANOVA analysis further confirmed substantial phenotypic variations among the BC2F4 lines, highlighting the noteworthy contribution of genomic regions or QTL from wild relatives in enhancing peanut yield. This study has identified moderate to high heritability values across traits and environments underscored the considerable genetic influence on trait expression. Correlation analysis identified 39 significant associations, with notable high significant positive correlations included total biomass with haulm weight, total biomass with pod weight, and hundred seed weight with hundred pod weight, among others. Positive associations were also found between traits like hundred pod weight, hundred seed weight, pod length, pod width, seed length and seed width each other. Strong positive correlations suggest shared genetic factors or pathways

To identify beneficial wild genomic regions, or QTLs, associated with yield traits, a high-resolution genetic map using SNP markers obtained through the DArTseq platform was constructed. It comprises 1,450 marker loci distributed across 20 linkage groups (LGs), covering a map distance of 1,358.02 cM, with an average marker density of 2.21 cM. Through QTL analysis, we identified a total of 44 QTL across 17 linkage groups, associated with 14 economically important yield traits across three different environments. Notably, four of these QTLs represent new loci that had not been previously mapped. Among the 44 QTLs identified, 13 were found to have a major effect (>10% phenotypic variance explained), suggesting their potential significance in peanut breeding programs through marker-assisted selection (MAS). Interestingly, our findings revealed that 20 out of the 44 QTLs (45%) showed an increase in the phenotypic value of the trait associated with alleles from wild relatives. This observation underscores the positive contribution of wild peanut species in peanut improvement efforts.

IV.2. RECOMMENDATIONS

- Due to the high phenotypic variability observed among environments for the studied traits, it is advisable to characterize this material for drought and foliar diseases resistance to identify potential sources of variation.
- The identified lines showing promising quality and yield traits should undergo testing across multiple locations and over multiple years to ensure their reliability and suitability for future breeding efforts.
- Recommended lines should be utilized in the development of mapping populations to precisely identify the genomic regions associated with the observed favourable chemical traits, facilitating targeted breeding efforts.
- Advanced lines that have not shown consistency for many traits across environments should undergo evaluation in different locations over multiple years to assess their stability and performance under diverse conditions.
- Further studies are necessary in multiple locations and over several years to validate the QTLs identified in the present study, ensuring their robustness and reliability for guiding breeding programs.

IV.3. PERSPECTIVEES

The observed genotypic variations among varieties and genotypes based on NIRS spectra suggest that nutritional traits could be enhanced through intra and interspecific breeding. Achieving breeding goals based on seed composition is feasible, given the influence of both environmental factors and genetic effects from different varieties and interspecific derivatives on seed chemical compounds. Despite the broad core-collection, significant genetic variation exists among interspecific genotypes not covered by it, implying the presence of untapped genetic diversity, possibly originating from wild relatives Varieties and interspecific lines identified in this study exhibited favourable spectra profiles, indicating their potential as valuable genetic materials for developing suitable varieties. Moreover, our identification of QTLs associated with enhanced phenotypic values from wild relatives highlights the valuable contribution of wild peanut species to peanut improvement efforts. This perspective underscores the potential for integrating wild germplasm into breeding programs to enhance genetic diversity and ultimately improve peanut yield and quality traits.

REFERENCES

Abady S., Shimelis H., Janila P., Yaduru S., Shayanowako A.I.T., Deshmukh D., Chaudhari S. & Manohar S.S., 2021a. Assessment of the genetic diversity and population structure of groundnut germplasm collections using phenotypic traits and SNP markers: Implications for drought tolerance breeding. *PLoS ONE,* 16(11): e0259883.

Abady S., Shimelis H., Pasupuleti J., Mashilo J., Chaudhari S. & Manohar S., 2021b. Assessment of the genetic diversity of groundnut (*Arachis hypogaea* L.) genotypes for kernel yield, oil and fodder quantity and quality under drought conditions. *Crop Science*, *61*(3): 1926-1943.

Al-Khayri J.M., Jain S.M. & Johnson D.V. (Ed.), 2019. Advances in Plant Breeding Strategies: Nut and Beverage Crops: Volume 4. Springer International Publishing, Cham.

Alam M., Neal J., O'Connor K., Kilian A. & Topp B., 2018. Ultra-high-throughput DArTseq-based silicoDArT and SNP markers for genomic studies in macadamia. *PloS One*, 13(8): e0203465.

Alhashim R. & Anandhi A., 2022. Global Warming and Toxicity Impacts: Peanuts in Georgia, USA Using Life Cycle Assessment. *Sustainability*, 14(6): 3671.

Ambika, Aski M.S., Gayacharan, Hamwieh A., Talukdar A. *et al.*, 2022. Unraveling Origin, History, Genetics, and Strategies for Accelerated Domestication and Diversification of Food Legumes. *Frontiers in Genetics*, 13:21.

Andersen P.C. & Gorbet D.W., 2002. Influence of Year and Planting Date on Fatty Acid Chemistry of High Oleic Acid and Normal Peanut Genotypes. *Journal of Agricultural and Food Chemistry*, 50(5): 1298-1305.

Andersen S.B. & Torp A.M., 2002. QTL Mapping in Crop Plants. In S. M. Jain, D. S. Brar, & B. S. Ahloowalia (Eds.), *Molecular Techniques in Crop Improvement* (pp. 541-562). Springer Netherlands ,Dordrecht.

Asíns M.J., 2002. Present and future of quantitative trait locus analysis in plant breeding. *Plant Breeding*, 121(4): 281-291.

Bakal H. & Arioglu H., 2019. The Determination of fatty acids composition and oil quality factors of some peanut varieties having different market types at different harvesting times in main and double crop growing seasons in Mediterranean Region. *Turkish Journal of Field Crops*, 24(2) :221-229.

Bakal H., 2020. Comparison of Fatty Acids Composition and Oil Quality Factors of Different Market Type Peanut (*Arachis hypogaea* L.) Varieties. Preprints, preprint.

Bansod B., Thakur R. & Holser R., 2015. Analysis of Peanut Seed Oil by NIR. *American Journal of Analytical Chemistry*, 6(12): 917-922.

Barkley N.A., Dean R.E., Pittman R.N., Wang M.L., Holbrook C.C. & Pederson G.A., 2007. Genetic diversity of cultivated and wild-type peanuts evaluated with M13-tailed SSR markers and sequencing. *Genetical research*, 89(2): 93-106.

Bates D., Mächler M., Bolker B. & Walker S., 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67: 1-48.

Bechara M.D., Moretzsohn M.C., Palmieri D.A., Monteiro J.P., Bacci M., Martins J., Valls J.F., Lopes C.R. & Gimenes M.A., 2010. Phylogenetic relationships in genus Arachis based on ITS and 5.8S rDNA sequences. *BMC Plant Biology*, 10(1): 1-12.

Bedru N., Matiwos T., Mekonnen T. & Belete T., 2024. Performance Evaluation of Different Sorghum Genotypes (*Sorghum bicolour* (L.) Moench) Using GGE Bi-plot Stability Analysis. *International Journal of Genetics and Genomics*, 12(2) :19-30

Bertioli D., Seijo G., Freitas F., Valls J., Leal-Bertioli S. & Moretzsohn M., 2011. An overview of peanut and its wild relatives. *Plant Genetic Resources*, 9(1): 134-149.

Bertioli D.J., Ozias-Akins P., Chu Y., Dantas K.M., Santos S.P. *et al.*, 2014. The Use of SNP Markers for Linkage Mapping in Diploid and Tetraploid Peanuts. *G3 Genes|Genomes|Genetics*, 4(1): 89-96.

Bertioli D.J., Jenkins J., Clevenger J., Dudchenko O., Gao D. *et al.*, 2019. The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nature Genetics*, 51(5): 877-884.

Bianchi-Hall C.M., Keys R.D., Stalker H.T. & Murphy J.P., 1993. Diversity of seed storage protein patterns in wild peanut (*Arachis, Fabaceae*) species. *Plant Systematics and Evolution*, 186(1): 1-15.

Bilal M., Xiaobo Z., Arslan M., Tahir H.E., Azam M., Junjun Z., Basheer S., & Abdullah, 2020. Rapid determination of the chemical compositions of peanut seed (*Arachis hypogaea*) Using portable near-infrared spectroscopy. *Vibrational Spectroscopy*, 110: 103138.

Bogie N.A., Bayala R., Diedhiou I., Dick R.P. & Ghezzehei T.A., 2018. Alteration of soil physical properties and processes after ten years of intercropping with native shrubs in the Sahel. *Soil and Tillage Research*, 182: 153-163.

Bomireddy D., Gangurde S.S., Variath M.T., Janila P., Manohar S.S. *et al.*, 2022. Discovery of Major Quantitative Trait Loci and Candidate Genes for Fresh Seed Dormancy in Groundnut. *Agronomy*, 12(2): 404.

Bonku R. & Yu J., 2020. Health aspects of peanuts as an outcome of its chemical composition. *Food Science and Human Wellness*, 9(1): 21-30.

Brandolini-Bunlon, M., Jallais B., Roger J.M. & Lesnoff M., 2023.R package rchemo: Dimension Reduction, Regression and Discrimination for Chemometrics. https://github.com/ChemHouse-group/rchemo.

Brown A.H.D., 1989. Core collections: a practical approach to genetic resources management. *Genome*, 31(2): 818-824.

Burow M.D., Simpson C.E., Starr J.L. & Paterson A.H., 2001. Transmission Genetics of Chromatin From a Synthetic Amphidiploid to Cultivated Peanut (*Arachis hypogaea* L.): Broadening the Gene Pool of a Monophyletic Polyploid Species. *Genetics*, 159(2): 823-837.

Burow M.D., Simpson C.E., Faries M.W., Starr J.L. & Paterson A.H., 2009. Molecular biogeographic study of recently described B- and A-genome Arachis species, also providing new insights into the origins of cultivated peanut. *Genome*, 52(2): 107-119.

Burow M.D., Starr J.L., Park C.-H., Simpson C.E. & Paterson A.H., 2014. Introgression of homeologous quantitative trait loci (QTLs) for resistance to the root-knot nematode [Meloidogyne arenaria (Neal) Chitwood] in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.). *Molecular Breeding*, 34(2): 393-406.

Carrillo C., Cavia M. del M. & Alonso-Torre S., 2012. Role of oleic acid in immune system; mechanism of action; a review. *Nutricion Hospitalaria*, 27(4): 978-990.

Chavarro C., Chu Y., Holbrook C., Isleib T., Bertioli D. *et al.*, 2020. Pod and Seed Trait QTL Identification to Assist Breeding for Peanut Market Preferences. *G3 Genes|Genomes|Genetics*, 10(7): 2297-2315.

Chen W., Jiao Y., Cheng L., Huang L., Liao B. *et al.*, 2016. Quantitative trait locus analysis for pod- and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.). *BMC Genetics*, 17(1): 25.

Chen Y., Ren X., Zheng Y., Zhou X., Huang L. *et al.*, 2017. Genetic mapping of yield traits using RIL population derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.). *Molecular Breeding*, 37(2): 17.

Chu Y., Wu C.L., Holbrook C.C., Tillman B.L., Person G. & Ozias-Akins P., 2011. Marker-Assisted Selection to Pyramid Nematode Resistance and the High Oleic Trait in Peanut. *The Plant Genome*, 4(2):110.

Chukwu S.C., Rafii M.Y., Ramlee S.I., Ismail S.I., Oladosu Y. *et al.*, 2019. Marker-assisted selection and gene pyramiding for resistance to bacterial leaf blight disease of rice (*Oryza sativa* L.). *Biotechnology & Biotechnological Equipment*, 33(1): 440-455.

Collard B.C.Y., Jahufer M.Z.Z., Brouwer J.B. & 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(1-2): 169-196.

Collard B.C.Y. & Mackill D.J., 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1491): 557-572.

Conde S., Rami J.-F., Okello D.K., Sambou A., Muitia A. *et al.*, 2023. The groundnut improvement network for Africa (GINA) germplasm collection: a unique genetic resource for breeding and gene discovery. *G3: Genes|Genomes|Genetics*, 14(1): 244.

Davis J.P., Price K.M., Dean L.L., Sweigart D.S., Cottonaro J.M. & Sanders T.H., 2016. Peanut Oil Stability and Physical Properties Across a Range of Industrially Relevant Oleic Acid/Linoleic Acid Ratios. *Peanut Science*, 43(1): 1-11.

Davis B.I., Agraz C.B., Kline M., Gottschall E., Nolt M. *et al.*, 2021. Measurements of High Oleic Purity in Peanut Lots Using Rapid, Single Kernel Near-Infrared Reflectance Spectroscopy. *Journal of the American Oil Chemists' Society*, 98(6): 621-632.

Desmae H., Janila P., Okori P., Pandey M.K., Monyo E. *et al.*, 2019. Genetics, genomics and breeding of groundnut (*Arachis hypogaea* L.) (B. Morris, Éd.). *Plant Breeding*, 138(4): 425-444.

Dwivedi S.L., Nigam S.N.& Rao R.C.N., 2000. Photoperiod Effects on Seed Quality Traits in Peanut. Crop Sci. 40(5), 1223-1227.

Essandoh D.A., Odong T., Okello D.K., Fonceka D., Nguepjop J. *et al.*, 2022. Quantitative Trait Analysis Shows the Potential for Alleles from the Wild Species Arachis batizocoi and A. duranensis to Improve Groundnut Disease Resistance and Yield in East Africa. *Agronomy*, 12(9): 2202.

Fang Y., Zhang X., Liu H., Wu J., Qi F., Sun Z., Zheng Z., Dong W. & Huang B., 2023. Identification of quantitative trait loci and development of diagnostic markers for growth habit traits in peanut (Arachis hypogaea L.). *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 136(5): 105.

FAOSTAT.Available online: URL (Accessed on 12 January 2024**.**
https://www.fao.org/faostat/en/#home)

Fávero A.P., Simpson C.E., Valls J.F.M. & Vello N.A., 2006. Study of the Evolution of Cultivated Peanut through Crossability Studies among *Arachis ipaënsis* , *A. duranensis* , and *A. hypogaea*. *Crop Science*, 46(4): 1546-1552.

Faye I., Pandey M.K., Hamidou F., Rathore A., Ndoye O., Vadez V. & Varshney R.K., 2015. Identification of quantitative trait loci for yield and yield related traits in groundnut (*Arachis hypogaea* L.) under different water regimes in Niger and Senegal. *Euphytica*, 206(3): 631-647.

Foncéka D., Hodo-Abalo T., Rivallan R., Faye I., Sall M. *et al.*, 2009. Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid. *BMC Plant Biology*, 9(1): 103.

Fonceka D., Tossim H.-A., Rivallan R., Vignes H., Faye I. *et al.*, 2012a. Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breediNg. *BMC Plant Biology*, 12(1): 1-16.

Fonceka D., Tossim H.-A., Rivallan R., Vignes H., Lacut E. *et al.*, 2012b. Construction of chromosome segment substitution lines in peanut (*Arachis hypogaea* L.) using a wild synthetic and QTL mapping for plant morphology. *PLoS One*, 7(11): e48642.

Fonceka D., Tossim H.-A., Rivallan R., Faye I., Leal-Bertioli S., Valls J., Bertioli D., Courtois B. & Rami J.-F., 2013. Interspecific Introgression and Trait Dissection in Peanut (*Arachis hypogaea*). In Proceedings of the Plant and Animal Genome Conference, San Diego, CA, USA,12–16 January 2013.

Fordellone M., Bellincontro A. & Mencarelli F., 2020. Partial least squares discriminant analysis: A dimensionality reduction method to classify hyperspectral data. *Statistica Applicata - Italian Journal of Applied Statistics*, 31(2): 181-200.

Garcia G.M., Stalker H.T. & Kochert G., 1995. Introgression analysis of an interspecific hybrid population in peanuts (Arachis hypogaea L.) using RFLP and RAPD markers. *Genome*, 38(1): 166-176.

Garcia G.M., Stalker H.T., Schroeder E., Lyerly J.H. & Kochert G., 2005. A RAPD-based Linkage Map of Peanut Based on a Backcross Population Between the Two Diploid Species *Arachis stenosperma* and *A. cardenasii*. *Peanut Science*, 32(1): 1-8.

Gautami B., Foncéka D., Pandey M.K., Moretzsohn M.C., Sujay V. *et al.*, 2012a. An International Reference Consensus Genetic Map with 897 Marker Loci Based on 11 Mapping Populations for Tetraploid Groundnut (*Arachis hypogaea* L.). *PLOS ONE*, 7(7): e41213.

Gautami B., Pandey M.K., Vadez V., Nigam S.N., Ratnakumar P. *et al.*, 2012b. Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Molecular Breeding*, 30(2): 757-772.

Gautam T., Dhillon G.S., Saripalli G., Rakhi, Singh V.P. *et al.*, 2020. Marker-assisted pyramiding of genes/QTL for grain quality and rust resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding*, 40(5): 49.

Getahun T., 2021. QTL Analysis Molecular Genetics Diversity and Agro-physiological Characterizations of Heat Tolerance in Chickpea (*Cicer Arietinum* L.): Addis Ababa University, Addis Ababa, Ethiopia.

Gimode D., Chu Y., Dean L., Holbrook C., Fonceka D. & Ozias-Akins P., 2020. Seed Composition Survey of a Peanut CSSL Population Reveals Introgression Lines with Elevated Oleic/Linoleic Profiles. *Peanut Science*, 47(3): 139-149.

Gokidi Y., Bhanu A. & MN S., 2016. Marker Assisted Recurrent Selection: An Overview. *Advances in Life Sciences*, 5(17): 6493-6499.

Goodwin M., Halvorsen K., Jiao L., Knausgård K., Martin A. *et al.*, 2022. Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook. *ICES Journal of Marine Science*, 79(2):319-336.

Govindarajan K.N., Kandala V. & Subbiah J., 2009. NIR Reflectance Spectroscopy for Nondestructive Moisture Content Determination in Peanut Kernels. *Transactions of the ASABE*, 52(5):1661-1666.

Gowda M., Motagi B., Naidu G., B D. & R S., 2002. GPBD 4: A Spanish bunch Groundnut Genotype Resistant to Rust and Late leaf spot. *International Arachis Newsletter*, 22: 29-32.

Grosso N.R., Nepote V. & Guzmán C.A., 2000. Chemical Composition of Some Wild Peanut Species (*Arachis hypogaea* L.) Seeds. *Journal of Agricultural and Food Chemistry*, 48(3): 806-809.

Guo Y., Khanal S., Tang S., Bowers J.E., Heesacker A.F. *et al.*, 2012. Comparative mapping in intraspecific populations uncovers a high degree of macrosynteny between A- and B-genome diploid species of peanut. *BMC Genomics*, 13(1): 608.

Guo J., Liu N., Li W., Wu B., Chen H. *et al.*, 2021. Identification of two major loci and linked marker for oil content in peanut (*Arachis hypogaea* L.). *Euphytica*, 217(2): 29.

Hancock, J. F. (2022). Portugal and the South Atlantic Exchange, in J. F. Hancock (Ed.), *World Agriculture Before and After 1492: Legacy of the Columbian Exchange*. Springer International Publishing. Cham, 83-96 p.

Halward T., Stalker H.T. & Kochert G., 1993. Development of an RFLP linkage map in diploid peanut species. *Theoretical and Applied Genetics*, 87(3): 379-384.

Harbau A.I. & Sanusi H., 2023. Comparative analysis of different varieties of ground nut (*Arachis hypogea* L) yield to different row spacing in Bichi and Tsanyawa local government areas of Kano State Nigeria. *Watari Multidisciplinary Journal of Science & Technology Education*, 7(1): 53-62.

Harch B.D., Basford K.E., DeLacy I.H., Lawrence P.K. & Cruickshank A., 1995. Patterns of diversity in fatty acid composition in the Australian groundnut germplasm collection. *Genetic Resources and Crop Evolution*, 42(3): 243-256.

Harris P. & James A.T., 1969. The effect of low temperatures on fatty acid biosynthesis in plants. *Biochemical Journal*, 112(3): 325-330.

Hasan N., Choudhary S., Naaz N., Sharma N. & Laskar R.A., 2021. Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *Journal of Genetic Engineering & Biotechnology*, 19(1): 128.

Herselman L., Thwaites R., Kimmins F.M., Courtois B., van der Merwe P.J.A. & Seal S.E., 2004. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. *Theoretical and Applied Genetics*, 109(7): 1426-1433.

Holaday C.E. & Pearson J.L., 1974. Effects of genotype and production area on the fatty acid composition, total oil and total protein in peanuts. *Journal of Food Science*, 39(6): 1206-1209.

Holbrook C.C., Anderson W.F. & Pittman R.N., 1993. Selection of a Core Collection from the U.S. Germplasm Collection of Peanut. *Crop Science*, 33(4): 859-861.

Holbrook C.C., Dong W., 2005. Development and Evaluation of a Mini Core Collection for the U.S. Peanut Germplasm Collection. Crop Science, 45(4):1540-1544.

Holbrook C.C., Ozias-Akins P., Chu Y., Culbreath A.K., Kvien C.K. & Brenneman T.B., 2017. Registration of 'TifNV-High O/L' Peanut. *Journal of Plant Registrations*, 11(3): 228-230.

Hong Y., Chen X., Liang X., Liu H., Zhou G., Li S., Wen S., Holbrook C.C. & Guo B., 2010. A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. *BMC Plant Biology*, 10(1): 17.

Hospital F, Chevalet C & Mulsant P (1992). Using markers in gene introgression breeding programms. Genetics, 132(1): 1199-1210.

Hu X.H., Zhang S.Z., Miao H.R., Cui F.G., Shen Y. *et al.*, 2018. High-Density Genetic Map Construction and Identification of QTLs Controlling Oleic and Linoleic Acid in Peanut using SLAF-seq and SSRs. *Scientific Reports*, 8(1): 5479.

Huang H., Liu L. & Ngadi M.O., 2014. Recent Developments in Hyperspectral Imaging for Assessment of Food Quality and Safety. *Sensors*, 14(4): 7248-7276.

Huang B.E., Verbyla K.L., Verbyla A.P., Raghavan C., Singh V.K., Gaur P., Leung H., Varshney R.K. & Cavanagh C.R., 2015a. MAGIC populations in crops: current status and future prospects. *Theoretical and Applied Genetics*, 128(6): 999-1017.

Huang L., He H., Chen W., Ren X., Chen Y. *et al.*, 2015b. Quantitative trait locus analysis of agronomic and quality-related traits in cultivated peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, 128(6): 1103-1115.

Huang B., Qi F., Sun Z., Miao L., Zhang Z. *et al.*, 2019. Marker-assisted backcrossing to improve seed oleic acid content in four elite and popular peanut (*Arachis hypogaea* L.) cultivars with high oil content. *Breeding Science*, 69(2): 234-243.

IBPGR and ICRISAT. 1992. Descriptors for groundnut. International Board for plant Genetic Resources, Rome, Italy; International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India

Janila P., Nigam S.N., Pandey M.K., Nagesh P. & Varshney R.K., 2013. Groundnut improvement: use of genetic and genomic tools. *Frontiers in Plant Science*, 4:23.

Janila P., Variath M.T., Pandey M.K., Desmae H., Motagi B.N. *et al.*, 2016. Genomic Tools in Groundnut Breeding Program: Status and Perspectives. *Frontiers in Plant Science*, 7:289.

Janila P., Manohar S.S., Deshmukh D.B., Chaudhari S., Papaiah V. & Variath M.T., 2018.Standard Operating Procedures for Groundnut Breeding and Testing. ICRISAT, https://oar.icrisat.org/10653/.

Jiang H.F., Ren X.P., Liao B.S., Huang J., Lei Y. & Chen B.Y., 2008. Peanut core collection established in china and compared with ICRISAT mini core collection. *Acta Agron Sin*, 34: 25-30.

Jiang H.-F., Ren X.-P., Zhang X.-J., Huang J.-Q., Lei Y., Yan L.-Y., Liao B.-S., Upadhyaya H.D. & Holbrook C.C., 2010. Comparison of Genetic Diversity Based on SSR Markers Between Peanut Mini Core Collections from China and ICRISAT. *Acta Agronomica Sinica*, 36(7): 1084-1091.

Jiang R., 2002. Nut and Peanut Butter Consumption and Risk of Type 2 Diabetes in Women. *JAMA*, 288(20): 2554.

Jin G., Liu N., Yu B., Jiang Y., Luo H. *et al.*, 2023. Identification and Pyramiding Major QTL Loci for Simultaneously Enhancing Aflatoxin Resistance and Yield Components in Peanut. *Genes*, 14(3): 625.

Kandala C.V.K., Butts C.L. & Lamb M.C., 2008. Moisture content determination for in-shell peanuts with a low-cost impedance analyzer and capacitor sensor. *Transactions of the ASABE*, 51(4): 1377-1381.

Kassie F.C., Nguepjop J.R., Ngalle H.B., Assaha D.V.M., Gesese M.K. *et al.*, 2023. An Overview of Mapping Quantitative Trait Loci in Peanut (*Arachis hypogaea* L.). *Genes*, 14(6): 1176.

Kassie F.C., Chaix G., Ngalle H.B., Seye M., Fall C. *et al.*, 2024. Application of near-infrared spectroscopy for fast germplasm analysis and classification in multi-environment using intact-seed peanut (*Arachis hypogaea* L.). *Oil Crop Science*, 9(2): 132-141.

Kayam G., Brand Y., Faigenboim-Doron A., Patil A., Hedvat I. & Hovav R., 2017. Fine-Mapping the Branching Habit Trait in Cultivated Peanut by Combining Bulked Segregant Analysis and High-Throughput Sequencing. *Frontiers in Plant Science*, 8:467.

Kenga R., Njoya A., M'biandoum M., 2005. Analysis of Constraints to Agricultural Production in the Sudano Savanna Zone of Cameroon and Implication for Research Priority Setting. Tropicultura, 23:19-99.

Khedikar Y.P., Gowda M.V.C., Sarvamangala C., Patgar K.V., Upadhyaya H.D. & Varshney R.K., 2010. A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, 121(5): 971-984.

Khedikar Y., Pandey M.K., Sujay V., Singh S., Nayak S.N. *et al.*, 2017. Identification of main effect and epistatic quantitative trait loci for morphological and yield-related traits in peanut (*Arachis hypogaea* L.). *Molecular Breeding*, 38(1): 1-12.

Khera P., Pandey M.K., Mallikarjuna N., Sriswathi M., Roorkiwal M. *et al.*, 2019. Genetic imprints of domestication for disease resistance, oil quality, and yield component traits in groundnut (*Arachis hypogaea* L.). *Molecular Genetics and Genomics*, 294(2): 365-378.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., et al. (2012). Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms, in F. Pompanon & A. Bonin (Eds.), *Data Production and Analysis in Population Genomics: Methods and Protocols*. Humana Press, 67-89 p.

Knoll J.E., Ramos M.L., Zeng Y., Holbrook C.C., Chow M., Chen S., Maleki S., Bhattacharya A. & Ozias-Akins P., 2011. TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). *BMC Plant Biology*, 11(1): 81.

Kochert G., Stalker H.T., Gimenes M., Galgaro L., Lopes C.R. & Moore K., 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (*Leguminosae*). *American journal of botany (USA)*.

Kolekar R., Sukruth M., Shirasawa K., Nadaf H., Motagi B., Sattigarahalli L., Patil P. & Bhat R., 2017. Marker-assisted backcrossing to develop foliar disease-resistant genotypes in TMV 2 variety of peanut (*Arachis hypogaea* L.). *Plant Breeding*, *136*(6):948-953.

Kosambi, D. D. (2016). The estimation of map distances from recombination values, in R. Ramaswamy (Ed.), *D.D. Kosambi: Selected Works in Mathematics and Statistics*. Springer India, New Delhi, 125-130 p.

Krapovickas A. & Gregory W.C., 1994. Taxonomía del género *Arachis* (*Leguminosae*). *Bonplandia*, 8(1-4): 1-186.

Kumari, M., Dubey, A. K., Kumar, R., & Kumar, A. (2024). Chapter 18 - Marker-assisted selection in plant breeding for stress tolerance, in M. A. Ahanger, J. A. Bhat, P. Ahmad, & R. John (Eds.), *Improving Stress Resilience in Plants*. Academic Press, 371-387 p.

Kumari V., Gowda M.V.C., Tasiwal V., Pandey M.K., Bhat R.S., Mallikarjuna N., Upadhyaya H.D. & Varshney R.K., 2014. Diversification of primary gene pool through introgression of resistance to foliar diseases from synthetic amphidiploids to cultivated groundnut (*Arachis hypogaea* L.). *The Crop Journal*, 2(2): 110-119.

Kumari V., Gowda M.V.C. & Yeri S.B., 2020. Utilization of Advanced Backcross Population Derived from Synthetic Amphidiploid for Dissecting Resistance to Late Leaf Spot in Peanut (*Arachis hypogaea* L.). *Tropical Plant Biology*, 13(1): 50-61.

Kumawat G., Kumawat C.K., Chandra K., Pandey S., Chand S., Mishra U.N., Lenka D. & Sharma R., 2020. Insights into Marker Assisted Selection and Its Applications in Plant Breeding. IntechOpen,

Kuznetsova A., Brockhoff P.B. & Christensen R.H.B., 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13):1-26.

Lander E.S. & Botstein D., 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1): 185-199.

Leal-Bertioli S.C., José A.C.V., Alves-Freitas D.M., Moretzsohn M.C., Guimarães P.M. *et al.*, 2009. Identification of candidate genome regions controlling disease resistance in Arachis. *BMC Plant Biology*, 9(1): 112.

Leal-Bertioli S.C.M., Moretzsohn M.C., Roberts P.A., Ballén-Taborda C., Borba T.C.O. *et al.*, 2016. Genetic Mapping of Resistance to Meloidogyne arenaria in *Arachis stenosperma*: A New Source of Nematode Resistance for Peanut. *G3: Genes, Genomes, Genetics*, 6(2): 377-390.

Lesnooff, M. (2021) *R package rnirs: Dimension reduction, Regression and Discrimination for Chemometrics*. Available at: https://github.com/mlesnoff/rnirs. CIRAD, UMR SELMET, Montpellier, France.

Li L., Yang X., Cui S., Meng X., Mu G. *et al.*, 2019a. Construction of High-Density Genetic Map and Mapping Quantitative Trait Loci for Growth Habit-Related Traits of Peanut (*Arachis hypogaea* L.). *Frontiers in Plant Science*, 10:745.

Li Z., Tang X., Shen Z., Yang K., Zhao L., Li Y., 2019b. Comprehensive comparison of multiple quantitative near-infrared spectroscopy models for Aspergillus flavus contamination detection in peanut. J. *Sci. Food Agric*. 99:5671-5679.

Liang Y., Baring M., Wang S. & Septiningsih E.M., 2017. Mapping QTLs for Leafspot Resistance in Peanut Using SNP-Based Next-Generation Sequencing Markers. *Plant Breeding and Biotechnology*, 5(2): 115-122.

Liang Y., Baring M.R. & Septiningsih E.M., 2018. Mapping of Quantitative Trait Loci for Yield and Grade Related Traits in Peanut (*Arachis hypogaea* L.) Using High-Resolution SNP Markers. *Plant Breeding and Biotechnology*, 6(4): 454-462.

Lin M., Long M., Li G., Chen X., Zheng J., Li C. & Kan J., 2016. Analysis of Peanut Using Near-Infrared Spectroscopy and Gas Chromatography–Mass Spectrometry: Correlation of Chemical Components and Volatile Compounds. *International Journal of Food Properties*, 19(3): 508-520.

Liu N., Guo J., Zhou X., Wu B., Huang L. *et al.*, 2020. High-resolution mapping of a major and consensus quantitative trait locus for oil content to a ~0.8-Mb region on chromosome A08 in peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, 133(1): 37-49.

Liu L., Jiang H., Chen Q., 2022. Rapid determination of acidity index of peanut during storage by a portable near-infrared spectroscopy system. *Infrared Phys. Technol*. 127: 104472.

Lu J. & Pickersgill B., 1993. Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L. - *Leguminosae*). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 85(5): 550-560.

Lu Q., Li H., Hong Y., Zhang G., Wen S. *et al.*, 2018. Genome Sequencing and Analysis of the Peanut B-Genome Progenitor (*Arachis ipaensis*). *Frontiers in Plant Science*, 9:604.

Luo H., Ren X., Li Z., Xu Z., Li X. *et al.*, 2017. Co-localization of major quantitative trait loci for pod size and weight to a 3.7 cM interval on chromosome A05 in cultivated peanut (*Arachis hypogaea* L.). *BMC Genomics*, 18(1): 58.

Luo H., Guo J., Ren X., Chen W., Huang L. *et al.*, 2018. Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 131(2): 267-282.

Lv J., Liu N., Guo J., Xu Z., Li X. *et al.*, 2018. Stable QTLs for Plant Height on Chromosome A09 Identified from Two Mapping Populations in Peanut (*Arachis hypogaea* L.). *Frontiers in Plant Science*, 9:684.

Mallikarjuna N., Senapathy S. & Hoisington D., 2010. Development of new sources of tetraploid Arachis to broaden the genetic base of cultivated groundnut (*Arachis hypogaea* L.). *Genetic Resources and Crop Evolution*, 58: 889-907.

Mallikarjuna N., Jadhav D.R., Reddy K., Husain F. & Das K., 2012. Screening new Arachis amphidiploids, and autotetraploids for resistance to late leaf spot by detached leaf technique. *European Journal of Plant Pathology*, 132(1): 17-21.

Manley, M., & Baeten, V. (2018). Spectroscopic Technique: Near Infrared (NIR) Spectroscopy. In D.-W. Sun (Ed.), *Modern Techniques for Food Authentication* (2nd). Academic Press, 51-102 p.

Massa A.N., Sobolev V.S., Faustinelli P.C., Tallury S.P., Stalker H.T., Lamb M.C. & Arias R.S., 2024. Genetic diversity, disease resistance, and environmental adaptation of Arachis duranensis L.: New insights from landscape genomics. *PLOS ONE*, 19(4): e0299992.

Meng L., Li H., Zhang L. & Wang J., 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal*, 3(3): 269-283.

Milla, S.R. (2003) Relationships and utilization of *Arachis* germplasm in peanut improvement. Ph.D. Thesis, North Carolina State University, North Carolina, USA.

Milla S.R., Isleib T.G. & Stalker H.T., 2005. Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. *Genome*, 48(1): 1-11.

Miller P. A., J. C, Williams H. F., Robinson and R. E., Comstock 1958. Estimates genotypic and environmental variances and covariances in upland cotton and their implications in selection. *Agronomic Journal*, 50: 126-131.

Mohammed A., Tesso B., Ojiewo C. & Ahmed S., 2019. Assessment of Genetic Variability and Heritability of Agronomic Traits of Ethiopian Chickpea (*Cicerarietinum* L) Landraces. *Black Sea Journal of Agriculture*, 2(1): 10-15.

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR and Sasaki T (1997) Genome mapping, molecular markers and marker-assistaed selection in crop plants. Mol Breeding, 3: 87-103.

Mondal S. & Badigannavar A.M., 2019. Identification of major consensus QTLs for seed size and minor QTLs for pod traits in cultivated groundnut (*Arachis hypogaea* L.). *3 Biotech*, 9(9): 347.

Moretzsohn M. de C., Hopkins M.S., Mitchell S.E., Kresovich S., Valls J.F.M. & Ferreira M.E., 2004. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. *BMC Plant Biology*, 4: 11.

Moretzsohn M.C., Leoi L., Proite K., Guimarães P.M., Leal-Bertioli S.C.M. *et al.*, 2005. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (*Fabaceae*). *Theoretical and Applied Genetics*, 111(6): 1060-1071.

Moretzsohn M.C., Barbosa A.V., Alves-Freitas D.M., Teixeira C., Leal-Bertioli S.C. *et al.*, 2009. A linkage map for the B-genome of *Arachis* (*Fabaceae*) and its synteny to the A-genome. *BMC Plant Biology*, 9(1): 40.

Moretzsohn M.C., Gouvea E.G., Inglis P.W., Leal-Bertioli S.C.M., Valls J.F.M. & Bertioli D.J., 2013. A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. *Annals of Botany*, 111(1): 113-126.

Motagi B.N., Bhat R.S., Pujer S., Nayak S.N., Pasupaleti J. *et al.*, 2022. Genetic Enhancement of Groundnut: Current Status and Future Prospects, in *Accelerated Plant Breeding, Volume 4*, In S. S. Gosal & S. H. Wani (Eds.). Springer International Publishing, Cham, 63-110 p.

Mozingo R.W., Coffelt T.A. & Wynne J.C., 1988. Market Grade Effects on Fatty Acid Composition of Five Peanut Cultivars. *Agronomy Journal*, 80(1): 73-75.

Nadeem M.A., Nawaz M.A., Shahid M.Q., Doğan Y., Comertpay G. *et al.*, 2018. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment*, 32(2): 261-285.

Nagy E.D., Guo Y., Tang S., Bowers J.E., Okashah R.A. *et al.*, 2012. A high-density genetic map of *Arachis duranensis,* a diploid ancestor of cultivated peanut. *BMC Genomics*, 13(1): 469.

Nawade B., Mishra G.P., Radhakrishnan T., Dodia S.M., Ahmad S., Kumar A., Kumar A. & Kundu R., 2018. High oleic peanut breeding: Achievements, perspectives, and prospects. *Trends in Food Science & Technology*, 78: 107-119.

Nawade B., Mishra G.P., Radhakrishnan T., Sangh C., Dobariya J.R. & Kundu R., 2019. Development of high oleic peanut lines through marker-assisted introgression of mutant ahFAD2 alleles and its fatty acid profiles under open-field and controlled conditions. *3 Biotech*, 9(6): 243.

Nayak S.N., Singh V.K. & Varshney R.K., 2017. Marker-Assisted Selection, in *Encyclopedia of Applied Plant Sciences*, Elsevier, 183-197 p.

Nguepjop J.R., Tossim H.-A., Bell J.M., Rami J.-F., Sharma S., Courtois B., Mallikarjuna N., Sane D. & Fonceka D., 2016. Evidence of Genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthetized Allotetraploid Hybrids. *Frontiers in Plant Science*, 7:1635.

Nian L., Chen H., Huai D., Xia F., Huang L. *et al.*, 2019. Four QTL clusters containing major and stable QTLs for saturated fatty acid contents in a dense genetic map of cultivated peanut (*Arachis hypogaea* L.). *Molecular Breeding*, 39(2):15.

Nigam S.N., 2014. Groundnut at a glance. ICRISAT, Patancheru,121 p.

Norden A.J., Gorbet D.W., Knauft D.A. & Young C.T., 1987. Variability in Oil Quality Among Peanut Genotypes in the Florida Breeding Program 1. *Peanut science*, 14(1): 7-11.

Ntare B.R., Diallo A.T., Ndjeunga J. & Waliyar F., 2008.Groundnut Seed Production Manual. InternationalCrops Research Institute for the Semi-Arid Tropics, https://oar.icrisat.org/420/.

Ogbole O.O., Akin-Ajani O.D., Ajala T.O., Ogunniyi Q.A., Fettke J. & Odeku O.A., 2023. Nutritional and pharmacological potentials of orphan legumes: Subfamily faboideae. *Heliyon*, 9(4): e15493.

Ouellette L.A., Reid R.W., Blanchard S.G. & Brouwer C.R., 2018. LinkageMapView—rendering high-resolution linkage and QTL maps (O. Stegle, Éd.). *Bioinformatics*, 34(2): 306-307.

Pan Y., Zhuang Y., Liu T., Chen H., Wang L., Varshney R.K., Zhuang W. & Wang X., 2023. Deciphering peanut complex genomes paves a way to understand its origin and domestication. *Plant Biotechnology Journal*, 21(11): 2173-2181.

Pandey M.K., Monyo E., Ozias-Akins P., Liang X., Guimarães P. *et al.*, 2012. Advances in *Arachis* genomics for peanut improvement. *Biotechnology Advances*, 30(3): 639-651.

Pandey M.K., Guo B., Holbrook C.C., Janila P., Zhang X., Bertioli D.J., Isobe S., Liang X. & Varshney R.K., 2014a. Molecular markers, genetic maps and QTLs for molecular breeding in peanut. *Genetics, genomics and breeding of peanuts*, 61-113 p.

Pandey M.K., Wang M.L., Qiao L., Feng S., Khera P. *et al.*, 2014b. Identification of QTLs associated with oil content and mapping FAD2 genes and their relative contribution to oil quality in peanut (*Arachis hypogaea*L.). *BMC Genetics*, 15(1): 133.

Pandey M.K., Roorkiwal M., Singh V.K., Ramalingam A., Kudapa H., Thudi M., Chitikineni A., Rathore A. & Varshney R.K., 2016. Emerging Genomic Tools for Legume Breeding: Current Status and Future Prospects. *Frontiers in Plant Science*, 7:455.

Pandey M.K., Gangurde S.S., Sharma V., Pattanashetti S.K., Naidu G.K. *et al.*, 2020a. Improved Genetic Map Identified Major QTLs for Drought Tolerance- and Iron Deficiency Tolerance-Related Traits in Groundnut. *Genes*, 12(1): 37.

Pandey M.K., Pandey A.K., Kumar R., Nwosu C.V., Guo B. *et al.*, 2020b. Translational genomics for achieving higher genetic gains in groundnut. *Theoretical and Applied Genetics*, 133(5): 1679-1702.

Panero J.S., Silva H.E.B. da, Panero P.S., Smiderle O.J., Panero F.S., Faria F.S.E.D.V., Rodriguez A.F.R., 2018. Separation of Cultivars of Soybeans by Chemometric Methods Using Near Infrared Spectroscopy. *Journal of Agricultural Science,* 10(4): 351.

Panero F.D.S., Smiderle O., Panero J.S., Faria F.S.D.V., Panero P.D.S., Rodriguez A.F.R., 2022. Non-Destructive Genotyping of Cultivars and Strains of Sesame through NIR Spectroscopy and Chemometrics. *Biosensors*, 12(2):69.

Pankaj Y.K., Kumar R., Pal L., Gill K.S., Nagarajan R., Kumar V. & Panigrahi S., 2022. Performance and yield stability of doubled haploid population of wheat (*Triticum aestivum* L.) under high-temperature regime. *Cereal Research Communications*, 50(4): 1185-1203.

Parilli-Moser I., Hurtado-Barroso S., Guasch-Ferré M. & Lamuela-Raventós R.M., 2022. Effect of Peanut Consumption on Cardiovascular Risk Factors: A Randomized Clinical Trial and Meta-Analysis. *Frontiers in Nutrition*, 11: 423.

Pascual L., Desplat N., Huang B., Desgroux A., Bruguier L. *et al.*, 2014. Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnology Journal*,13(4):13.

Patil A.S., Popovsky S., Levy Y., Chu Y., Clevenger J., Ozias-Akins P. & Hovav R., 2018. Genetic insight and mapping of the pod constriction trait in Virginia-type peanut. *BMC Genetics*, 19(1): 93.

Phuc T., Hoa T.T.T., Quynh C.T.T., Dzung N.H. & Hung P.N., 2023. Rapid Identification of the Geographical Origin of Black Pepper in Vietnam Using Near-Infrared Spectroscopy and Chemometrics. *Journal of Science and Technology: Engineering and Technology for Sustainable Development*, 33(1): 1-8.

Phung B.T.M., Van Hoa N., Tanh N.T.N. & Chiem N.H., 2023. Indigenous Knowledge of Peanut Cultivation of the Ethnic Khmer Resident in the Mountainous Area in An Giang Province, Vietnam. *International Journal on Advanced Science, Engineering & Information Technology*, 13(1):10.

Qin H., Feng S., Chen C., Guo Y., Knapp S. *et al.*, 2012. An integrated genetic linkage map of cultivated peanut (*Arachis hypogaea* L.) constructed from two RIL populations. *Theoretical and Applied Genetics*, 124(4): 653-664.

R Core Team (2021) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. Available at: https://www.R-project.org.

Rabanera J.D., Guzman J.D. & Yaptenco K.F., 2021. Rapid and Non-destructive measurement of moisture content of peanut (*Arachis hypogaea* L.) kernel using a near-infrared hyperspectral imaging technique. *Journal of Food Measurement and Characterization*, 15(4): 3069-3078.

Radhakrishnan, T., Kona, P., Ajay, B. C., & Kumar, N. (2022a). Groundnut breeding, in D. K. Yadava, H. K. Dikshit, G. P. Mishra, & S. Tripathi (Eds.), *Fundamentals of field crop breeding*. Springer Nature., Singapore, 837-906 p.

Radhakrishnan, T., Rathnakumar, A. L., Mahatma, M. K., Chandramohan, S., & Patel, S. (2022b). Genetic Resources of Groundnut, in: P. M. Priyadarshan & S. M. Jain (Eds.), *Cash Crops: Genetic Diversity, Erosion, Conservation and Utilization*. Springer International Publishing, Cham, 341-406 p.

Raj, S. P., Solomon, P. R., & Thangaraj, B. (2022). Fabaceae, in: S. P. Raj, P. R. Solomon, & B. Thangaraj (Eds.), *Biodiesel from Flowering Plants*. Springer, Singapore, 291-363 p.

Rajwade Y., Swain D., Tiwari K., Mohanty U.C. & Goswami P., 2015. Evaluation of Field Level Adaptation Measures Under the Climate Change Scenarios in Rice Based Cropping System in India. *Environmental Processes*, 2:669-687.

Ramos M.L., Fleming G., Chu Y., Akiyama Y., Gallo M. & Ozias-Akins P., 2006. Chromosomal and phylogenetic context for conglutin genes in *Arachis* based on genomic sequence. *Molecular genetics and genomics: MGG*, 275(6): 578-592.

Ravi K., Vadez V., Isobe S., Mir R.R., Guo Y. *et al.*, 2011. Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachishypogaea* L.). *Theoretical and Applied Genetics*, 122(6): 1119-1132.

Recal F. & Demirel T., 2021. Comparison of machine learning methods in predicting binary and multi-class occupational accident severity. *Journal of Intelligent & Fuzzy Systems*, 40(6): 10981-10998.

Robledo G., Seijo G., 2010. Species relationships among the wild B genome of *Arachis species* (section *Arachis*) based on FISH mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. *Theoretical and Applied Genetics*, 121: 1033-1046.

Sambou A., 2017. Genetic studies and mapping of beneficial alleles from wild species to improve agronomic traits in cultivated groundnut (*Arachis hypogaea L.*); University of Legon: Accra, Ghana,1119-1132.

Sampaio P. & Brites C., 2021. Near-Infrared Spectroscopy and Machine Learning: Analysis and Classification Methods of Rice. In: Integrative Advances in Rice Research. IntechOpen, 33 p.

Sarvamangala C., Gowda M.V.C. & Varshney R.K., 2011. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). *Field Crops Research*, 122(1): 49-59.

Seijo J.G., Lavia G.I., Fernández A., Krapovickas A., Ducasse D. & Moscone E.A., 2004. Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (*Leguminosae*). *American Journal of Botany*, 91(9): 1294-1303.

Seijo G., Lavia G.I., Fernández A., Krapovickas A., Ducasse D.A., Bertioli D.J. & Moscone E.A., 2007. Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH. *American Journal of Botany*, 94(12): 1963-1971.

Servin B., Martin O.C., Mézard M. & Hospital F., 2004. Toward a Theory of Marker-Assisted Gene Pyramiding. *Genetics*, 168(1): 513-523.

Shang H., Shang L., Wu J., Xu Z., Zhou S., Wang Z., Wang H. & Yin J., 2023. NIR spectroscopy combined with 1D-convolutional neural network for breast cancerization analysis and diagnosis. *Spectrochemical Acta Part A: Molecular and Biomolecular Spectroscopy*, 287: 121990.

Sharma S, Upadhyaya HD, Varshney RK and Gowda CLL (2013) Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. *Frontiers in Plant Science,* 4:309.

Shasidhar Y., Vishwakarma M.K., Pandey M.K., Janila P., Variath M.T., Manohar S.S., Nigam S.N., Guo B. & Varshney R.K., 2017. Molecular Mapping of Oil Content and Fatty Acids Using Dense Genetic Maps in Groundnut (*Arachis hypogaea* L.). *Frontiers in Plant Science*, 8: 794.

Shasidhar Y., Variath M.T., Vishwakarma M.K., Manohar S.S., Gangurde S.S. *et al.*, 2020. Improvement of three popular Indian groundnut varieties for foliar disease resistance and high oleic acid using SSR markers and SNP array in marker-assisted backcrossing. *The Crop Journal*, 8(1): 1-15.

Shirasawa K., Bertioli D.J., Varshney R.K., Moretzsohn M.C., Leal-Bertioli S.C.M. *et al.*, 2013. Integrated Consensus Map of Cultivated Peanut and Wild Relatives Reveals Structures of the A and B Genomes of *Arachis* and Divergence of the Legume Genomes. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes*, 20(2): 173-184.

Simpson C.E., Nelson S.C., Starr J.L., Woodard K.E. & Smith O.D., 1993. Registration of TxAG-6 and TxAG-7 Peanut Germplasm Lines. *Crop Science*, 33(6): 1418-1418.

Simpson C.E., 2001. Use of Wild *Arachis* Species/Introgression of Genes into *A. hypogaea* L. *Peanut Science*, 28(2): 114-116.

Simpson C.E., Starr J.L., Baring M.R., Burow M.D., Cason J.M. & Wilson J.N., 2013. Registration of 'Webb' Peanut. *Journal of Plant Registrations*, 7(3): 265-268.

Singh A.K. & Moss J.P., 1982. Utilization of wild relatives in genetic improvement of *Arachis hypogaea* L. *Theoretical and Applied Genetics*, 61(4): 305-314.

Singh B.D. & Singh A.K., 2015. Marker-Assisted Plant Breeding: Principles and Practices. Springer India, New Delhi.514 p.

Singh A. K & Nigam S.N (2016). *Arachis* gene pools and genetic improvement in groundnut, in: Rajpal VR, Rao SM, Raina SN (Eds.) Gene Pool Diversity and Crop Improvement, Springer International Publishing, Switzerland, 17-77 p.

Singh T., Garg N.M., Iyengar S.R.S. & Singh V., 2023. Near-infrared hyperspectral imaging for determination of protein content in barley samples using convolutional neural network. *Journal of Food Measurement and Characterization*, 17:1-13.

Snee R.D., 1977. Validation of Regression Models: Methods and Examples. *Technometrics*, 19(4): 415-428.

Song L., Wang R., Yang X., Zhang A. & Liu D., 2023. Molecular Markers and Their Applications in Marker-Assisted Selection (MAS) in Bread Wheat (*Triticum aestivum* L.). *Agriculture*, 13(3): 642.

Song W., Wang H., Maguire P., Nibouche O., 2018. Nearest clusters based partial least squares discriminant analysis for the classification of spectral data. Analytica Chimica Acta, 1009: 27-38.

Stalker H.T., 2017. Utilizing Wild Species for Peanut Improvement. *Crop Science*, 57(3): 1102-1120.

Statista 2023. https://www.statista.com/statistics/263933/production-of-vegetable-oils-worldwide-since-2000/,n.d.Acceced March, 2023.

Sujay V., Gowda M.V.C., Pandey M.K., Bhat R.S., Khedikar Y.P. *et al.*, 2012. Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Molecular Breeding*, 30(2): 773-788.

Sundaram J., Kandala C.V., Holser R.A., Butts C.L. & Windham W.R., 2010. Determination of In-Shell Peanut Oil and Fatty Acid Composition Using Near-Infrared Reflectance Spectroscopy. *Journal of the American Oil Chemists' Society*, 87(10): 1103-1114.

Sun Z., Qi F., Liu H., Qin L., Xu J. *et al.*, 2021. QTL Mapping of Quality Related Traits in Peanut Using Whole-Genome Resequencing. In Review, preprint,10:177-184.

Smith, C., & Reeves, A. (2023). European Exploration and Colonization. *He Huaka'i Honua: Journeys in World History I, to 1500 CE Honolulu CC HIST 151*.

Tang Y., Qiu X., Hu C., Li J., Wu L. *et al.*, 2022. Breeding of a new variety of peanut with high-oleic-acid content and high-yield by marker-assisted backcrossing. *Molecular Breeding*, 42(7): 42.

Tao F., Yao H., Hruska Z., Liu Y., Rajasekaran K. & Bhatnagar D., 2019. Use of Visible-Near-Infrared (Vis-NIR) Spectroscopy to Detect Aflatoxin B1 on Peanut Kernels. *Applied Spectroscopy*, 73(4): 415-423.

Temga J.P., Sababa E., Mamdem L.E., Bijeck M.L.N., Azinwi P.T. *et al.*, 2021. Rare earth elements in tropical soils, Cameroon soils (Central Africa). *Geoderma Regional*, 25: 23.

Teres S., Barceló-Coblijn G., Benet M., Alvarez R., Bressani R., Halver J. & Escriba P., 2008. Oleic acid content is responsible for the reduction in blood pressure induced by olive oil. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37): 13811-13816.

Tian W., Zang L., Nie L., Li L., Zhong L., Guo X., Huang S. & Zang H., 2023. Structural Analysis and Classification of Low-Molecular-Weight Hyaluronic Acid by Near-Infrared Spectroscopy: A Comparison between Traditional Machine Learning and Deep Learning. *Molecules*, 28(2): 809.

Tossim H.-A., Nguepjop J.R., Diatta C., Sambou A., Seye M., Sane D., Rami J.-F. & Fonceka D., 2020. Assessment of 16 Peanut (*Arachis hypogaea* L.) CSSLs Derived from an Interspecific Cross for Yield and Yield Component Traits: QTL Validation. *Agronomy*, 10(4): 583.

Upadhyaya H.D. & Gowda C.L.L., 2009.Managing and Enhancing the Use of Germplasm – Strategies and Methodologies. International Crops Research Institute for the Semi-Arid Tropics, https://oar.icrisat.org/1316/.

Upadhyaya H.D., Bramel P.J., Ortiz R. & Singh S., 2002. Developing a Mini Core of Peanut for Utilization of Genetic Resources. *Crop Science*, 42(6): 2150-2156.

Upadhyaya H.D., 2005. Variability for Drought Resistance Related Traits in the Mini Core Collection of Peanut. *Crop Science*, 45(4): 1432-1440.

Upadhyaya H.D., Furman B.J., Dwivedi S.L., Udupa S.M., Gowda C.L.L., Baum M., Crouch J.H., Buhariwalla H.K. & Singh S., 2006. Development of a Composite Collection for Mining Germplasm Possessing Allelic Variation for Beneficial Traits in Chickpea. *Plant Genetic Resources*, 4(01): 13-19.

Upadhyaya H.D., Dwivedi S.L., Nadaf H.L. & Singh S., 2011. Phenotypic diversity and identification of wild *Arachis* accessions with useful agronomic and nutritional traits. *Euphytica*, 182:103-115.

Valls J.F.M. & Simpson C.E., 2005. New Species of *Arachis* (*leguminosae*) from Brazil, Paraguay and Bolivia. *Bonplandia*, 14(1/2): 35-63.

Van Ooijen, J. W. (2006) *JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations*. Kyazma BV, Wageningen, Netherlands.

Varshney R.K., Bertioli D.J., Moretzsohn M.C., Vadez V., Krishnamurthy L. *et al.*, 2009. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, 118(4): 729-739.

Varshney R.K., Pandey M.K., Janila P., Nigam S.N., Sudini H. *et al.*, 2014. Marker-assisted introgression of a QTL region to improve rust resistance in three elite and popular varieties of peanut (*Arachis hypogaea* L.). *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 127(8): 1771-1781.

Varshney R.K., Pandey M.K., Puppala N., 2017. The Peanut Genome: An Introduction, in: Varshney, R.K., Pandey, M.K., Puppala, N. (Eds.), The Peanut Genome, Compendium of Plant Genomes. Springer International Publishing, Cham. 1-6 p.

Vassiliou E.K., Gonzalez A., Garcia C., Tadros J.H., Chakraborty G. & Toney J.H., 2009. Oleic acid and peanut oil high in oleic acid reverse the inhibitory effect of insulin production of the inflammatory cytokine TNF-α both in vitro and in vivo systems. *Lipids in Health and Disease,* 8(1): 25.

Vishwakarma M.K., Pandey M.K., Shasidhar Y., Manohar S.S., Nagesh P., Janila P. & Varshney R.K., 2016. Identification of two major quantitative trait locus for fresh seed dormancy using the diversity arrays technology and diversity arrays technology-seq based genetic map in Spanish-type peanuts. *Plant Breeding*, 135(3): 367-375.

Vishwakarma M.K., Nayak S.N., Guo B., Wan L., Liao B., Varshney R.K. & Pandey M.K., 2017. Classical and molecular approaches for mapping of genes and quantitative trait loci in peanut, in: The Peanut Genome. *Springer*, 93-116 p.

Vishwakarma M., Yadav P., Rai V., Kumar U. & Joshi A., 2022. Molecular markers and genomics assisted breeding for improving crop plants. *Elsevier*, 303-334 p.

Wang L., Yang X., Cui S., Wang J., Hou M., Mu G., Li Z. & Liu L., 2020. Identification of main effect and epistatic QTLs controlling initial flowering date in cultivated peanut (*Arachis hypogaea* L.). *Journal of Integrative Agriculture*, 19(10): 2383-2393.

Wang L., Wang L., Tan M., Yu H., Wang J., Li Y., Wang W., Yan X. & Wang L., 2022. Rapid identification and preliminary evaluation of quality characters of oilseed sunflower by near infrared spectroscopy. *Oil Crop Science*, 7(3): 142-148.

Wang M.L., Khera P., Pandey M.K., Wang H., Qiao L. *et al.*, 2015. Genetic Mapping of QTLs Controlling Fatty Acids Provided Insights into the Genetic Control of Fatty Acid Synthesis Pathway in Peanut (*Arachis hypogaea* L.). *PLOS ONE*, 10(4): 21.

Wang Y. & Song S., 2023. Variety identification of sweet maize seeds based on hyperspectral imaging combined with deep learning. *Infrared Physics & Technology*, 130: 104611.

Watson A., Hickey L.T., Christopher J., Rutkoski J., Poland J. & Hayes B.J., 2019. Multivariate Genomic Selection and Potential of Rapid Indirect Selection with Speed Breeding in Spring Wheat. *Crop Science*, 59(5): 1945-1959.

Wilson J.N., Chopra R., Baring M.R., Selvaraj M.G., Simpson C.E., Chagoya J. & Burow M.D., 2017. Advanced backcross quantitative trait loci (QTL) analysis of oil concentration and oil quality traits in peanut (*Arachis hypogaea* L.). *Tropical plant biology*, 10(1): 1-17.

World Resource Institute (WRI) (2020). Country Rankings. https://www.wri.org/applications/aqueduct/country-rankings/

Worthington R.E. & Hammons R.O., 1977. Variability in fatty acid composition among *Arachis* genotypes: A potential source of product improvement. *Journal of the American Oil Chemists' Society*, 54(2):4.

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

Xiaojing Z., Xia Y., Ren X., Chen Y., Huang L. *et al.*, 2014. Construction of a SNP-based genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). *BMC genomics*, 15: 351.

Xu Y., Li P., Yang Z. & Xu C., 2017. Genetic mapping of quantitative trait loci in crops. *The Crop Journal*, 5(2): 175-184.

Xu P., Sun W., Xu K., Zhang Y., Tan Q., Qing Y. & Yang R., 2023. Identification of Defective Maize Seeds Using Hyperspectral Imaging Combined with Deep Learning. *Foods*, 12(1): 144.

Young C.T., Worthington R.E., Hammons R.O., Matlock R.S., Waller G.R., Morrison R.D., 1974. Fatty Acid Composition of Spanish Peanut Oils as Influenced by Planting

Location, Soil Moisture Conditions, Variety, and Season. *J. Am. Oil Chem. Soc, 51*(7): 312-315.

Yu B., Huai D., Huang L., Kang Y., Ren X. *et al.*, 2019. Identification of genomic regions and diagnostic markers for resistance to aflatoxin contamination in peanut (*Arachis hypogaea* L.). *BMC Genetics*, 20(1): 32.

Yu H., Liu H., Erasmus S.W., Zhao S., Wang Q., van Ruth S.M., 2020. Rapid high-throughput determination of major components and amino acids in a single peanut kernel based on portable near-infrared spectroscopy combined with chemometrics. *Ind. Crops Pro*d. 158:112956.

Zahran H.A. & Tawfeuk H.Z., 2019. Physicochemical properties of new peanut (*Arachis hypogaea* L.) varieties. *OCL*, 26(19): 7.

Zhao S.-Z., Li A., Li C., Xia H., Zhao C., Zhang Y., Hou L. & Wang X., 2016. Development and application of KASP marker for high throughput detection of AhFAD2 mutation in peanut. *Electronic Journal of Biotechnology*, 25: 9-12.

Zhou X., Xia Y., Ren X., Chen Y., Huang L. *et al.*, 2014. Construction of a SNP-based genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). *BMC Genomics*, 15(1): 351.

Zhuang W., Chen H., Yang M., Wang J., Pandey M.K. *et al.*, 2019. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nature Genetics*, 51(5): 865-876.

Zhang S., Hu X., Wang F., Chu Y., Yang W. *et al.*, 2023. A stable and major QTL region on chromosome 2 conditions pod shape in cultivated peanut (*Arachis hyopgaea* L.). *Journal of Integrative Agriculture*, 22(8) :2323-2334.

APPENDICES

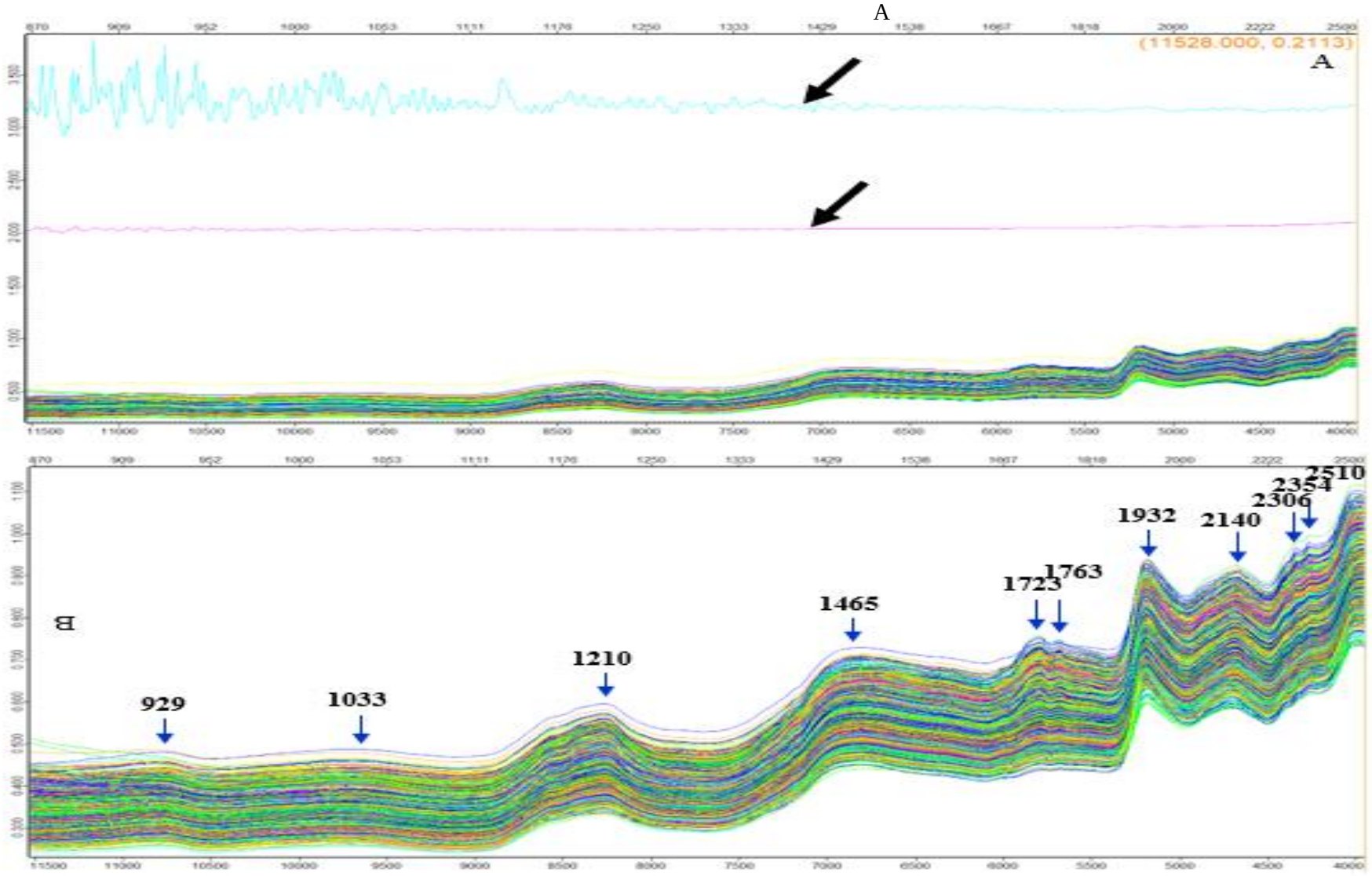
Appendix I. The logical framework of the thesis

General objective	Specific objectives	Materials and Methods	Results	Conclusions
To explore both, rapid germplasm screening for quality traits, and the favourable wild QTL linked to yield-traits	1. To screen peanut core-collection and interspecific population for quality traits;	The evaluation was conducted at Mbalmayo and Bafia in Cameroon and Nioro in Senegal. 133 AB genotypes long Fleur11 and 300 African core collections, were evaluated NIRS spectra were acquired from 680 samples from four sets. The spectra were acquired on six subsets of each sample after three rotation scans. The spectral resolution was 16 cm^-1 over the spectral range of 3952-11528 cm^-1 (867–2530 nm). The acquired spectra were processed using principal component analysis (PCA) coupled with partial least squares discriminant analysis (PLS-DA)	Eleven absorbance peaks were observed at specific wavelengths: 929 nm, 1033 nm, 1210 nm, 1465 nm, 1723 nm, 1763 nm, 1932 nm, 2306 nm, 2350 nm, 2140 nm, and 2510 nm. – A huge variation of absorbance was observed among the varieties and AB lines within and between environments. The strata were ordered core-collection, Bafia, Mbalmayo and Nioro. PCA was performed and the first 3 PCs represent 93 % of the total variability. A PLS-DA model classified samples-based environment and genetic origin with 99.6% prediction accuracy.	NIRS screened the germplasm rapidly because the only 6 days were needed to obtain the all spectra of 680(680 x 6) samples. Based on the NRIS spectra and PCA result, a wide variability of the seed composition was found within and between environments indicating seed composition affected by environment and genetic materials. A PLS-DA model was developed with a strong classification accuracy.

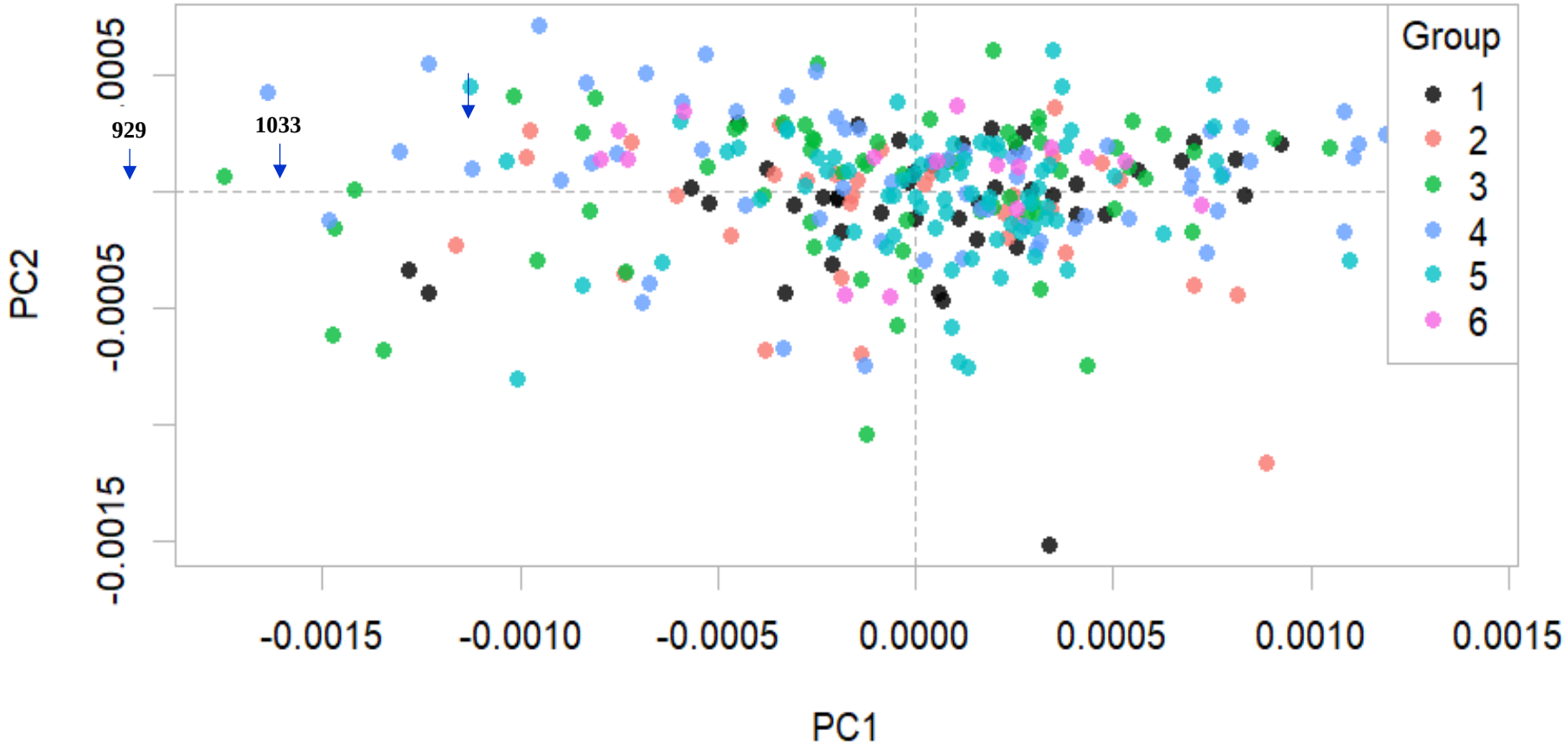
General objective	Specific objectives	Materials and Methods	Results	Conclusions
To explore both, rapid germplasm screening for quality traits, and the favourable wild QTL linked to yield-traits	2. To assess phenotypic variability, heritability, and trait correlations in an interspecific population for yield-related traits	Evaluated at Maroua, Bafia and Mbalmayo. 133 AB genotypes and Fleur11. were used. A total of 15 yield related traits, were collected. An analysis of variance (ANOVA) was performed to estimate the genetic and replication effects on each trait under each environment. Broad sense heritability (H2) was computed from ANOVA. Relationships between traits were estimated by the Pearson correlation coefficients.	Moderate to high phenotypic variability was observed for all quantitative traits. The plant growth habit (GH) showed morphological variation from semi-erect to the ground to totally erect. A similar variation was observed for pod construction (PC) and pod beak (PB) from slight to prominent and from slight to deep, respectively. Two AB lines were recorded based on the mean values of for HPW, HSW, PL, PW, SL and SW. Analysis of variance (ANOVA) showed significant differences at ($P < 0.001$) for 13 traits and at ($P < 0.05$) for 2 traits. Moderate to high) heritability was estimated for all traits. The highest positive and significant association were estimated between (TB and HaW), (TB and PWT, PW and HPW), (HPW and HSW) and (PC and PB).	The ANOVA result showed significant among the AB lines, highlighting the segregation of these traits. -The computed broad sense heritability showed moderate to high confirmed the impact environment condition for quantitative traits. The main yield related traits such as 100 pod and seed weight, and pod and seed length were corelated each other.

General objective	Specific objectives	Materials and Methods	Results	Conclusions
To explore both, rapid germplasm screening for quality traits, and the favourable wild QTL linked to yield-traits	3. To identify wild genomic regions (QTLs) associated with yield-related traits	-Evaluated at Maroua, Bafia and Mbalmayo. -133 AB lines. were used. -A total of 15 yield related traits, were collected. -Genomic DNA was extracted from 20-day-old leaves using CTAB extraction method. -Silico DArT markers and SNP markers were scored. -A total of 1,450 SNP markers were used for construction of a genetic map. -It is constructed using Joinmap software. -QTL was analysed for 15 yield traits in QTL IciMapping software v4.1.0.0. -The final high resolution linkage map for major effect QTL was generated using LinkageMap view pakage in R software.	-The genetic map was constructed contained 1,450 SNPs markers with a total length of 1,358.02 cM and 2.21 cM distance between adjacent markers on 20 linkage groups. A total of 44 main effect -QTL were detected for 14 traits except pod beak on 17 LGs. Four QTL new were mapped related to 100 pod and seed weight and seed width. Of 44 QTL, 13 were mapped as major effect for eight chromosomes for 8 traits such as: 100 pod and seed weight, halum weight, harvest index, seed length, pod weight, plant height and, pod constriction.	A total of 20 QTL were identified to increase phenotypic value from wild parent. This indicates the positive contribution of the wild relatives in peanut improvement

Appendix II. Raw spectra with two (arrows) (A) and without (B) non-atypical intact-seed spectra.



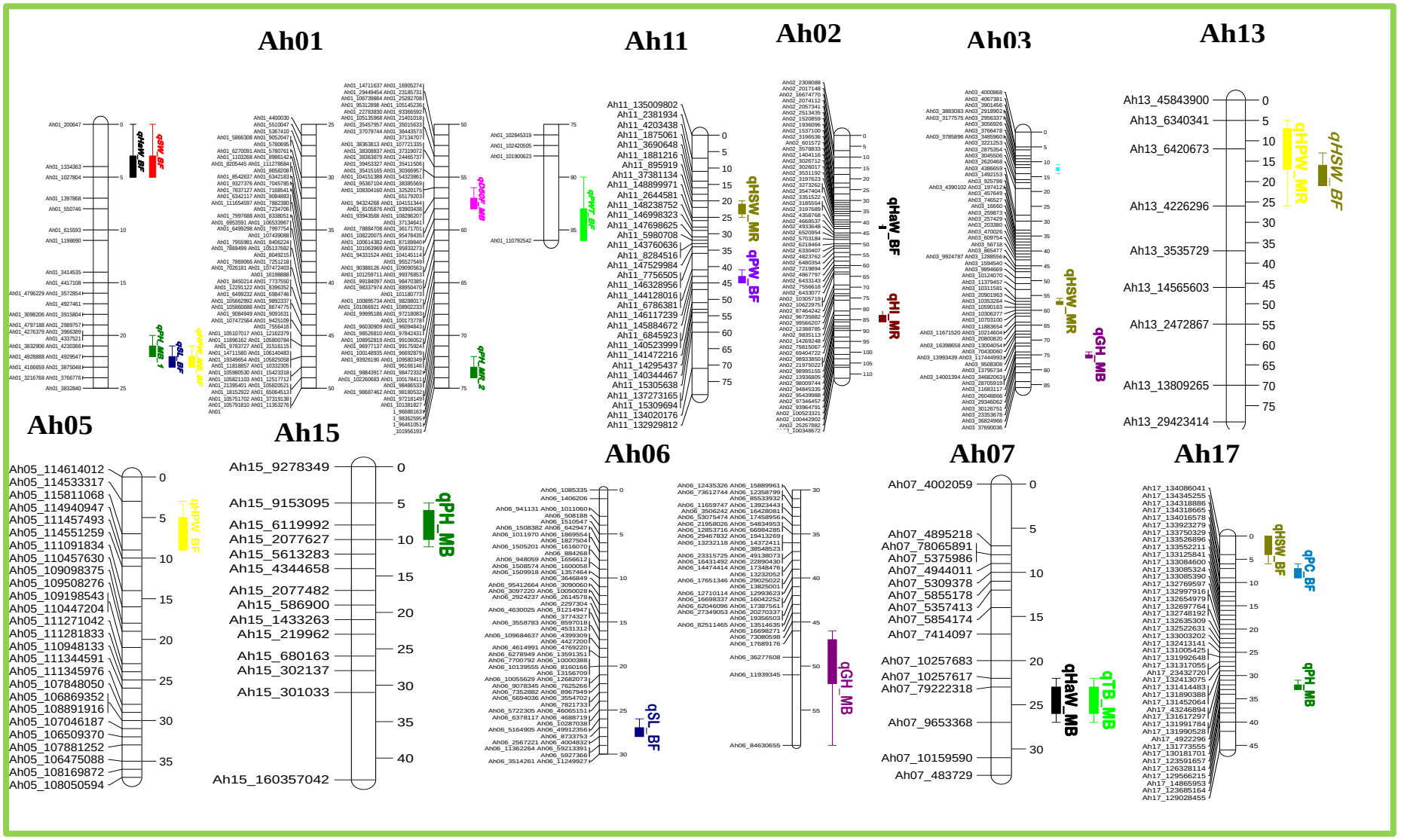
Appendix III. PCA 2-dimensional score plot of PC2 versus PC1 for the date of spectral acquisition.

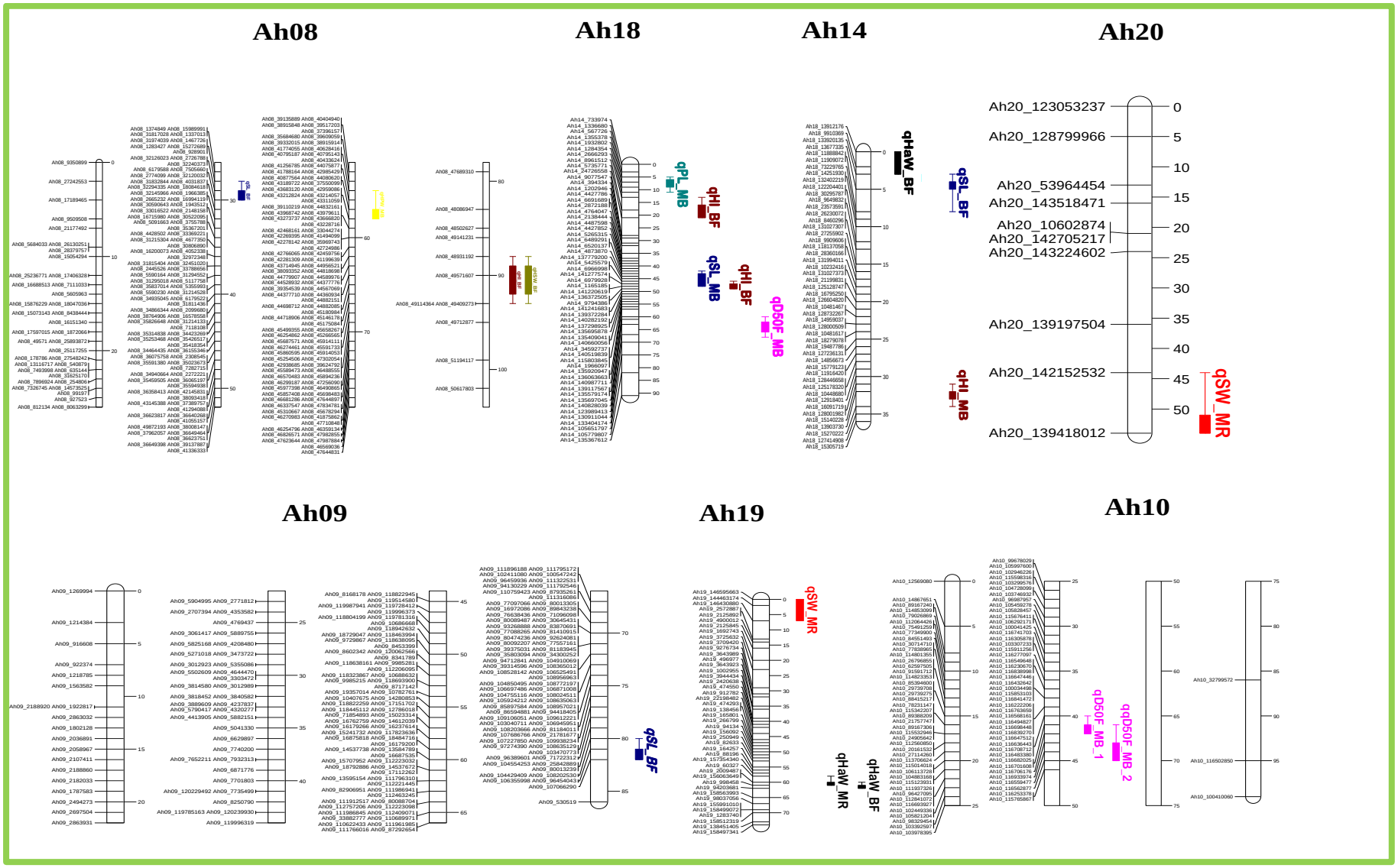


Appendix IV. The variability of yield related traits mean value in each environment.

Traits	Environments					
	Maroua		**Mbalmayo**		**Bafia**	
	Range	Mean	Range	Mean	Range	Mean
Plant height (PH) (cm)	6.00–19.33	12.32	11.08–20.33	16.41		
Total biomass (TB) (g)	20–111.67	67.24	20.33–59.00	39.21	13.5–89	50.25
Pod weight (PWT) (g)	6.67–27.33	16.34	3.33–16	9.22	3.00–16	7.92
Halum weight (HaW) (g)	11.33–95.67	51.2	15–45	30.12	12.00–76	43.9
Harvest index (HI) (ratio)	12.00–46	27.12	12.00–37.00	24.23	5.00–28	15.27
Hundred pod weight (HPW) (g)	59.67–91.33	76.94	49.00–90	69.59	46–95	72.59
Hundred seed weight (HSW) (g)	29.67–41.33	34.86	26.00–44.33	35.52	27.00–50.5	39.45
Pod length (PL) (mm)	21.06–29.33	25.	21.82–28.73	25.29	21.51–28.29	24.67
Pod width (PW) (mm)	10.36–12.73	11.49	9.69–12.25	10.98	8.75–12.8	11.09
Seed length (SL) (mm)	10.96–13.93	12.34	10.53–12.79	11.88	10.29–13.85	12.02
Seed width (SW) (mm)	6.13–7.81	6.95	6.69–8.46	7.71	5.88–8.77	7.12

Appendix V. Genetic linkage map showing the location of main effect QTLs identified using inclusive composite interval mapping (ICIM) for 14 yield related traits among the BC2F4 lines.





D50%F: days to 50% flowering, **GH**: growth habit, **PH**: plant height, **HPW**: hundred pod weight, **HSW**: hundred seed weight, **TB**: total biomass, **HaW**: halum weight, **PWT**: pod weight, **PL**: pod length, **PW**: pod width, **SL**: seed length, **HI**: harvest index, **PC**: pod constriction, **SW**: seed width, **MR:** Maroua, **MB**: Mbalmayo, **BF**: Bafiya and **q**:QTL.

Oil Crop Science 9 (2024) 132–141

Contents lists available at ScienceDirect

Oil Crop Science

journal homepage: www.keaipublishing.com/en/journals/oil-crop-science

Application of near-infrared spectroscopy for fast germplasm analysis and classification in multi-environment using intact-seed peanut (*Arachis hypogaea* L.)

Fentanesh Chekole Kassie [a,b], Gilles Chaix [c,d,e], Hermine Bille Ngalle [a], Maguette Seye [f,g], Coura Fall [f,g], Hodo-Abalo Tossim [f,g], Aissatou Sambou [f,g], Olivier Gibert [c,d], Fabrice Davrieux [h], Joseph Martin Bell [a], Jean-François Rami [c,d,g], Daniel Fonceka [c,d,f,g], Joël Romaric Nguepjop [c,d,f,g,*]

[a] Department of Plant Biology and Physiology, Faculty of Sciences, University of Yaounde I, Yaounde, P.O. Box 337, Cameroon
[b] Department of Plant Science, College of Agriculture, Wolaita Sodo University, Sodo, P.O. Box 138, Ethiopia
[c] CIRAD, UMR AGAP Institut, Montpellier, France
[d] AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France
[e] ChemHouse Research Group, Montpellier, France
[f] Centre D'Etudes Régional Pour L'Amélioration de L'Adaptation à La Sécheresse (CERAAS/ISRA), Route de Khombole, Thiès, BP 3320, Senegal
[g] Dispositif de Recherche et de Formation en Partenariat, Innovation et Amélioration Variétale en Afrique de L'Ouest (IAVAO), CERAAS, Route de Khombole, Thiès, BP 3320, Senegal
[h] CIRAD, UMR QUALISUD, Saint Pierre, France

ARTICLE INFO

Keywords:
Groundnut
Oilseed
Near infrared spectroscopy
Germplasm analysis
Environment
Nutritional
Breeding

ABSTRACT

Peanut is a worldwide oilseed crop and the need to assess germplasm in a non-destructive manner is important for seed nutritional breeding. In this study, Near Infrared Spectroscopy (NIRS) was applied to rapidly assess germplasm variability from whole seed of 699 samples, field-collected and assembled in four genetic and environment-based sets: one set of 300 varieties of a core-collection and three sets of 133 genotypes of an interspecific population, evaluated in three environments in a large spatial scale of two countries, Mbalmayo and Bafia in Cameroon and Nioro in Senegal, under rainfed conditions. NIR elemental spectra were gathered on six subsets of seeds of each sample, after three rotation scans, with a spectral resolution of 16 cm^{-1} over the spectral range of 867 nm to 2530 nm. Spectra were then processed by principal component analysis (PCA) coupled with Partial least squares-discriminant analysis (PLS-DA). As results, a huge variability was found between varieties and genotypes for all NIR wavelength within and between environments. The magnitude of genetic variation was particularly observed at 11 relevant wavelengths such as 1723 nm, usually related to oil content and fatty acid composition. PCA yielded the most chemical attributes in three significant PCs (i.e., eigenvalues >10), which together captured 93% of the total variation, revealing genetic and environment structure of varieties and genotypes into four clusters, corresponding to the four samples sets. The pattern of genetic variability of the interspecific population covers, remarkably half of spectrum of the core-collection, turning out to be the largest. Interestingly, a PLS-DA model was developed and a strong accuracy of 99.6% was achieved for the four sets, aiming to classify each seed sample according to environment origin. The confusion matrix achieved for the two sets of Bafia and Nioro showed 100% of instances classified correctly with 100% at both sensitivity and specificity, confirming that their seed quality was different from each other and all other samples. Overall, NIRS chemometrics is useful to assess and distinguish seeds from different environments and highlights the value of the interspecific population and core-collection, as a source of nutritional diversity, to support the breeding efforts.

* Corresponding author. CIRAD, UMR AGAP Institut, Montpellier, France.
E-mail address: joel.romaric.nguepjop@cirad.fr (J.R. Nguepjop).

https://doi.org/10.1016/j.ocsci.2024.03.003
Received 18 January 2024; Received in revised form 7 March 2024; Accepted 12 March 2024
Available online 15 May 2024
2096-2428/© 2024 the Oil Crop Research Institute, Chinese Academy of Agriculture Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Peanut is an annual oilseed crop cultivated globally on 36.18 million hectares of area in the world, yielding 71.68 million metric tons of pods in 2020 (FAOSTAT, 2020). As a functional food, peanut seed contains 34% – 56% oil, 22% – 30% protein, 10% – 25% carbohydrates, and 0.05% – 1% of various secondary metabolites, beneficial to human health, such as vitamin E, K and B complex, folic acid, niacin, and minerals (Ca, P, Mg, Zn, and Fe) (Desmae et al., 2019; Harch et al., 1995; Janila et al., 2013; Parilli-Moser et al., 2022). The main production constraints of the crop include drought, pests, diseases, and environmental changes. The oil content of seeds, shelf life, aroma, flavor and cooking quality are all affected by these constraints. Consequently, seed quality traits are targets in genetic breeding (Nawade et al., 2018; Parmar et al., 2022; Tang et al., 2022).

Peanut (*Arachis hypogaea* L.) is an autogamous species, allotetraploid (AABB genome; 2n = 4x = 40) with narrow genetic base (Burow et al., 2009; Simpson et al., 2001). The narrow genetic diversity coupled with low utilization of genetic resources are the major factors limiting peanut breeding. Thus, interspecific hybridization is currently used as a realistic strategy for introgressing prospective diversity from wild species into the cultivated gene pool (Favero et al., 2006; Fonceka et al., 2012a, 2012b; Mallikarjuna et al., 2011b; Tossim et al., 2020). Likewise, genetic diversity assessment and the detection of promising quality-related genotypes are fundamental to germplasm utilization and management in breeding strategies to support food security. To facilitate the investigation of large germplasm, it is reasonable to begin by examining subsets of germplasm that embody appropriate diversity and of manageable size, such as core collections or interspecific populations derived from wild × elite crosses, using appropriate characterization procedures.

Although significant efforts have been devoted to characterizing germplasm for simple traits and for the most important agronomic traits (yield and resistance to pests and diseases) (Fan et al., 2020; Kumari et al., 2014; Mallikarjuna et al., 2011a; Upadhyaya, 2005; Upadhyaya et al., 2011), less is known about quality traits across environments (Grosso et al., 2000; Wang et al., 2023). This is mainly due to the fact that these traits are quantitative and multigenic, with low heritability, and strong genotype environment interactions (Grosso et al., 1994; Isleib et al., 2008). Moreover, the phenotyping of these traits, regularly base on chemical survey, is expensive in terms of both direct monetary input and human labor, time-consuming, complex, and irreversibly destructive. Another main factor limiting chemical studies are the difficulties to assess many samples, each requiring many seeds (Davis et al., 2021; Nawade et al., 2018). Efforts to improve the knowledges of seed attributes might be supported by rapid and non-destructive tools. These include modified refractive index, capacitance sensor (Kandala et al., 2008), hyperspectral imaging (Huang et al., 2014; Rabanera et al., 2021) and near infrared (NIR) spectroscopy (Davis et al., 2021; Govindarajan et al., 2009; Tao et al., 2019; Wang et al., 2022). Among these, NIR-based methods are rapid, make it possible to analyze large number of samples. Moreover, some scholars have already applied machine learning as promising statistical methods to assist humans in the modeling and analysis of complex spectral data (Fordellone et al., 2019; Song et al., 2018) in many research fields including seed quality detection, genotyping of cultivars (Panero et al., 2022), varieties identification (Panero et al., 2018, 2022; Wang and Song, 2023; Xu et al., 2023) and classification (Sampaio et al., 2021; Singh et al., 2023; Tian et al., 2023). Some works previously described the feasibility of near infrared spectrometers to achieve some quick prediction of various peanut chemical compounds (proximal components and secondary metabolites) (Bilal et al., 2020; Li et al., 2019; Liu et al., 2022; Yu et al., 2020). In this paper, we focused on the non-destructive approach by NIR spectroscopy to investigate the environment and the genetic contribution of germplasm variability from intact-peanut-seed spectra without chemical references.

In this study, NIR spectroscopy was applied and coupled with chemometrics to assess germplasm variability from peanut intact-seed of a core-collection and of an interspecific population field-evaluated in three different environments. The objectives were specifically to i) perform a rapid NIR measurement on seeds and check the quality of spectra data, ii) assess genetic variation of varieties and genotypes from seed spectra, iii) study the pattern of genetic variability of the interspecific population in comparison to the core-collection, iv) potentially discriminate genetically related interspecific genotypes within and between environments by developing a classification model using PLS-DA.

2. Materials and methods

2.1. Genetic materials

Two distinct genetic materials were used in this study: an interspecific advanced backcross QTL (AB-QTL) population of 133 genotypes and a core collection of 300 cultivars. The AB-QTL population of 133 BC_2F_4 derivatives was developed from an interspecific cross, using Fleur11 as recurrent cultivated parent and the wild synthetic tetraploid 'ISATGR 278-18' (Nguepjop et al., 2016). The cultivated parent, Fleur 11, is an elite Spanish-type variety, widely cultivated in West Africa. The wild parent, ISATGR 278-18 is derived from a cross between *A. batizocoi* ICG 13160 (GKBSPSc 30,082, PI 468328) and *A. duranensis* ICG 8138 (GKP 10038, PI 262133) (Mallikarjuna et al., 2011b). The core collection of 300 cultivars was defined based on the knowledge of breeders and on diversity data from a collection of 1050 accessions (breeding lines and landraces) held by 10 breeding programs in East, Southern and West Africa (Conde et al., 2023). The detailed information of the 300 varieties of the core-collection and the 133 genotypes of the population are presented in the Supplementary Tables 1 and 2, respectively.

2.2. Trials environment and field experimental design

Whole seed used were collected from field experiments. Experiments were conducted in 3 different locations in 2 countries, Mbalmayo and Bafia in Cameroon and Nioro in Senegal, under rainfall conditions in 2021. The 3 locations were chosen to meet environmental differences, based on different criteria, largely based on ecology (climate and vegetation) and tradition of peanut cultivation and crop rotation (Table 1). Bafia is one of the main areas of peanut production in Cameroon. It is located in tropical savanna and has yellow vertisol soil (Temga et al., 2021) and equatorial climate of the Sudano-Guinean type with an average temperature of 25.1 °C and annual rainfall of 1500 mm. Mbalmayo is located in the tropical forest of Cameroon and has ocher vertisol soil (Temga et al., 2021) with a bimodal humid-forest rainfall climate with an average temperature of 26.5 °C and rainfall of 2402 mm. Nioro is located in the South of the Senegalese peanut basin and have Sahelian

Table 1
Characteristics of the field environments.

	Trial Environments		
	Bafia	Mbalmayo	Nioro
Country	Cameroon	Cameroon	Senegal
Location	Bafia	Mbalmayo	Nioro
Peanut cultivation	++++	+++	+++++
Ecology type	Tropical savanna	Tropical forest	Sahelian
Climate type	Sudano-guinean equatorial	Humid-forest bimodal rainfall	Sahelian semi-humid
Temperature (°C)	25.1	26.5	30.0
Rainfall (mm)	1500	2403	758
Soil type	Yellow vertisol	Ocher vertisol	Deck-Dior (leached ferruginous)
Previous crop	Maize	Maize	Millet
Experiment period	April–July	April–July	July–October

semi-humid ecology with a Deck Dior soil, a leached ferruginous tropical soil (Bogie et al., 2018), and annual rainfall of 758 mm and average temperature of 30 °C. The fields at Bafia and Mbalmayo were one-year fallow land after maize cultivation by farmers and were cleared and plowed for the study. The previous crop at Nioro was millet. The experiments in Bafia and Mbalmayo were conducted during one of the two rainy seasons from April–July, while the Nioro experiment was done during the rainfall season between July and October, at the Research Station of National Agricultural Research Center.

The same experimental design with common agricultural practices, from sowing to harvest, were used in each of the 3 environments. Within each environment, an alpha-lattice design was used with 3 replications, with 10 elementary plots within blocks. A plot consisted of rows of 3 m long on which 10 plants of the same genotype were sown with a spacing of 30 cm between plants on the same row, and 50 cm between two adjacent rows. The seeds were treated with Granox (captafol 10%, benomyl 10%, and carbofuran 20%) before planting to protect them against parasitic attacks and one seed per hill was sown manually at 4 cm depth. According to usual cultural practice, one hundred and fifty kg/ha of mineral fertilizer (6-20-10) were added 20 days after sowing. Throughout the vegetative development, weeds were manually harvested. The harvest was done at 95 days after sowing, followed by free-air drying for one month. At the end of the pod-drying stage, pods of each plant were separated from haulms, stored and dehulled.

2.3. Whole seed sample preparation

Whole seeds from pods of the three agronomic replicates of each genotype were bulked into one specific sample, stored in plastic bag, and labelled according to their respective name and environment. Thus, seeds of each sample for NIR analysis came from pods of 25–30 harvested plants of each genotype. From the expected 699 samples, we discarded 21 who had less than 100 seeds, 3 from Bafia, 9 from Mbalmayo and 9 from the core-collection. Finally, a total of 680 samples of seed were assembled in four genetic- and environment-based sets: one set of the 291 samples from cultivars from Nioro and three sets of the interspecific genotypes (130 samples for Bafia, 124 samples for Mbalmayo, 135 samples for Nioro, including the 133 genotypes and the CS16 variety and the cultivated parent Fleur11, both commonly used as check varieties in Nioro). All sample sealed in hermetic plastic bags were conveyed to the laboratory and kept at ambient temperature prior to spectra acquisition.

2.4. NIR spectra acquisition

Spectra acquisition was performed to generate a reference database. Prior to recording spectra, a gold reference was used. Spectra were then acquired on six subsets of each 680 samples. The six subsets of each sample were used, as six replicates, to minimize uncertainties due to the hypothetical heterogeneity of seed. Specifically, seeds of each sample, were six-fold randomly sampled to provide biological and analytical replicates, from each other to cover the whole sample. Seeds of each subset were then loaded in the ring cup with an internal diameter of 5 cm and the six subsets of each sample were measured in sequence. Spectra of each of the six subsets were gathered after 3 rotation scans with a spectral resolution of 16 cm^{-1} over the spectral range of 3952 cm^{-1} – 11528 cm^{-1} (867 nm – 2530 nm), using Tango spectrometer from Bruker. At the end, each sample was analyzed in six replicates, and the mean spectra were used for data analyses.

2.5. Statistics and PCA analysis

R software (R Core Team, 2021) with rchemo (Brandolini-Bunlon, et al., 2023) and rnirs packages (Lesnoff, 2021) were used to visualize raw spectra and perform data analysis. PCA over the spectral range selected from 1000 nm to 2500 nm was applied to describe variability according to the varieties and interspecific genotypes within and between environments. PCA is a multivariate unsupervised statistical method able to project multivariate data and describe relevant trends in the analyzed dataset. PCA can also reveal variables with loading that determine some inherent structure of the data, which can be interpreted in chemical terms. The reduction of the number of variables is achieved by making a linear combination of original variables, which yields the so-called principal components (PC) that are decorrelated with each other. PCA was conducted on the pretreated spectra. The full whole spectra have been pre-processed to improve the signal by reducing uncontrolled variations as noise and baseline through Savitsky Golay (SavGol) and derivative.

Mahalanobis distance was computed after PCA to check the 6 replicates distances for each sample. The Mahalanobis distances were determined in units of standard deviations from the center (mean) of the dataset. The 6 replicates were averaged or each sample, and the Mahalanobis distances were computed again.

In this study, the following PCA results were considered (i) the score plot, to visualize the projection of the sample on each PC; and (ii) the loading plot, to evaluate the influence of wavelength, on each PC. Thus, PCA allows emphasizing and interpreting variables and all relevant differences among genotypes within and between environments.

2.6. Classification using PLS-DA modeling on NIR spectra

PLS-DA was used to classifying varieties and interspecific genotypes thorough modeling and prediction of genotype-specific spectra, according to genetic and environmental origin. Data have been split by Duplex method (Snee, 1977) into train set (N = 541, 201, 108, 109, 126 respectively for Core population, AB-QTL Bafia, AB-QTL Mbalmayo, AB-QTL Nioro) and test set (N = 139, 42, 32, 31, 34 respectively for the previous populations) in each group (to keep the same proportionality). The train set was used to train the model, while the test set is used to evaluate its performance. Prior to applying PLS-DA algorithms, the train set spectra were pre-processed by SavGol filter and derivative. The best preprocessing was selected according to the error of classification by cross validation (2 K-fold group repeated 20 times) and the number of latent values was fixed. Then, these parameters were used to build the PLS-DA model and applied on test set spectra.

The resulting confusion matrix of each model were further evaluated to assess the model's performance using the following metrics for each group and for all.

- Recall (the proportion of samples of a specific class that have been predicted by the model as belonging to that class; also known as sensitivity)

$$RECALL = \frac{TP}{FN + TP}$$

- Specificity (The number of samples predicted correctly to be in the negative class out of all the samples in the dataset that actually belong to the negative class.)

$$SPECIFICITY = \frac{TN}{FP + TN}$$

- Precision (the proportion of correct predictions among all predictions for a particular class)

$$PRECISION = \frac{TP}{FP}$$

- Accuracy (the number of samples correctly classified out of all the samples present in the test set)

$$ACCURACY = \frac{TP + TN}{TP + FN + FP + TN}$$

- the proportion of false-negatives (FNR)

$$FNR = \frac{FN}{TP + FN}$$

- the proportion of false-positives (TNR)

$$TNR = \frac{TN}{FP + TN}$$

- the F1-score (the harmonic mean of precision and recall)

$$F_1 score = \frac{2TP}{2TP + FP + FN}$$

True Positive (TP) refers to a sample belonging to the positive class being classified correctly. True Negative (TN) refers to a sample belonging to the negative class being classified correctly. False Positive (FP) refers to a sample belonging to the negative class but being classified wrongly as belonging to the positive class. False Negative (FN) refers to a sample belonging to the positive class but being classified wrongly as belonging to the negative class. Model performances were evaluated by their classification accuracy, which was calculated as the ratio of the number of correctly classified samples to the total number of samples.

3. Results

3.1. Spectra profiles and quality control

From the raw spectra, eleven relevant absorbance peaks were observed around the wavelengths of 929 nm, 1033 nm, 1465 nm, 1763 nm, 2306 nm, 2350 nm and 2510 nm, with four wide spectral peaks appearing close to 1210 nm, 1723 nm, 1932 nm and 2140 nm (Fig. S1). Quality control of spectra was performed to identify atypical spectra and to check variation among the six subsets of each samples. As results, 2 of 4080 spectra (0.04%), were identified as an outlier (Fig. S1) and were discarded for analyses. PCA was performed to check the effect of date on spectra acquisition and no cluster related to date was found (Fig. S2), indicating that there were stable lab conditions during the 6 days of spectra acquisition. With few exceptions, the Mahalanobis distance (MD) among the 6 subsets of each sample was consistent among the 680 samples (Fig. 1). Thus, the spectra graph was presented in Fig. 2 as the average absorption of each sample from the 6 replicated spectra.

3.2. Genetic variability and environmental impact on intact-seed composition

The mean absorbance spectra of varieties and interspecific genotypes, according to their environment are presented in Fig. 2. A huge variation of absorbance along the spectra was observed among varieties and interspecific genotypes within and between environments. Four spectra group, superimposed on each other, was observed for all wavelength

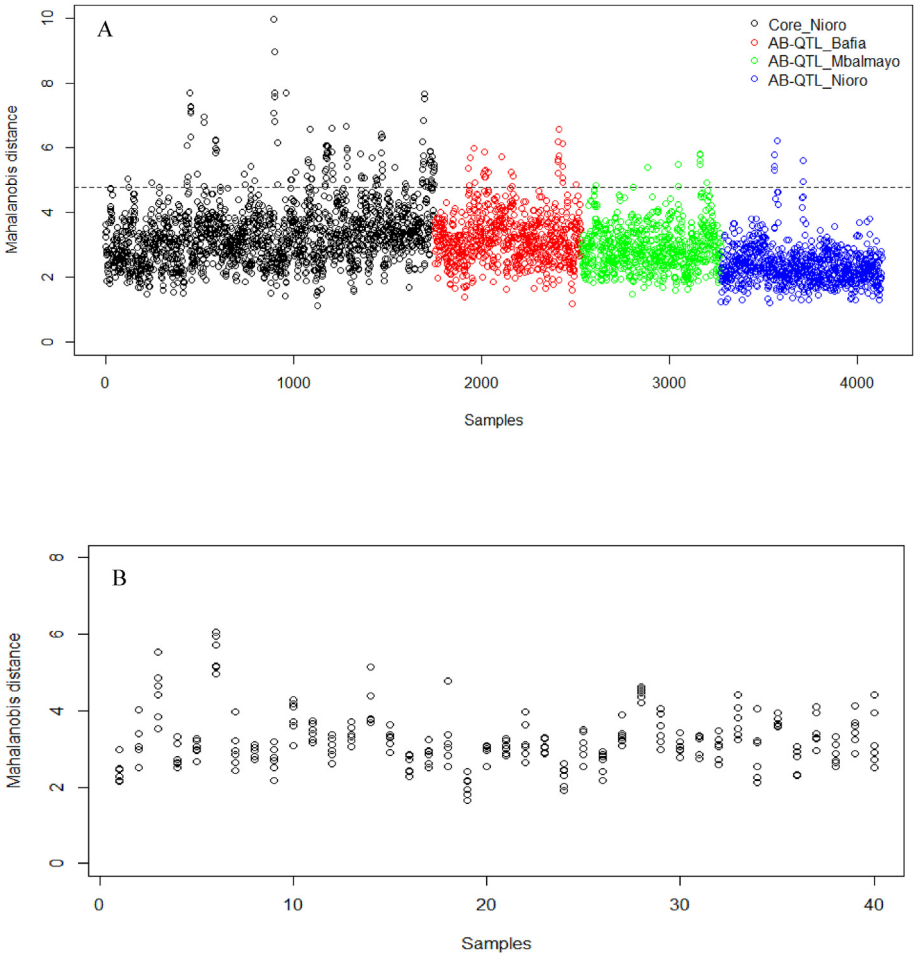


Fig. 1. Plot showing Mahalanobis distance among the six subsets of each sample of four populations. Each dot represents one spectrum. MD Details of 40 samples (B) is figured from the 4080 spectra (A) for a better MD visualization among the 6 spectra of each sample (dotted line: distance cutoff - Chi-squared distribution for Standard Deviation squared (Brandolini-Bunlon, et al., 2023)).

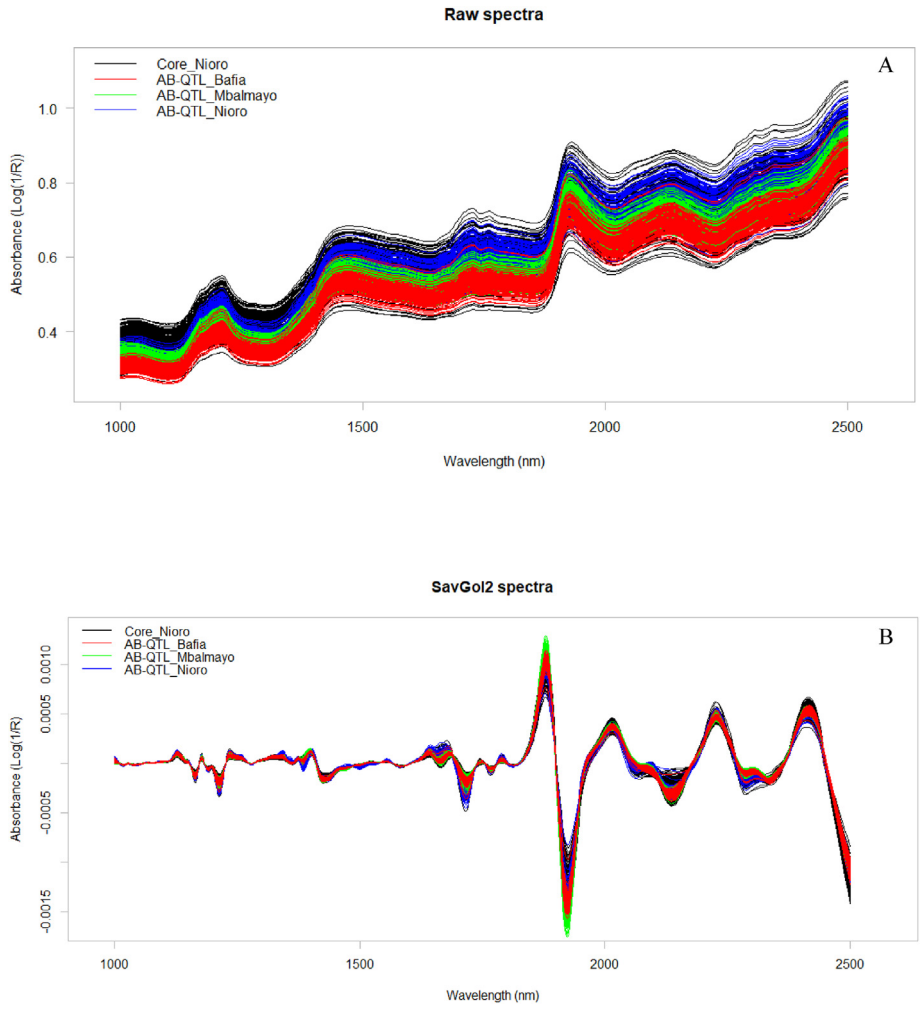


Fig. 2. NIR spectra of intact-seed according to genetic and environment origin of samples without treatment (A) and after Stavisky Golay filter with derivative 2 pre-processing (B). Spectra of varieties of the core collection are labelled in black. Spectra of interspecific AB-QTL genotypes from Bafia, Mbalmayo and Nioro, environments are labelled in red, green and blue, respectively.

from 1000 nm to 2500 nm (Fig. 2 A). Each spectra group corresponds to each of the four studied sets. The widest spectra group corresponded to the set of the core collection while the three other ones were each specific to the three sets of the interspecific population, each from one of the 3 studied environments, Bafia, Mbalmayo and Nioro (Fig. 2 A). The absorbance range of interspecific population was highest in Bafia followed by Mbalmayo and Nioro, pointing out the effect of environmental factors on chemical composition of seeds.

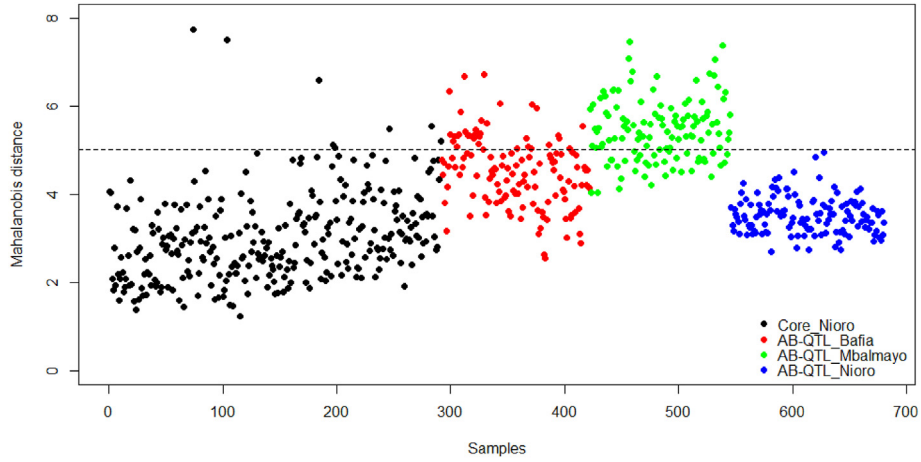


Fig. 3. Plot showing Mahalanobis distance among varieties and interspecific genotypes Each dot on the plot represents a variety or genotype. Varieties of the core collection are labelled in black. Interspecific AB-QTL genotypes from Bafia, Mbalmayo and Nioro, environments are labelled in red, green and blue, respectively **(**dotted line: distance cutoff - Chi-squared distribution for Standard Deviation squared (Brandolini-Bunlon, et al., 2023)**).**

3.3. Pretreatments effects on spectra

Different spectra pretreatments, SNV, Detrend and SavGol were applied in the raw spectra since spectral measurements can be affected by many factors leading to interference (light diffusion, scattering, …) with consequence observed as additive and multiplicative effects on raw spectra data. As example, the absorbance spectra pre-treated by SavGol filter with a window width of 15 points and first derivative was shown in Fig. 2B. As expected, the pretreatments eliminated physical effects due to seed dimension, surface of seed, etc., with consequences on light diffusion. Thus, from pretreated spectra, a huge MD variation, from 1 to 8, was found among varieties and genotypes (Fig. 3). Likewise, a distinct MD was found between 3 environments with a highest value at Mbalmayo followed by Bafia and Nioro.

3.4. Principal component analysis

PCA was performed using pretreated spectra after Savitzky-Golay filter with a window width of 15 points and first derivative. The first 5 PC represent more than 95% of the total variability with the values 60.5, 17.0, 15.5, 3.6 and 1.6, respectively. The PC1/PC2 and PC3/PC4 score plots are shows in Fig. 5. As expected, these figures show greater variability in the core collection and less variability in the other groups. The PC3/PC4 plot allows to distinguish easily the 4 seed lots. These plots showed that samples from different genetic and environmental origins are able to be well clustered and that they have great potential to be correctly identified.

Loading plots showing how each variable correlates with PC are shown in Fig. 4. The first loading indicates that the regions around 1900 nm and 2150 nm have a higher influence on PC1. Likewise, regions around 1210 nm, 1720 nm and 2300 nm seemed more related to PC2. For PC3, the region around 2400 nm seemed to be more important. PC4 was more related to 1400 nm, 1800 nm, 1950 nm and 2150 nm regions. The varieties and interspecific genotypes demonstrating contrasted scores in the top PCs were recorded (Fig. S3) and could be used further in peanut breeding programs.

3.5. Discrimination of genetically related interspecific genotypes among environments

The score plots illustrated that data could be grouped into four clusters, with overlapping the main clusters at the margin, with some interspecific genotypes and varieties superimposed, particularly, at the Nioro environment-set cluster (Fig. 5). The two most separated environments in the plane determined by plot scores were Mbalmayo and Bafia. With few exceptions, all interspecific genotypes from Mbalmayo exhibited high positive values at the PC3 compared to the other environments. This suggests that Mbalmayo environment might positively increases the seed traits associated with PC3. Finally, the African varieties studied in one environment added genetic variability to the environmental variability, resulting in a wide range of differences.

3.6. Classification based on whole seed spectra

A PLS-DA model was developed and the classification results of the model were shown in Table 2. The classification accuracy on the test set was 99.6% with correctly classified instances of the 4 samples sets i.e. African varieties in one environment and the interspecific genotypes from the 3 environments (Table 2). Interestingly, the confusion matrix achieved for the two sets, Bafia and Nioro shows 100% of instances classified correctly with 100% at both sensitivity and specificity. These two sets did not show incorrect instances, even in the model generated when all other sets were considered, thus confirming that their seed composition seemed very different from each other and from those of the other seed samples. These results showed that NIRS combined with discrimination analysis based on PLS regression is a simple and efficient tool for the classification of peanut genotypes, depending on each combination of the genetic and environment origins, which determine plant nutritional availability.

4. Discussion

The efficiency of NIR spectroscopy, as tool for fast and non-destructive large germplasm characterization in multi-environment were later discussed under the umbrella of breeding in intra and interspecific context.

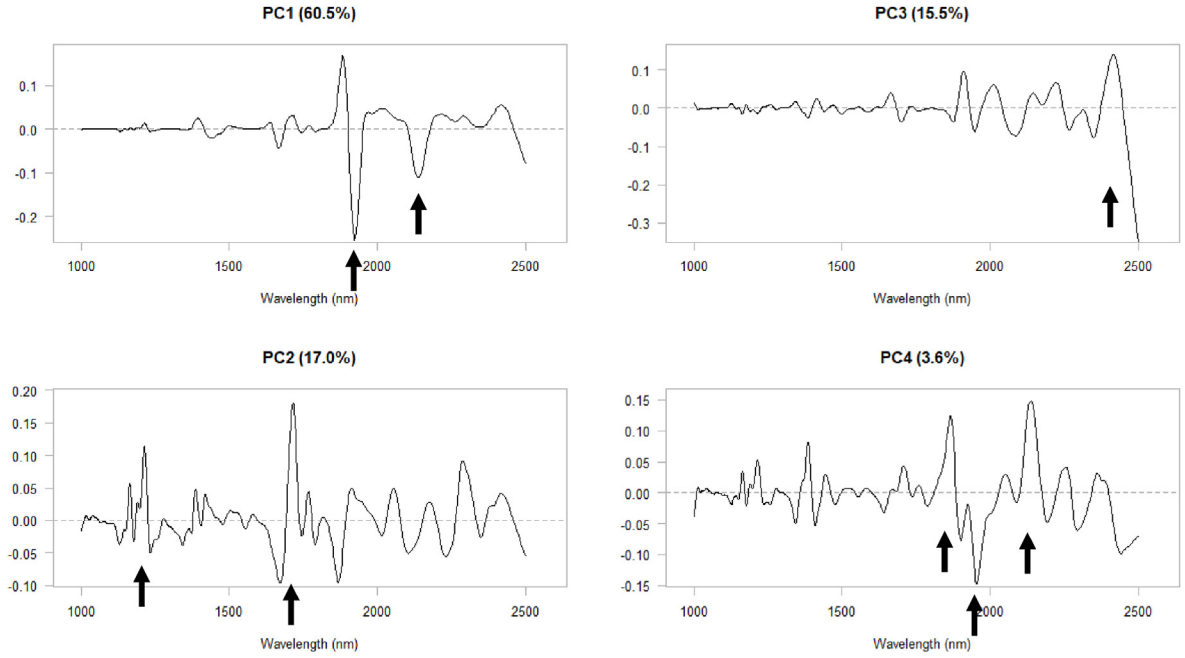


Fig. 4. PCA loading plots for the fourth first PCs showing how each variable correlate to each PC for wavelength.

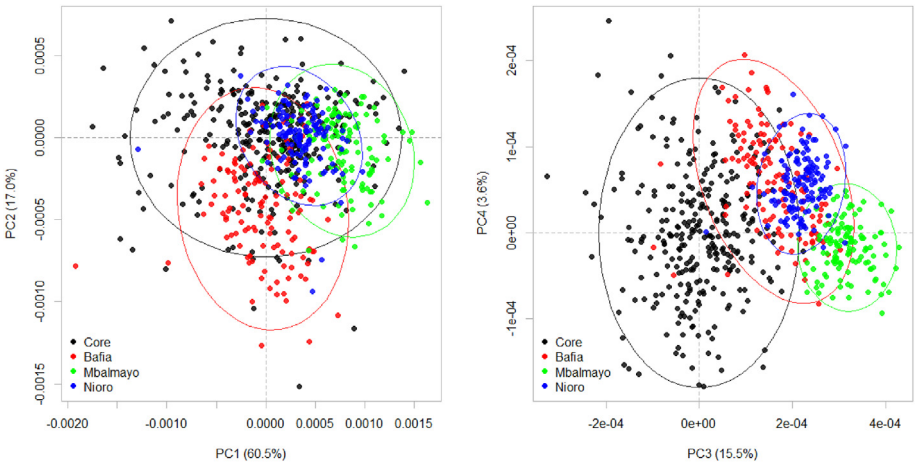


Fig. 5. PCA visualization of core varieties and interspecific genotypes among environments. PCA 2-dimensional score plots of PC2 and PC1 (A) and PC3 and PC4 (B) using NIRS spectra. Each dot on the plot represents a variety or genotype. Varieties of the core collection are labelled in black. Interspecific AB-QTL genotypes from Bafia, Mbalmayo and Nioro, environments are labelled in red, green and blue, respectively.

Table 2
Confusion matrix showing classification performance of PLS-DA model applied to test set sample samples (N = 139, Class 1: Core, Class 2: AB-QTL Bafia, Class 3: AB-QTL Mbalmayo, Class 4: AB-QTL Nioro).

		Predicted								
		1	2	3	4	Actual	Accuracy	Precison	Recall	F1-score
Actual	1	41	0	0	0	42	0.993	0.976	1.000	0.988
	2	0	32	0	0	32	1.000	1.000	1.000	1.000
	3	1	0	31	0	31	0.993	1.000	0.969	0.984
	4	0	0	0	34	34	1.000	1.000	1.000	1.000
	Pred	41	32	31	34	139				
Accuracy										0.996
Specificity										0.998
Recall										0.993
Precision										0.993
Proportion of false-negatives										0.007
Proportion of false-positives										0.002

Peanut is an important oil seed crop and the need to characterize peanut germplasm is essential as the demand for peanut is increasing continuously in various end product applications. According to the rapid and non-destructive attributes of the NIR, a total of 6 days was required to obtain all spectra of the six subsets of seed of 680 samples. The low level (0.04%) of outlier spectra on the global data set was considered as a good basis for analysis. Typical spectra observed in this study were in accordance with those reported in past studies. From raw spectra, eleven major peaks were observed. The region around 1210 nm, 1720 nm, 1763 nm, 2306 nm and 2350 nm could be assigned to fatty acids or oil content, which are generally considered as major components of peanut kernels (Govindarajan et al., 2009; Sundaram et al., 2009; Tao et al., 2019). The spectral peak around 2140 nm would likely result from the absorbance of proteins. The absorbance peak around 1465 nm might be related to the O–H overtone bond. The sharp peak around 1932 nm was due to the strong absorption of water contained in peanut kernels (Govindarajan et al., 2009; Sundaram et al., 2009; Tao et al., 2019). In the future, predictive models will be developed for nutritional content of peanut seeds.

A wide genetic variation was found among varieties and interspecific genotypes within environments. An environmental effect on seed compounds was highlighted by using the same interspecific population, thorough 3 environments. The largest variation was found in Bafia, followed by Mbalmayo and Nioro. Bafia in savanna and Mbalmayo in forest, grown under yellow and ocher vertisol, respectively in Cameroun while Nioro in Sahel in Senegal exhibited leached ferruginous soil. The interaction between all agroecological scenarios (climate, vegetation and soil) and spatial factors create a complex system of environments that affect peanut plant growth and development, leading to a discrimination among genotypes within and between environments. As previously reported by chemical studies, seed composition is influenced by environment but also has a strong genetic component. The variation of oil composition has been related to temperature (Harris and James, 1969), planting date (Andersen and Gorbet, 2002), location and soil moisture (Holaday and Pearson, 1974; Young et al., 1974), photoperiod (Dwivedi et al., 2000), market grade (Mozingo et al., 1988) and genotype (Gimode et al., 2020; Harch et al., 1995; Holaday and Pearson, 1974; Norden et al., 1987; Worthington and Hammons, 1971). However, with multiple environmental factors mentioned above, it is difficult to decipher factors underlining variation in this study. Likewise, identifying suitable peanut genotypes for global ecological zones remains a challenging task due to the significant genotype variability across environments. Finally, the African varieties studied in one environment added genetic variability to the environmental one, resulting in a wide range of variability.

According to the spectra profiles and PCA plot, the genetic pattern of interspecific population covers, remarkably half of the spectrum of the core-collection, that turned out to be largest, as we expected. Interestingly, we found specific genetic variation among interspecific genotypes that was not subtle cover by the core-collection at the common Nioro environment. Interspecific genotypes with positive value on the main PCA axis were recorded as promising genotypes for quality traits. These genotypes could be recommended for further breeding for developing suitable varieties. In this respect, evaluation of the segregating interspecific population could further ease the discovery of QTL and valuable

wild genes that contribute to improved seed quality.

A PLS-DA model was successful developed from seed spectra to classify varieties and genotypes according to their genetic and environmental origin. A robust prediction accuracy of 99.6% was achieved. The confusion matrix achieved for the two environments, Nioro and Bafia shows 100% of instances classified correctly with 100% at both sensitivity and specificity. This confirm that their seed chemical composition was very different from each other and from those of the other seed samples. These results suggested that PLS-DA model could be used to classify peanut genotypes depending on the combination of the genetic and environment origins of seeds, which influence plant nutritional properties. In further studies, the current model would be confronted to wide others breeding populations in different environment to predict genetic and environment origin and nutritional content of whole seeds.

Breeding programs need germplasm diversity with extreme values for any nutritional trait. The magnitude of the genetic influence among varieties and genotypes suggested that nutritional related traits were amenable to improvement through intra and interspecific breeding. Nowadays, the availability of NIR data, might accelerate the utilization of germplasm and genetic diversity both in breeding programs. The observed genotypic variations and their variability across environments have deep implications for breeding programs. It seems feasible to achieve a fruitful goal in breeding on the basis of seed composition, because both the environmental effects found in the different locations and the genetic effects of the different varieties and interspecific derivatives influence the seed chemical compounds. Interestingly, even if the core-collection turned out to be the widest, a huge, specific and subtle genetic variation was found among interspecific genotypes, that was not covered by the 300 varieties. This offers the possibility of discovering new sources of diverse nutritional polymorphisms from wild derivatives. As early reported, three introgression lines with elevated Oleic/Linoleic profiles were found using chemical survey of 77 interspecific lines (Gimode et al., 2020). Interspecific hybridization has recently played an important role in accessing useful alleles from the wild (Favero et al., 2006; Mallikarjuna et al., 2011b; Simpson, 2001). We recorded varieties and interspecific lines with favorable spectra profiles. Thus, those potential chemotypes, with favorable chemical profiles could be further evaluated and promoted as a valuable genetic material to develop suitable varieties. Moreover, the comprehension of the genetic and environments determinants of nutritional traits might help in marker-assisted selection, accelerating the breeding of superior varieties tailored for specific environments and end-user demands.

5. Conclusion

The present study was carried out to investigate the potential of NIR coupled with chemometric to rapidly assess peanut germplasm from whole seed of a core-collection and an interspecific population, field-evaluated in 3 environments. This paper describes the NIR inputs to control breeding populations and assess germplasm variability, as we expected before the genetic studies. A wide variability of seed compounds was observed in the given germplasm, within and between environments, as revealed by spectra and multivariate analysis. A PLS-DA model was developed with a strong classification accuracy, aiming to properly predict each whole seed sample according to environment. These results indicate that NIR coupled with chemometric seem useful to accurately assess and distinguish intact-seed within different environments, that would ease further prediction of intact-seed nutritional content and utilization of germplasm in breeding programs.

CRediT authorship contribution statement

Fentanesh Chekole Kassie: Writing – original draft, Investigation. **Gilles Chaix:** Writing – review & editing, Validation, Formal analysis. **Hermine Bille Ngalle:** Supervision, Resources, Project administration, Funding acquisition. **Maguette Seye:** Investigation. **Coura Fall:** Investigation. **Hodo-Abalo Tossim:** Investigation. **Aissatou Sambou:** Investigation. **Olivier Gibert:** Writing – review & editing. **Fabrice Davrieux:** Data curation. **Joseph Martin Bell:** Supervision, Resources, Project administration, Funding acquisition. **Jean-François Rami:** Supervision, Resources, Project administration, Funding acquisition. **Daniel Fonceka:** Supervision, Resources, Project administration, Funding acquisition. **Joël Romaric Nguepjop:** Writing – review & editing, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the GENES intra-Africa Academic Mobility scheme of the European Union (EU-GENES: EACEA/2917/2552) and by the DESIRA-ABEE project funded by European Union. The authors thank Paulo Salgado, Rokhaya Diarra, Aissatou Bathily and Laurent Bonnal, for technical advices for spectra acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ocsci.2024.03.003.

References

Andersen, P.C., Gorbet, D.W., 2002. Influence of year and planting date on fatty acid chemistry of high oleic acid and normal peanut genotypes. J. Agric. Food Chem. 50, 1298–1305. https://doi.org/10.1021/jf0113171.

Bilal, M., Xiaobo, Z., Arslan, M., Tahir, H.E., Azam, M., Junjun, Z., Basheer, S., Abdullah, 2020. Rapid determination of the chemical compositions of peanut seed (Arachis hypogaea.) Using portable near-infrared spectroscopy. Vib. Spectrosc. 110, 103138. https://doi.org/10.1016/j.vibspec.2020.103138.

Bogie, N.A., Bayala, R., Diedhiou, I., Dick, R.P., Ghezzehei, T.A., 2018. Alteration of soil physical properties and processes after ten years of intercropping with native shrubs in the Sahel. Soil Tillage Res. 182, 153–163. https://doi.org/10.1016/j.still.2018.05.010.

Brandolini-Bunlon, M., Jaillais, B., Roger, J.M., Lesnoff, M., 2023. rchemo: Dimension Reduction, Regression and Discrimination for Chemometrics. R package version 0.1-1. https://CRAN.R-project.org/package=rchemo.

Burow, M.D., Simpson, C.E., Faries, M.W., Starr, J.L., Paterson, A.H., 2009. Molecular biogeographic study of recently described B- and A-genome Arachis species, also providing new insights into the origins of cultivated peanut. Genome 52, 107–119.

Conde, S., Rami, J.-F., Okello, D.K., Sambou, A., Muitia, A., Oteng-Frimpong, R., Makweti, L., Sako, D., Faye, I., Chintu, J., Coulibaly, A.M., Miningou, A., Asibuo, J.Y., Konate, M., Banla, E.M., Seye, M., Djiboune, Y.R., Tossim, H.-A., Sylla, S.N., Hoisington, D., Clevenger, J., Chu, Y., Tallury, S., Ozias-Akins, P., Fonceka, D., 2023. The Groundnut Improvement Network for Africa (GINA) Germplasm collection: a unique genetic resource for breeding and gene discovery. G3 GenesGenomesGenetics jkad244. https://doi.org/10.1093/g3journal/jkad244.

Davis, B.I., Agraz, C.B., Kline, M., Gottschall, E., Nolt, M., Whitaker, T.B., Osborne, J.A., Tengstrand, E., Ostrowski, K., Teixeira, R., Davis, J.P., 2021. Measurements of high oleic purity in peanut lots using rapid, single kernel near-infrared reflectance spectroscopy. J. Am. Oil Chem. Soc. 98, 621–632. https://doi.org/10.1002/aocs.12487.

Desmae, H., Janila, P., Okori, P., Pandey, M.K., Motagi, B.N., Monyo, E., Mponda, O., Okello, D., Sako, D., Echeckwu, C., Oteng-Frimpong, R., Miningou, A., Ojiewo, C., Varshney, R.K., 2019. Genetics, genomics and breeding of groundnut (Arachis hypogaea L.). Plant Breed. 138, 425–444. https://doi.org/10.1111/pbr.12645.

Dwivedi, S.L., Nigam, S.N., Rao, R.C.N., 2000. Photoperiod effects on seed quality traits in peanut. Crop Sci. 40, 1223–1227. https://doi.org/10.2135/cropsci2000.4051223x.

Fan, P., Song, W., Kang, Y., Wan, L., Lei, Y., Huai, D., Chen, Y., Wang, X., Jiang, H., Yan, L., Liao, B., 2020. Phenotypic identification of peanut germplasm for resistance to southern stem rot. Oil Crop Sci 5, 174–179. https://doi.org/10.1016/j.ocsci.2020.12.001.

FAOSTAT, 2020. Statistical Database. Food and Agriculture Organization of the United Nations, Rome.

Favero, A.P., Simpson, C.E., Valls, J.F.M., Vello, N.A., 2006. Study of the evolution of cultivated peanut through crossability studies among Arachis ipaensis, A. Duranensis, and A. hypogaea. Crop Sci. 46, 1546–1552. https://doi.org/10.2135/cropsci2005.09-0331.

Fonceka, D., Tossim, H.-A., Rivallan, R., Vignes, H., Faye, I., Ndoye, O., Moretzsohn, M.C., Bertioli, D.J., Glaszmann, J.-C., Courtois, B., Rami, J.-F., 2012a. Fostered and left

behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. BMC Plant Biol. 12, 26. https://doi.org/10.1186/1471-2229-12-26.
Fonceka, D., Tossim, H.-A., Rivallan, R., Vignes, H., Lacut, E., de Bellis, F., Faye, I., Ndoye, O., Leal-Bertioli, S.C.M., Valls, J.F.M., Bertioli, D.J., Glaszmann, J.-C., Courtois, B., Rami, J.-F., 2012b. Construction of chromosome segment substitution lines in peanut (Arachis hypogaea L.) using a wild synthetic and QTL mapping for plant morphology. PLoS One 7, e48642. https://doi.org/10.1371/journal.pone.0048642.
Fordellone, M., Bellincontro, A., Mencarelli, F., 2019. Partial least squares discriminant analysis: a dimensionality reduction method to classify hyperspectral data. Stat. Appl. - Ital. J. Appl. Stat. 181–200. https://doi.org/10.26398/IJAS.0031-010.
Gimode, D., Chu, Y., Dean, L., Holbrook, C., Fonceka, D., Ozias-Akins, P., 2020. Seed composition survey of a peanut CSSL population reveals introgression lines with elevated oleic/linoleic profiles. Peanut Sci. 47, 139–149. https://doi.org/10.3146/PS20-17.1.
Govindarajan, K.N., Kandala, V., Subbiah, J., 2009. NIR reflectance spectroscopy for nondestructive moisture content determination in peanut kernels. Trans. ASABE (Am. Soc. Agric. Biol. Eng.) 52. https://doi.org/10.13031/2013.29117.
Grosso, N.R., Lamarque, A., Maestri, D.M., Zygadlo, J.A., Guzmán, C.A., 1994. Fatty acid variation of runner peanut (Arachis hypogaea L.) among geographic localities from Córdoba (Argentina). J. Am. Oil Chem. Soc. 71, 541–542. https://doi.org/10.1007/BF02540669.
Grosso, N.R., Nepote, V., Guzman, C.A., 2000. Chemical composition of some wild peanut species (Arachis L.) seeds. J. Agric. Food Chem. 48, 806–809. https://doi.org/10.1021/jf9901744.
Harch, B.D., Basford, K.E., DeLacy, I.H., Lawrence, P.K., Cruickshank, A., 1995. Patterns of diversity in fatty acid composition in the Australian groundnut germplasm collection. Genet. Resour. Crop Evol. 42, 243–256. https://doi.org/10.1007/BF02431259.
Harris, P., James, A.T., 1969. The effect of low temperatures on fatty acid biosynthesis in plants. Biochem. J. 112, 325–330.
Holaday, C.E., Pearson, J.L., 1974. Effects of genotype and production area on the fatty acid composition, total oil and total protein in peanuts. J. Food Sci. 39, 1206–1209.
Huang, H., Liu, L., Ngadi, M.O., 2014. Recent developments in hyperspectral imaging for assessment of food quality and safety. Sensors 14, 7248–7276. https://doi.org/10.3390/s140407248.
Isleib, T.G., Tillman, B.L., Pattee, H.E., Sanders, T.H., Hendrix, K.W., Dean, L.O., 2008. Genotype-by-Environment interactions for seed composition traits of breeding lines in the uniform peanut performance test. Peanut Sci. 35 (1), 130–138. https://doi.org/10.3146/PS08-001.
Janila, P., Nigam, S.N., Pandey, M.K., Nagesh, P., Varshney, R.K., 2013. Groundnut improvement: use of genetic and genomic tools. Front. Plant Sci. 4, 23. https://doi.org/10.3389/fpls.2013.00023.
Kandala, C.V.K., Butts, C.L., Lamb, M.C., 2008. Moisture content determination for in-shell peanuts with a low-cost impedance analyzer and capacitor sensor. Moisture Content Determ. -Shell Peanuts Low-Cost Impedance Anal. Capacit. Sens. 51, 1377–1381.
Kumari, V., Gowda, M.V.C., Tasiwal, V., Pandey, M.K., Bhat, R.S., Mallikarjuna, N., Upadhyaya, H.D., Varshney, R.K., 2014. Diversification of primary gene pool through introgression of resistance to foliar diseases from synthetic amphidiploids to cultivated groundnut (Arachis hypogaea L.). Crops J 2, 110–119. https://doi.org/10.1016/j.cj.2014.03.002.
Lesnoff, M., 2021. R package rnirs: dimension reduction, regression and discrimination for chemometrics [WWW Document]. https://github.com/mlesnoff/rnirs.URL. (Accessed 8 January 2023).
Li, Z., Tang, X., Shen, Z., Yang, K., Zhao, L., Li, Y., 2019. Comprehensive comparison of multiple quantitative near-infrared spectroscopy models for Aspergillus flavus contamination detection in peanut. J. Sci. Food Agric. 99, 5671–5679. https://doi.org/10.1002/jsfa.9828.
Liu, L., Jiang, H., Chen, Q., 2022. Rapid determination of acidity index of peanut during storage by a portable near-infrared spectroscopy system. Infrared Phys. Technol. 127, 104472. https://doi.org/10.1016/j.infrared.2022.104472.
Mallikarjuna, N., Jadhav, D.R., Reddy, K., Husain, F., Das, K., 2011a. Screening new Arachis amphidiploids, and autotetraploids for resistance to late leaf spot by detached leaf technique. Eur. J. Plant Pathol. 132, 17–21. https://doi.org/10.1007/s10658-011-9859-2.
Mallikarjuna, N., Senthilvel, S., Hoisington, D., 2011b. Development of new sources of tetraploid Arachis to broaden the genetic base of cultivated groundnut (Arachis hypogaea L.). Genet. Resour. Crop Evol. 58, 889–907. https://doi.org/10.1007/s10722-010-9627-8.
Mozingo, R.W., Coffelt, T.A., Wynne, J.C., 1988. Market grade effects on fatty acid composition of five peanut cultivars. Agron. J. 80, 73–75. https://doi.org/10.2134/agronj1988.00021962008000010017x.
Nawade, B., Mishra, G.P., Radhakrishnan, T., Dodia, S.M., Ahmad, S., Kumar, Abhay, Kumar, Atul, Kundu, R., 2018. High oleic peanut breeding: achievements, perspectives, and prospects. Trends Food Sci. Technol. 78, 107–119. https://doi.org/10.1016/j.tifs.2018.05.022.
Nguepjop, J.R., Tossim, H.-A., Bell, J.M., Rami, J.-F., Sharma, S., Courtois, B., Mallikarjuna, N., Sane, D., Fonceka, D., 2016. Evidence of genomic exchanges between homeologous chromosomes in a cross of peanut with newly synthetized allotetraploid hybrids. Front. Plant Sci. 7, 1635. https://doi.org/10.3389/fpls.2016.01635.
Norden, A.J., Gorbet, D.W., Knauft, D.A., Young, C.T., 1987. Variability in oil quality among peanut genotypes in the Florida breeding program 1. Peanut Sci. 14, 7–11.
Panero, F.D.S., Smiderle, O., Panero, J.S., Faria, F.S.D.V., Panero, P.D.S., Rodriguez, A.F.R., 2022. Non-destructive genotyping of cultivars and strains of sesame through NIR spectroscopy and chemometrics. Biosensors 12, 69. https://doi.org/10.3390/bios12020069.
Panero, J.S., Silva, H.E.B. da, Panero, P.S., Smiderle, O.J., Panero, F.S., Faria, F.S.E.D.V., Rodriguez, A.F.R., 2018. Separation of cultivars of soybeans by chemometric methods using near infrared spectroscopy. J. Agric. Sci. 10, p351. https://doi.org/10.5539/jas.v10n4p351.
Parilli-Moser, I., Hurtado-Barroso, S., Guasch-Ferré, M., Lamuela-Raventós, R.M., 2022. Effect of peanut consumption on cardiovascular risk factors: a randomized clinical trial and meta-analysis. Front. Nutr. 9.
Parmar, S., Sharma, V., Bomireddy, D., Soni, P., Joshi, P., Gangurde, S., Wang, J., Bera, S.K., Bhat, R., Desmae, H., Shirasawa, K., Varshney, R., Pandey, M., 2022. Recent Advances in Genetics, Genomics, and Breeding for Nutritional Quality in Groundnut, pp. 111–137. https://doi.org/10.1007/978-3-030-81107-5_4.
R Core Team, 2021. R: the R Project for Statistical Computing [WWW Document]. https://www.r-project.org/. (Accessed 8 January 2023).
Rabanera, J.D., Guzman, J.D., Yaptenco, K.F., 2021. Rapid and Non-destructive measurement of moisture content of peanut (Arachis hypogaea L.) kernel using a near-infrared hyperspectral imaging technique. J. Food Meas. Char. 15, 3069–3078. https://doi.org/10.1007/s11694-021-00894-x.
Sampaio, P.S., Brites, C.M., Sampaio, P.S., Brites, C.M., 2021. Near-infrared spectroscopy and machine learning: analysis and classification methods of rice. Integrative Advances in Rice Research. IntechOpen. https://doi.org/10.5772/intechopen.99017.
Simpson, C.E., 2001. Use of wild Arachis species/introgression of genes into A. hypogaea L. Peanut Sci. 28, 114–116. https://doi.org/10.3146/i0095-3679-28-2-12.
Simpson, C.E., Krapovickas, A., Valls, J.F.M., 2001. History of Arachis including evidence of A. hypogaea L. Progenitors. Peanut Sci. 28, 78–80. https://doi.org/10.3146/i0095-3679-28-2-7.
Singh, T., Garg, N.M., Iyengar, S.R.S., Singh, V., 2023. Near-infrared hyperspectral imaging for determination of protein content in barley samples using convolutional neural network. J. Food Meas. Char. 1–13.
Snee, R.D., 1977. Validation of regression models: methods and examples. Technometrics 19, 415–428. https://doi.org/10.1080/00401706.1977.10489581.
Song, W., Wang, H., Maguire, P., Nibouche, O., 2018. Nearest clusters based partial least squares discriminant analysis for the classification of spectral data. Anal. Chim. Acta 1009, 27–38. https://doi.org/10.1016/j.aca.2018.01.023.
Sundaram, J., Kandala, C.V., Butts, C.L., 2009. Application of near infrared spectroscopy to peanut grading and quality analysis: overview. Sens. Instrum. Food Qual. Saf. 3, 156–164. https://doi.org/10.1007/s11694-009-9081-5.
Tang, Y., Qiu, X., Hu, C., Li, J., Wu, L., Wang, W., Li, Xin, Li, Xiaoting, Zhu, H., Sui, J., Wang, J., Qiao, L., 2022. Breeding of a new variety of peanut with high-oleic-acid content and high-yield by marker-assisted backcrossing. Mol. Breed. New Strateg. Plant Improv. 42, 42. https://doi.org/10.1007/s11032-022-01313-9.
Tao, F., Yao, H., Hruska, Z., Liu, Y., Rajasekaran, K., Bhatnagar, D., 2019. Use of visible–near-infrared (Vis-NIR) spectroscopy to detect aflatoxin B1 on peanut kernels. Appl. Spectrosc. 73, 415–423. https://doi.org/10.1177/0003702819829725.
Temga, P., Sababa, E., Mamdem, L., Bijeck, M., Azinwi Tamfuh, P., Tehna, N., Zame, P., Onana, V.L., Jean Pierre, N., Bitom, L., Ndjigui, P.-D., 2021. Rare earth elements in tropical soils, Cameroon soils (Central Africa). Geoderma Reg. 25, e00369. https://doi.org/10.1016/j.geodrs.2021.e00369.
Tian, W., Zang, L., Nie, L., Li, L., Zhong, L., Guo, X., Huang, S., Zang, H., 2023. Structural analysis and classification of low-molecular-weight hyaluronic acid by near-infrared spectroscopy: a comparison between traditional machine learning and deep learning. Molecules 28, 809. https://doi.org/10.3390/molecules28020809.
Tossim, H.-A., Nguepjop, J.R., Diatta, C., Sambou, A., Seye, M., Sane, D., Rami, J.-F., Fonceka, D., 2020. Assessment of 16 peanut (Arachis hypogaea L.) CSSLs derived from an interspecific cross for yield and yield component traits: QTL validation. Agronomy 10, 583. https://doi.org/10.3390/agronomy10040583.
Upadhyaya, H.D., 2005. Variability for drought resistance related traits in the mini core collection of peanut. Crop Sci. 45, 1432–1440. https://doi.org/10.2135/cropsci2004.0389.
Upadhyaya, H.D., Dwivedi, S.L., Nadaf, H.L., Singh, S., 2011. Phenotypic diversity and identification of wild Arachis accessions with useful agronomic and nutritional traits. Euphytica. https://doi.org/10.1007/s10681-011-0518-7.
Wang, Ling, Wang, Lei, Tan, M., Yu, H., Wang, J., Li, Y., Wang, W., Yan, X., Wang, Lijun, 2022. Rapid identification and preliminary evaluation of quality characters of oilseed sunflower by near infrared spectroscopy. Oil Crop Sci 7, 142–148. https://doi.org/10.1016/j.ocsci.2022.08.003.
Wang, Y., Song, S., 2023. Variety identification of sweet maize seeds based on hyperspectral imaging combined with deep learning. Infrared Phys. Technol. 130, 104611. https://doi.org/10.1016/j.infrared.2023.104611.
Wang, Z., Zhang, Y., Yan, L., Chen, Y., Kang, Y., Huai, D., Wang, X., Liu, K., Jiang, H., Lei, Y., Liao, B., 2023. Correlation and variability analysis of yield and quality related traits in different peanut varieties across various ecological zones of China. Oil Crop Sci 8, 236–242. https://doi.org/10.1016/j.ocsci.2023.12.001.
Worthington, R.E., Hammons, R.O., 1971. Genotypic Variation in Fatty Acid Composition and Stability of Arachis hypogaea L. Oil. Oleagineux.

Xu, P., Sun, W., Xu, K., Zhang, Y., Tan, Q., Qing, Y., Yang, R., 2023. Identification of defective maize seeds using hyperspectral imaging combined with deep learning. Foods 12, 144. https://doi.org/10.3390/foods12010144.

Young, C.T., Worthington, R.E., Hammons, R.O., Matlock, R.S., Waller, G.R., Morrison, R.D., 1974. Fatty acid composition of Spanish peanut oils as influenced by planting location, soil moisture conditions, variety, and season. J. Am. Oil Chem. Soc. 51, 312–315. https://doi.org/10.1007/BF02633004.

Yu, H., Liu, H., Erasmus, S.W., Zhao, S., Wang, Q., van Ruth, S.M., 2020. Rapid high-throughput determination of major components and amino acids in a single peanut kernel based on portable near-infrared spectroscopy combined with chemometrics. Ind. Crops Prod. 158, 112956. https://doi.org/10.1016/j.indcrop.2020.112956.

genes

MDPI

Review

An Overview of Mapping Quantitative Trait Loci in Peanut (*Arachis hypogaea* L.)

Fentanesh C. Kassie [1,2], Joël R. Nguepjop [3,4,5,6,*], Hermine B. Ngalle [1], Dekoum V. M. Assaha [7], Mesfin K. Gessese [2], Wosene G. Abtew [8], Hodo-Abalo Tossim [5,6], Aissatou Sambou [5,6], Maguette Seye [5,6], Jean-François Rami [3,4,6], Daniel Fonceka [3,4,5,6] and Joseph M. Bell [1]

1 Department of Plant Biology and Physiology, Faculty of Sciences, University of Yaounde I, Yaounde P.O. Box 337, Cameroon
2 Department of Plant Science, College of Agriculture, Wolaita Sodo University, Sodo P.O. Box 138, Ethiopia
3 UMR AGAP, CIRAD, F-34398 Montpellier, France
4 AGAP Institute, Institut Agro, CIRAD, INRAE, University of Montpellier, F-34060 Montpellier, France
5 Centre d'Etudes Régional Pour l'Amélioration de l'Adaptation à la Sécheresse (CERAAS/ISRA), Route de Khombole, Thiès BP 3320, Senegal
6 Dispositif de Recherche et de Formation en Partenariat, Innovation et Amélioration Variétale en Afrique de l'Ouest (IAVAO), CERAAS, Route de Khombole, Thiès BP 3320, Senegal
7 Department of Agriculture, Higher Technical Teachers Training College, University of Buea, Kumba P.O. Box 249, Cameroon
8 Department of Horticulture and Plant Science, College of Agriculture and Veterinary Medicine, Jimma University, Jimma P.O. Box 378, Ethiopia
* Correspondence: joel-romaric.nguepjop@cirad.fr

Abstract: Quantitative Trait Loci (QTL) mapping has been thoroughly used in peanut genetics and breeding in spite of the narrow genetic diversity and the segmental tetraploid nature of the cultivated species. QTL mapping is helpful for identifying the genomic regions that contribute to traits, for estimating the extent of variation and the genetic action (i.e., additive, dominant, or epistatic) underlying this variation, and for pinpointing genetic correlations between traits. The aim of this paper is to review the recently published studies on QTL mapping with a particular emphasis on mapping populations used as well as traits related to kernel quality. We found that several populations have been used for QTL mapping including interspecific populations developed from crosses between synthetic tetraploids and elite varieties. Those populations allowed the broadening of the genetic base of cultivated peanut and helped with the mapping of QTL and identifying beneficial wild alleles for economically important traits. Furthermore, only a few studies reported QTL related to kernel quality. The main quality traits for which QTL have been mapped include oil and protein content as well as fatty acid compositions. QTL for other agronomic traits have also been reported. Among the 1261 QTL reported in this review, and extracted from the most relevant studies on QTL mapping in peanut, 413 (~33%) were related to kernel quality showing the importance of quality in peanut genetics and breeding. Exploiting the QTL information could accelerate breeding to develop highly nutritious superior cultivars in the face of climate change.

Keywords: QTL mapping; quality traits; interspecific; genetic; breeding; peanut

check for updates

Citation: Kassie, F.C.; Nguepjop, J.R.; Ngalle, H.B.; Assaha, D.V.M.; Gessese, M.K.; Abtew, W.G.; Tossim, H.-A.; Sambou, A.; Seye, M.; Rami, J.-F.; et al. An Overview of Mapping Quantitative Trait Loci in Peanut (*Arachis hypogaea* L.). *Genes* **2023**, *14*, 1176. https://doi.org/10.3390/genes14061176

Academic Editors: Liying Yan and Huaiyong Luo

Received: 17 April 2023
Revised: 24 May 2023
Accepted: 25 May 2023
Published: 28 May 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Peanut (*Arachis hypogaea* L.) is a grain legume mainly grown in the tropics, subtropics, and warm temperate regions of the world. The genus *Arachis* originated in South America and all of its species produce their fruit underground [1]. Peanut is a self-pollinated, segmental allotetraploid [2,3], with 2n = 4x = 40 chromosomes [3,4]. It is an oilseed crop with global importance to food and nutritional security and a source of livelihood for millions of smallholder growers of Asia and Sub-Saharan Africa. World production of peanut was approximately 54 million metric tons (MT) harvested from 30 million hectares

(Mha) in 2020. China is the world's largest producer with 18 million metric tons (MT). Africa accounts for 32 % of worldwide production and the annual production and harvested area were 17 MT and 17.43 Mha, respectively, in 2020 [5]. Peanut ranked fifth among vegetable oilseed crops in terms of edible oil production (6.26 MT), preceded respectively by sunflower seed (21.56 MT), rapeseed (27.85 MT), soybean (57.74 MT), and palm (72.77 MT) [6].

Peanut is a major oilseed crop used for a variety of purposes, such as direct consumption of the seed (kernel), which can be eaten raw, roasted, boiled, or processed into confectionary and peanut flour for flavor enhancement, or crushed to edible oil. Nutritionally, it is a source of high-quality edible oil (35–60%), protein (22–30%), carbohydrates (10–25%), vitamins (E, K, and B complex), and minerals (Ca, P, Mg, Zn, and Fe) [7]. Peanut oil contains about 12 different kinds of fatty acids (FAs), with oleic acid (C18:1) and linoleic acid (C18:2) accounting for nearly 80% of the total [8]. The presence of relative proportions of various FAs affects the nutritional quality, flavor, and shelf-life of peanut seeds and products. The high linoleic acid content in peanut oil induces low oxidative and frying stability, resulting in rancidity, off-flavors, and short shelf life in produced foods [8,9]. Compared to a normal ratio, a high oleic-to-linoleic ratio leads to longer shelf-life and improved flavors [10]. The main consumption and production constraints of the crop include drought, pests, diseases, and environmental changes. These constraints have an impact on the content of oil and protein present, as well as the oil's composition, which also has an indirect impact on the oil's shelf life, aroma, flavor, cooking quality, and cooking time [4].

According to genetic, cytogenetic, phylogeography, and molecular data, the cultivated peanut, *A. hypogaea* is an allotetraploid, derived from hybridization between the diploids, *A. duranensis* (*A genome*) and *A. ipaensis* (*B genome*) [3,11]. These two species are members of the section *Arachis* [12]. Molecular analysis has shown that cultivated peanut has a limited polymorphism at the DNA level due to the crop origin, from single to a few hybridization events and the transition from diploid to tetraploid [1]. Various studies have revealed that wild relatives harbor high levels of genetic variation [13,14] that can be used for improving cultivated peanut. Several accessions of those species have been used to board the genetic diversity of cultivated peanut in different studies for a variety of traits including abiotic stress [15], disease resistance [16,17], oil content, and composition [18].

Chromatin introgression of a tiny fraction of the wild species genome while maintaining the genome background of the cultivated peanut is a means to discover the most untapped wild genes/alleles. As several authors have mentioned, direct gene transfer from wild diploid species has been hampered by ploidy differences, fertility barriers caused by species incompatibilities, linkage drag of desirable wild alleles with those conferring agronomically unadopted traits and, finally, difficulties in confirming hybrid identities and tracking introgressed segments [16,19,20]. These issues have been partly resolved by the production of synthetic allotetraploid that can be crossed with cultivated peanut [21–23]. as well as a number of molecular markers and genomics tools to ease introgression and genetic analysis [24]. GPBD 4, Span cross, Tamnut 74, TxAG 7, COAN, NemaTAM, and Tifguard are improved germplasm/cultivars that were developed using genes from wild *Arachis* species [13,24]. Among them, GPBD4 resulting from a cross between *A. hypogaea* and *A. cardenasii* derived introgression line has been widely used for improving disease resistance in a variety of breeding programs [25].

Mapping quantitative trait loci (QTL) and identifying markers that are linked to target traits are important steps toward accelerating the rate of genetic gains in breeding programs. QTL mapping is a routine technique used to identify genetic loci governing traits of interest [26]. QTL mapping in family-based populations requires (i) the development of appropriate mapping population and traits phenotyping; (ii) selection of appropriate molecular marker(s) and generation of molecular data with an adequate number of uniformly-spaced polymorphic markers; (iii) construction of genetic linkage maps to locate QTL using statistical programs.

The success of QTL mapping depends on the size of the mapping populations and the quality of the genotyping and phenotyping data. The availability of a tremendous number

of genomic resources, including molecular markers, and genetic and physical maps have greatly eased the mapping of QTL and/or genes [27,28]. Many QTL have been reported for seed and pod-related traits [15,29–35], fresh seed or seed dormancy [36–40], nutritional quality traits [18,41–47], drought resistance [48–50], and disease resistance [51–56], providing potential tools for peanut improvement. However, few QTL are being effectively utilized in peanut improvement programs to produce elite cultivars.

The general aim of this paper is to review the recently mapped QTL in peanut, with a particular emphasis on traits related to kernel quality. Genome-wide distribution of QTL and their effective use in innovative schemes of marker-assisted selection (MAS) in peanut breeding are discussed.

2. Mapping Populations of QTL in Peanut

Genetic mapping populations are required for marker–trait associations both for oligogenic and polygenic traits. Genetic mapping can be broadly divided into two types: (i) family-based mapping, which is conducted on offspring of biparental or multiparent crosses, and (ii) natural population-based mapping, also known as association or linkage disequilibrium mapping, which is conducted on unrelated natural populations [26].

2.1. Family-Based Mapping Populations

The biparental mapping population is a kind of family-based mapping population created by crossing two parents that differ on the target of interest [57]. F2 [29,36,43,55,58], backcrosses (BC) [15,18,59–63], recombinant inbred lines (RIL) [44–50], chromosome segment substitution lines (CSSL) [64–66], and near-isogenic lines (NIL) [67] are the examples of biparental mapping populations. In peanut, F2, RIL, advanced backcross (AB) populations, and CSSL have been used for QTL mapping (Figure 1). The relative simplicity of construction, the high QTL detectability, and the low rate of linkage disequilibrium decay within chromosomes are the three main benefits of biparental populations. However, there are two main limitations of biparental populations: a lack of mapping precision due to the limited amount of effective recombination that occurs during population development and low genetic diversity because of the genetic bottleneck caused by the selection of two founders [68].

To overcome the limitations of biparental populations, multiparental mapping populations have been created; these populations increase the recombination rates and the genetic pool, leading to higher-resolution genetic maps [68–70]. Usually, they are suitable for high-resolution QTL mapping, although they also have some limitations, such as labor-intensive crossing, managing large population sizes, being time-consuming, and requiring significant investments in phenotyping and genotyping (Table 1). Nested association mapping (NAM) and multi-parent advanced generation inter-crossing (MAGIC) [71], are the two most important multiparent mapping populations [68]. Four NAM populations, ICGV 91114 and 22 genotypes, and ICGS 76 and 21 genotypes, using Florida-07 and Tifrunner as common parents with eight genotypes, were developed at ICRISAT and by USDA ARS and the University of Georgia (UGA), respectively [69,70,72]. Using the last two populations, a total of 42 SNP markers linked with 100-pod weight and 100-seed weight were detected [73]. Three MAGIC populations also were developed by crossing eight parental genotypes targeting multiple traits, including fresh seed dormancy, oil content, aflatoxin, and drought resistance [69]. Despite the discovery of QTL mainly using linkage analysis, natural population-based mapping has also been used in a couple of studies [74–80]. It is discussed in the next section.

Table 1. Advantages and disadvantages of populations used in peanut mapping.

Populations	Advantages	Disadvantages	References
BC	- Useful for the introgression of wild chromatin and specific genes/QTL - Easy analysis	- Time requirement to construct in such cases - Estimating dominance effects is impossible - Not suitable for Meiotic behavior analysis	[15,59,60]
F2	- Require less time to construct - Impossible to determine the degree of dominance and additive effects - Easy analysis - Suitable for Meiotic behavior analysis	- Precision is low - Temporary nature - It is not repeated across years and locations	[2,43,55,58]
NIL	- Suitable for fine mapping - Useful for tagging the gene - Highly reliable and accurate statistically - Suitable for quantitative and qualitative trait tagging	- It takes time to construct - Not suitable for whole linkage mapping - Problem of linkage drag	[67]
RIL	- Abundance of recombination - Immortality: replicable throughout locations and years - Very useful in identifying tightly linked markers	- Impossible to estimate dominant effects - Time requirement: many seasons and generations are needed for the development	[44–50]
CSSL	- Immortality: replicable throughout locations and years - Very useful in identifying tightly linked markers	- Time requirement: many seasons and generations are needed for the development	[65]
NAM or MAGIC	- Several alleles than biparental populations - Several QTL segregating than biparental populations - Rapid fine mapping - Useful for candidate	- Time to construction/establish - Require more markers than biparental - Require larger population than biparental	[69–73]
Natural	- Available collections - High diversity - Natural recombination - When LD limited, Precise mapping	- Time to construction/establish - Require more markers than biparental and MAGIC or NAM population - Population structure - Spurious association - When high LD, coarse mapping - Rare alleles poorly identified	[74–80]

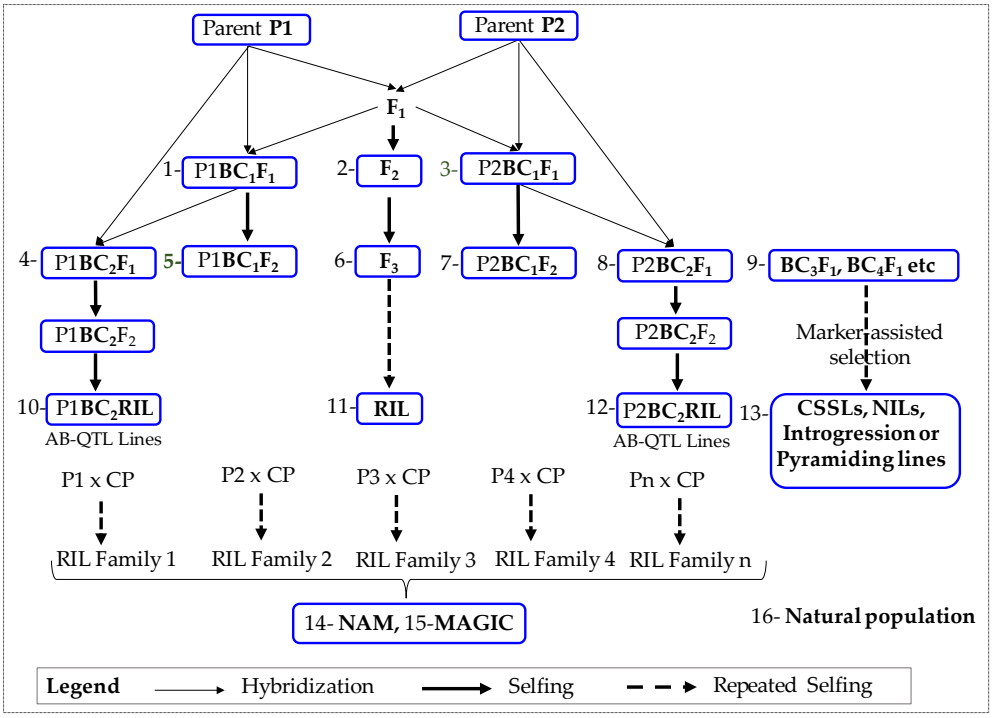


Figure 1. Derivation of biparental and other populations used for peanut. F2 [29,36,43,55,58], BC2F2 [15], BC1F1 [59], BC2F4 [60–62], BC2RIL [18], BC2F1 [63], RIL [44–48], CSSL [64–66], NIL [67] MAGIC [69,71], NAM [69,70,73], Natural population [74–78].

2.2. Natural-Based Mapping Populations

Natural population-based mapping is a method of mapping QTL that takes advantage of historic linkage disequilibrium to link phenotype to genotype by sampling distantly related individuals. However, the method comes with limitations: it is predominantly influenced by unknown population structure, leading to spurious associations, and also requires very large samples to have sufficient power to detect genomic regions of interest [74–76]. Natural population-based mapping utilizes diverse germplasm sets with high variability for economically important traits in a crop species with the advantages, in such cases, of high resolution and high allelic richness, with no investment in crossing. The main steps in natural population-based mapping include (i) collection of a sample population including elite cultivars, landraces, wild relatives, and exotic accessions; (ii) phenotyping target traits, estimation of broad-sense heritability, genotyping the population; (iii) quantification of LD extent of the selected population; (iv) identification of the influence of population structure and kinship; and (v) testing the association between genotype and phenotype using appropriate statistical approaches and validation of detected QTL [26,28].

Using this approach, several QTL have been reported in peanut. Two functional single nucleotide polymorphism (SNP) markers for two fatty acid desaturases (FAD2 for oleic acid, linoleic acid, and oleic-to-linoleic ratio) were found by phenotyping for quality traits and genotyping of the US "Mini Core Collection" using 81 SSR markers [74]. A total of 50 SSR markers linked with oil content, protein content, oleic to linoleic acid ratio, and fatty acid concentrations, with phenotypic variation explained (PVE) from 5.81 to 47.45 percent, were detected using 300 genotypes. Additionally, 12 QTL that linked to seed length and seed width, explaining phenotypic variance from 11.81 to 30.09 percent were also detected using these genotypes [75]. The other 107 significant SNP markers underlying pod weight, pod length, pod width, seed length, seed width, and 100-pod weight and 100-seed weight [76,79] were discovered using 158 and 250 genotypes, respectively. Similarly, 12 QTL associated with oil content were identified using 292 accession numbers [78], and 253 loci controlling oil content, protein content, oleic to linoleic acid ratio, and fatty acid composition were

identified [77,80] using 120 and 250 genotypes, respectively. Among all the genotypes that have been used as a mapping population, more than 50% are from core collections. QTL that have been mapped for important quantitative traits, particularly utilizing a biparental approach, have been discussed in the next sections.

3. Molecular Markers for Linkage and QTL Mapping

A major application of molecular markers is the construction of linkage maps required for QTL mapping and marker-assisted breeding. Past molecular markers used in peanut include RFLP, RAPD, and AFLP. The first RFLP-based map for *Arachis* was created with 117 loci [81]. The first tetraploid RFLP-based genetic map was developed with 370 loci using an advanced population [60]. RAPD and RFLP markers were used to create a genetic linkage map using an interspecific diploid BC population derived from *A. stenosperma* x (*A. stenosperma* x *A. cardenasii)* [59]. In this study, 167 RAPD and 39 RFLP loci were mapped to 11 linkage groups, covering 800 cM. AFLP markers were also used to find DNA markers linked to aphid resistance and to construct a partial genetic linkage map of cultivated peanut [82]. Currently, the most commonly used molecular markers in peanut include SSR, SNP, and DArT. In 2009, the first genetic map based on SSR markers was constructed with 135 loci [83]. In the same year, 298 loci were mapped in 21 linkage groups, spanning a total map distance of 1843.7 cM in an advanced backcross population [84]. Among many SSR-based maps, these were developed for populations derived from TG26 and GPBD4, Sun Oleic 97R, NC94022, Tifrunner and GT-C20, and amphidiploid ''TxAG-6'' and ''Florunner'', to locate the genomic region underlying seed quality traits [18,41,42]. The maps contain 45 to 378 SSR loci spanning 671.1 to 2487.4 cM distance. In addition, with these maps, approximately 33 SSR-based genetic maps have been developed to date to identify the genomic areas responsible for disease resistance, drought resistance, yield, and yield component traits as reviewed previously [7,27,72].

SNP are the most common molecular markers in the genome, and they can be analyzed with high-throughput genotyping techniques. The first SNP-based genetic map in peanut was created with 1621 SNPs and 64 SSR markers on 20 linkage groups [85]. Aiming at oil and protein content and oil composition, three linkage maps were constructed with 2266–4561 SNP loci spanning a total map length ranging from 2032.39–2586.37 cM derived from Huayu28 and P76, Xuhua 13 and Zhonghua 6, and Yuhua 15 and W1202 [44,45,47], on 20 linkage groups. To date, SNP-based maps have been presented in peanut [7,27,72]. By using DArT markers, five genetic maps were constructed using F2 and advanced backcross populations. A genetic map using the F2 population derived from ICGV 00350 and ICGV 97045 has 1152 loci spanning a map distance of 2423.12 cM and a map density of 2.96 cM/loci developed [36]. A total of 854 (ICGV 07368 and ICGV 06420) and 1435 (ICGV 06420 and SunOleic 95R) marker loci were used to create two genetic maps, with total map distances of 3526 and 1869 cM, respectively [43]. The other two additional genetic maps, with 253 DArT and five SSR loci, and 1035 DArT and eight SSR loci, covering 1415.7 and 1500.8 cM of map length, respectively, were created using advanced backcross populations [63].

The availability of genome sequencing for peanut speeds up the development of different types of genotyping platforms/assays, including Kompetitive Alelle Specific PCR (KASP) assays, Golden Gate assays, Vera-code assays, micro-array-based markers, next-generation sequencing (NGS)-based markers, genotyping by sequencing (GBS), InDel markers and Affymetrix axiom SNP array. All these genotyping platforms are SNP-based since SNP markers are considered markers of choice and are amenable to high-throughput genotyping for several applications including QTL mapping.

4. Mapping of QTL

A QTL is a genomic region that is responsible for the quantitative variation of a trait. A quantitative trait is a measurable attribute based on the combined activity of one or many genes and their interactions with the environment, which can vary between

individuals over a given range to generate a continuous distribution of phenotypes [86]. QTL mapping is important for identifying responsible genes, understanding variation mechanisms, determining how many QTL contribute significantly to the trait, determining how much variation is due to additive, dominant and epistatic effects, and determining the nature of the genetic correlation between different traits in a genomic region [87]. The steps involved in biparental QTL mapping are presented (Figure 2). To date, quantitative or metric traits in peanut include traits related to yield and yield component traits, flowering, agro-morphology, seed dormancy, quality and nutritional traits, and resistance to viral, bacterial, and fungal diseases.

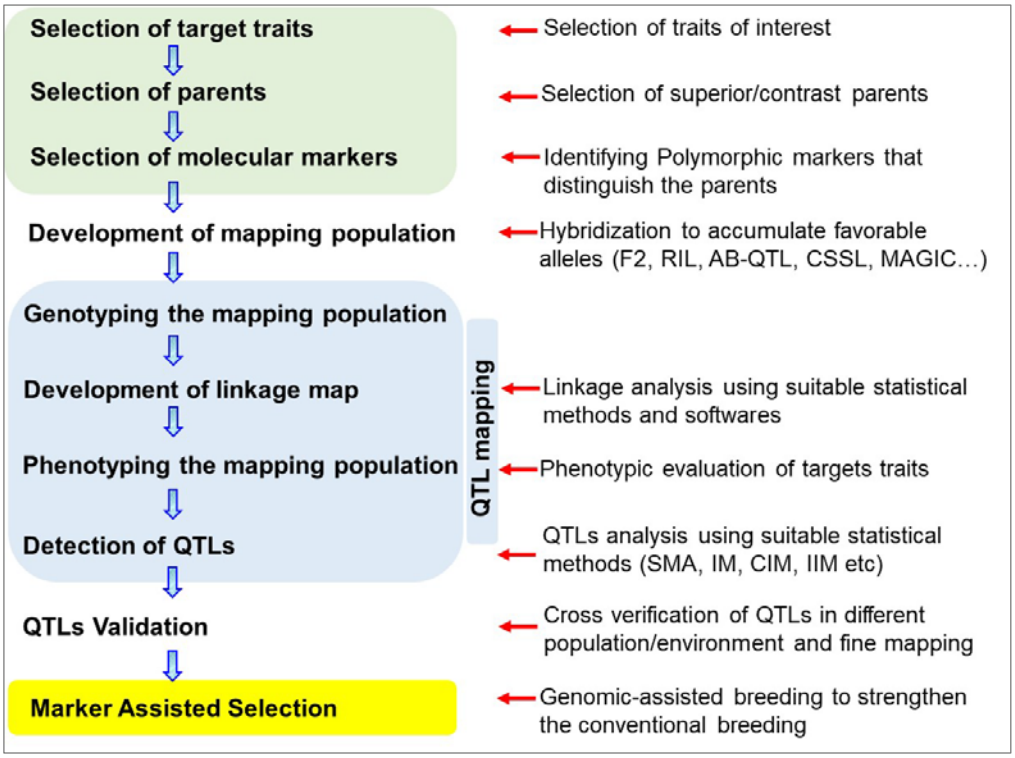


Figure 2. Steps involved in biparental QTL mapping and further use of the QTL in breeding programs.

4.1. Mapping QTL for Seed Quality Traits

Important quality traits that can be assessed by biochemical analysis of the peanut kernel include oil, protein, and sugar content, as well as fatty acid (FA), amino acid, and carbohydrate composition. The proportion of different FAs, such as saturated, monounsaturated, and polyunsaturated (PUFA), present in the oil determines the nutritional quality, flavor, and shelf life of both peanut kernels and products [8]. In particular, the concentration of oleic acid is one of the most important quality traits because it can increase the shelf life of peanut products and is beneficial for human health [88].

QTL mapping for traits related to oil and protein content as well as fatty acid composition in peanut has been reported [41–43,45–47,89]. Furthermore, QTL for unsaturated FA and the oleic acid to linoleic acid ratio [44], as well as QTL for saturated fatty acid composition [90], were also reported. A mapping population of 146 recombinant inbred lines (RILs) generated from a cross of TG26 x GPBD4 was used to discover QTL for protein, oil, oleic, and linoleic acid content, and for the oleic acid to linoleic acid ratio [41]. As the same authors have mentioned, GPBD4 has a desirable combination of early maturity, high yield, high pod growth rate, desirable pod and kernel features, high oil, and protein content, and an optimum oleic/linoleic acid (O/L) ratio, whereas TG26 is a semi-dwarf, erect cultivar with high linoleic acid content. Although the genetic map has low coverage (45 SSR markers on eight linkage groups), the authors reported 17 QTL on four genomic regions, including two major QTL for protein content. Likewise, several QTL were identified using

two genetic maps developed from RIL populations derived from the crosses between Sun Oleic 97R and NC94022 and between Tifrunner and GT-C20 [42]. They found two major QTL for oil content on chromosomes A05 and A08 and 11 major QTL for oleic acid, linoleic acid, and the ratio of oleic acid to linoleic acid on the homeologous chromosomes A09 and B09. Using these mapping populations, 16 major QTL on B04 and A09/B09 were identified for palmitic acid, stearic acid, arachidic acid, gadoleic acid, behenic acid, and lignoceric acid content [90]. One consistent QTL for oil content was mapped on chromosome B03, explaining 14.36% of phenotypic variance [89]. Likewise, two genetic maps were developed using two F2 mapping populations, one for fatty acid composition (FA-population, ICGV 06420 x Sun Oleic 95R) and the other for oil content (OC-population, ICGV 07368 x ICGV 06420) [43], with 1435 and 854 SNP loci spanning 1869 and 3526 cM distances, respectively. In these two maps, 23 major QTL were identified on 11 genomic regions; 2 for oil content and 21 for fatty acid composition variation, explaining up to 41% of PEV.

Another high-resolution genetic map, with 2334 SNP markers and a total length of 2586.37 cM, was constructed using a RIL population developed from the cross between high- and normal oleic cultivars [44]. The authors reported 29 major QTL for oleic and linoleic acid content as well as oleic to linoleic acid ratio, explaining 10 to 57.6% of the phenotypic variance, which were mapped on chromosomes A03 and A09/B09.

More recently, 14 major QTL involving oil content on A05, A06, A08, B06, and B10, explaining up to 27.19% PEV, were discovered from the three mapping populations derived from Xuhual13 and Zhonghua6, Yuhual15 and W1205, and Zhonghua10 and ICG12625 [45–47]. Moreover, major QTL associated with protein stearic acid, behenic acid, and arachidic acid contents were mapped on chromosomes A05, A06, and A08 [47]. The locations of the chromosomes where the aforementioned QTL are located are shown in Figure 3.

We performed a comparative QTL analysis using data from the studies above in order to gain more insight into the genome-wide distribution of kernel-quality QTL and to document the most consistent ones for future use in marker-assisted breeding. The map location of the QTL is presented in Figure 3 and the detailed data of all the 1261 QTL reported in this review, including information on the 413 quality-related QTL are found in Supplementary Table S1. Except for chromosome B01, QTL related to quality traits are mapped on 19 out of the 20 peanut chromosomes. We found that QTL for quality traits are mainly clustered on chromosomes A05, A08, and A09 for the A genome, and B04, B08, and B09 for the B genome. For instance, many QTL for oil and protein content as well as fatty acid compositions (arachidic, arachidonic, behenic, stearic, palmitic, linoleic, and oleic) co-localized on chromosome A05 and were consistent among environments (Figure 3 and Supplementary Table S1). Furthermore, QTL for oleic acid, linoleic acid, and the oleic/linoleic ratio from different studies were found in common genomic regions on chromosomes A05, A08, A09, B04, and B09. In chromosome B09, the common QTL are closely linked to markers ahFAD2B and SNP markers, Marker2575339 or Marker239598. The AhFAD2B QTL, on chromosome B09, explained up to 57% of phenotypic variation of oleic acid or linoleic acid content [42,44]. Similarly, the AhFAD2A and Marker4391589 or Marker4463600 on chromosome A09, are common among studies and explained up to 29% of phenotypic variation [42,44]. Additionally, AhMXZ190701 was discovered to be tightly linked to a major and stable QTL A08 for oil content [42,45]. These consistent markers, AhMXZ190701, ahFAD2B, ahFAD2A, Marker2575339, or Marker2379598 have been used for QTL validation and MAS of quality traits [45,91]. Likewise, several QTL for arachidic, behenic, stearic, palmitic, linoleic, and oleic acid and oil content, mapped in three studies are linked to the marker RN34A10 on chromosome A7 (Figure 3) [42,90]. Furthermore, consistent QTL among traits and environments were also reported. From a QTL mapping study on four environments, among the 110 QTL related to nine quality traits, 36 pleiotropic QTL were associated with two or more traits and showed consistent effects in more than one environment [47].

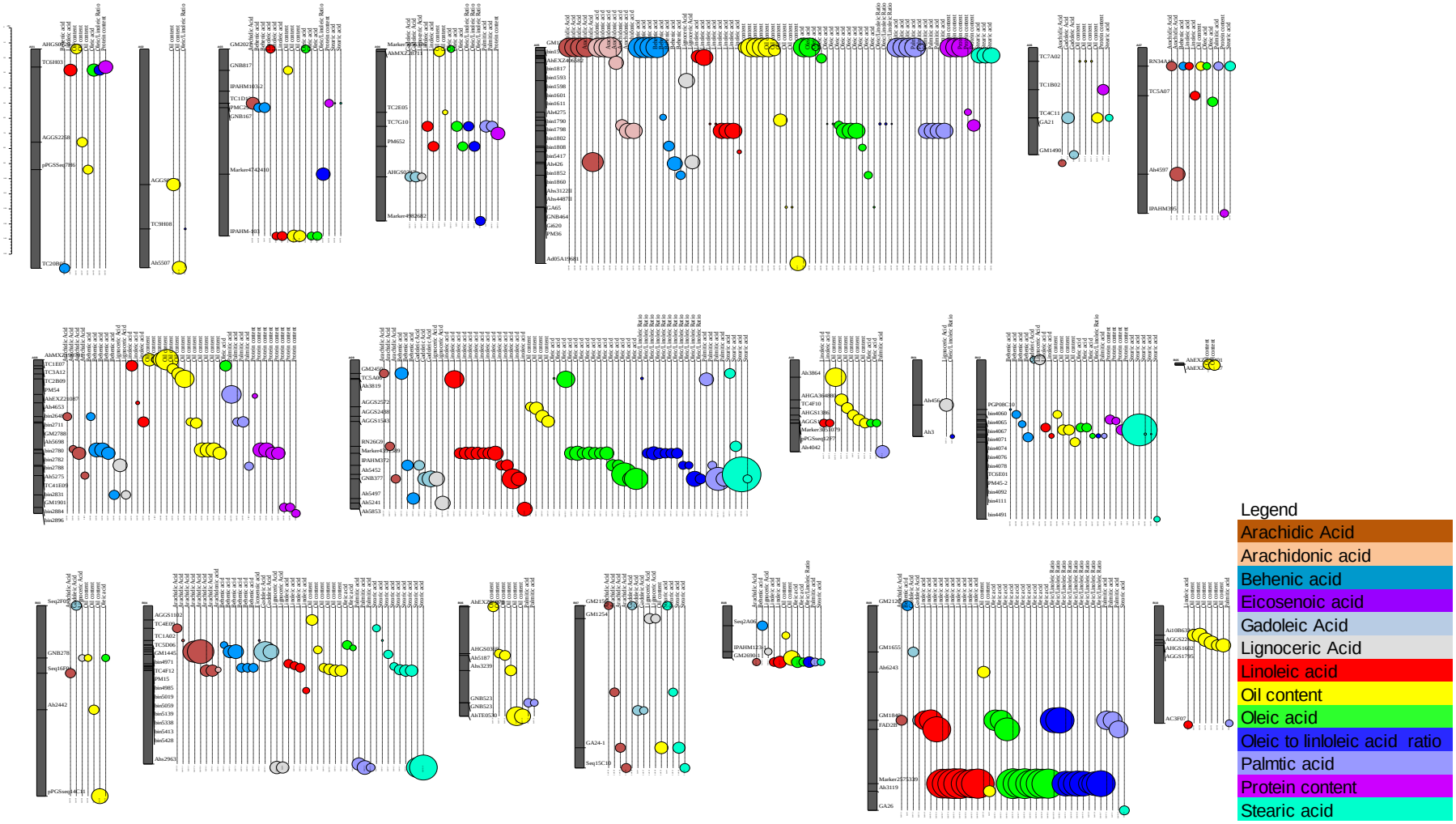


Figure 3. Schematic map of known QTL related to quality traits in peanut. The QTL detected are distributed over the GLs of the A genome, named from A1 to A10, and those of the B genome named from B1 to B10 indicated by gray-colored segments. Locus names related to the QTL are shown on the right of each GL. The QTL detected for each trait are indicated by colors in the legend. The peaks of the QTL are respectively indicated circles. The size of the circles is proportional to the phenotypic variance of the trait indicated by the authors.

The consistent QTL identified for peanut quality traits, thanks to published studies, reviewed here can be used in breeding special-purpose peanut cultivars. However, some QTL need to be validated with fine mapping considering their positions on chromosomes differed in different studies, probably due to the genetic material, large QTL intervals, and statistical imprecisions.

4.2. Mapping QTL for Agro-Morphological Traits

In order to meet the food needs of a growing world population, the main goal of the plant breeding program has been to increase pod yield. In this paper, SSR, SNP, and DArT markers linked to agro-morphological traits utilizing various mapping populations, including F2, BC2F1, BC2F3, BC3F2, BC4F3, and recombinant inbred lines (RILs), have been discussed. Given this, a total of 266 main-effect QTL were mapped for pod- and seed-related traits: 100-seed weight, 100-pod weight, pod weight, pod length, pod width, seed length, seed width, and pod number, using F2 [36,89] and RIL [31–35,92,93] populations. A total of 44 QTL for 100-pod weight were identified, explaining up to 38.15% of the variance on chromosomes A05, A07, A08, B02, B03, B07, and B08, as well as homeologous chromosomes A07 and B07 [30,31,89]. On chromosomes A02, A03, A04, A05, A06, A07, A08, B02, B03, B04, B05, B06, and B08 with A05 and B05 homoeologous loci, 35 QTL were reported with 5.68 to 35. 9% phenotypic variance explained linked with 100-seed weight [31,32,89]. These 05A1430-A05A1601, A05A1344-A05A1562, and A05A1430-A05A1601 major and stable SSR markers increased pod length, pod width, and 100-pod weight by 27.84, 14.12, and 26.82%, respectively [30]. Similarly, major QTL for pod weight and seed weight were reported [35].

Along with traits related to seeds and pods, QTL for flowering, plant height, and fresh seed dormancy were also identified. A total of 30 QTL were reported, ranging in PVE from 1.15 to 21.82%, which underlie the days for 50% flowering and the first days of flowering, using three genetic maps created from TAG 24 x GPBD 4, and TAG 24 x ICGV 86031 [32,94]. For plant height, 71 main-effect QTL [32,92,95], were identified, accounting for up to 26.27% of the phenotypic variance. Several main-effect QTL linked to seed dormancy or fresh seed dormancy in peanut have recently been reported to explain up to 71.21% of the phenotypic variance using F2 and RIL populations [36–40]. For fresh seed germination, QTL were detected on seven homoeologous chromosomes, which are in both A and B genomes with one major stable marker [38]. The co-localization of QTL for the studied traits was reported. Several QTL associated with plant architecture as growth habit and plant height co-localized with those associated with flowering [96]. Moreover, the QTL of yield components, such as 100-pod weight, pod weight, and pod length [30,33], 100-pod weight, 100-seed weight, and pod weight [29,32], were co-localized on chromosomes A05 and A07. Overall, the QTL mapped related to agro-morphological traits include QTL related to plant architecture, flowering, fresh seed or seed dormancy, and yield component traits.

In this review, QTL underlying drought resistance traits were also highlighted. For these drought resistance traits, shoot dry weight, transpiration efficiency, leaf area, transpiration rate, transpiration, and SPAD chlorophyll meter readings (SCMR), 127 QTL were discovered with the phenotypic explained variation ranging from 4.2 to 22.09% [48–50]. From a field experiment using well-watered and water-limited treatments a total of 13 QTL, individually explained 10.4%–20.1% of the phenotypic variance, were significant for the stress tolerance indices (STI): two for total biomass on chromosomes B06 and A05, one for pod weight on chromosome A05, one for seed weight on chromosome A05, two for haulm weight on chromosomes A02 and A05, two for 100 pod weight on chromosomes B02 and B05, and two for 100 seed weight on chromosome A05. In most cases, the STI-related QTL co-localized with the yield component-related trait for which they were calculated [15]. Main-effect QTL reviewed for the important traits are found in Table 2.

Table 2. Main-effect QTL reviewed for the important traits of peanut.

Quality-Related Traits			
Traits Studied	**QTL Identified**	**Phenotypic Variance Explained**	**References**
Oil content	80	0.76–27.19	[18,41–43,46,47,89,90]
Protein content	22	0.76–26.99	[41,47]
Oleic acid	58	0.13–57.56	[18,41–44,47]
Linoleic acid	54	0.17–57.56	[18,41–44,47]
Oleic/linoleic acid ratio	32	1.04–43.41	[18,41,42,44]
Palmitic acid	32	0.3–34.35	[43,47,90]
Arachidic acid	32	0.13–36.93	[18,43,47,90]
Stearic acid	31	0;13–78.6	[18,47,90]
Behenic acid	32	0.76–26.99	[18,43,47,90]
Eicosanoid	1	0.2	[18]
Lignoceric acid	15	2.89–12.61	[43,90]
Gadoleic acid	16	2.55–15.11	[90]
Arachidonic acid	8	0.76–26.99	[47]
Agro-Morphological-Related Traits			
Traits Studied	QTL Identified	Phenotypic Variance Explained	References
Plant height	77	0.01–26.7	[15,32,92,95]
Hundred-pod weight	48	3.33–38.15	[15,30,31,33]
Fresh seed/seed dormancy	54	69.3–74.7	[36–40]
Days to flowering	31	1.15–21.82	[15,32,94]
Pod weight	20	7.7–29.7	[15,35,48,92]
Pod length	S2	1.25–26.46	[15,30,31,33,89]
Pod width	54	5.1–43.63	[15,29–31,33,89]
Seed length	32	3.03–20.8	[15,35,48]
Seed width	33	2.21–23.7	[15,29,31]
Harvest index	15	11.0–18.1	[15,49,50]
Hundred-seed weight	42	5.68–35.9	[15,31,34,63,89,92]
Haulm weight	11	2.9–33.36	[15,48,92]
Pod number	24	3.91–14.2	[15,32,93]
Total biomass	15	4.34–22.39	[15,48,50]
Growth habit	48	4.55–27.14	[15,64,96]
Drought-Tolerance-Related Traits			
Traits Studied	QTL Identified	Phenotypic Variance Explained	References
Shoot dry weight	16	4.2–22.09	[49,50]
Transpiration efficiency	27	4.47–18.12	[48–50]
Leaf area	26	5.0–16.2	[48,50]
Transpiration rate	13	4.3–17.3	[50]
Transpiration	16	4.36–18.17	[48,49]
SPAD chlorophyll meter readings (SCMR)	29	4.00–19.53	[48]
Stress Tolerance Index	13	10.4–20.1	[15]
Pest- and Disease-Related Traits			
Traits Studied	QTL Identified	Phenotypic Variance Explained	References
Nematode	7	1.3–22.18	[17]
Leaf spot	80	1.7–50.9	[51,53–55,62,63,97]
Bacterial wilt	3	0.12–0.22	[58]
Rust	34	7.24–48.7	[52,55,62,63]
Smut	10	7.24–11.4	[56]

4.3. Mapping QTL for Disease Resistance Traits

The most efficient and environmentally friendly way to fight against pests and diseases to control yield losses is to develop disease-resistant cultivars by finding the responsible QTL or genes. For these particular important quantitative traits, 134 QTL were reviewed

using RFLP, SSR, SNP, and DArT markers with F2, BC2F1, BC3F1, BC2F4, and RILs, mapping populations. Of the 134 QTL, 82 have been found to be linked to resistance to late leaf spot, early leaf spot, rust, smut, and bacterial wilt using F2 [55,58], and RILs [51–54,56,97], mapping populations with SSR and SNP markers. Using 103 RIL genotypes derived from a cross between JS17304-7-B and JS1806, 10 QTL underlying smut resistance were identified with a phenotypic variance of up to 11.4% [56]. A total of 12 main-effect QTL linked to rust resistance traits were identified in a RIL population [52,55]. From a RIL mapping population derived from Tamrun 0L07 and Tx964117, eight QTL were identified as linked to leaf spots that explained phenotypic variance ranging from 8 to 20% [97]. Similarly, 36 QTL that linked early and late leaf spots were discovered using a 192 RIL population produced from a hybrid of Florida-7 x GP-NCWS16 cultivars [53,54]. Three minor effects of QTL explaining up to 0.26% PEV for bacterial wilt were discovered using the mapping population derived from the cross of Yueyou 92 and Xinhuixiaoli [58]. Given this review, QTL linked to disease resistance traits have been mapped on all linkage groups except A04 and B07 having resistance QTL from the donor parents on A02/B02, A03, B03, and B05 for late leaf spot, rust, and smut [53–56]. Homoeologous QTL were discovered on A05/B05 for late leaf spot [53,54], and on A02/B02 for rust [55]. QTL on A02, B03, and A05 are common among studies and common for both disease and yield component traits, which is also supported by their strong genetic correlation.

4.4. QTL Mapping Using Interspecific Synthetic Tetraploids

From the reported relevant studies, we found that several interspecific populations have been used for QTL mapping in peanut. Those populations are developed from crosses between synthetic tetraploids and elite varieties and allowed the broadening of the genetic base of cultivated peanut and helped with the mapping of QTL and identifying wild beneficial alleles for economically important traits. Indeed, cultivated peanut has low genetic variation due to its origin in a single hybridization event between two diploid species, followed by chromosomal doubling and crossing barriers with wild diploid species [60,98]. The low genetic variability for traits of importance and polyploidy is a bottleneck to peanut improvement. The primary gene pool of peanut includes mainly tetraploids such as cultivars, advanced breeding lines, and landraces of *A. hypogaea*, as well as *A. monticola* [22,87]. This gene pool is cross-compatible, allowing fertile hybrids to be produced. The secondary gene pool, on the other hand, consists of wild diploid species (2n = 2x = 20) [24,27], which possess desirable alleles for several economically important traits, such as biotic and abiotic resistance. Despite rich diversity with desirable alleles, the use of wild relatives has been limited in breeding programs because of ploidy differences with cultivated peanut [1,27]. However, the ploidy level difference-induced bottleneck has been solved using interspecific synthetic allotetraploids [22,23,60]. The effect of polyploidization and hybridization on various traits in *Arachis* interspecific synthetic tetraploids has been studied [20,62]. Wild chromatin introgression into cultigen from the synthetic tetraploid increase DNA polymorphism helping to map QTL by using the AB-QTL or CSSL analysis [15,17,18,63,98].

To the best of our knowledge, six synthetic tetraploids have been used to date for QTL mapping of traits related to disease resistance, drought resistance, agronomic traits, and oil quality traits utilizing BC2F1, BC3F1, BC2F3, BC2F4, BC3F2, BC4F3, and BC3F6 mapping populations (Figure 1). A total of seven QTL were identified from BC3F1 in the population developed from cultivated "Florunner" and synthetic TxAG-6 for root-knot nematode resistance using RFLP markers [17]. Using the BC3F6 generations of the mapping population developed from the aforementioned parents genotyped using SSR markers, 29 QTL associated with oil content, six fatty acid traits, and the oleic to linoleic acid ratio were detected on 20 genomic regions [18]. Of the 20, two are major and stable and linked to oil content and the oleic to linoleic acid ratio with phenotypic variance explained at 17–21 and 13–31% PEV, respectively. Another genetic mapping based on SSR markers was performed using a cross between the cultivated variety "Fleur 11" and the synthetic tetraploid AiAd derived from *A. ipaensis* x *A. duranensis*, the two diploid ancestors of the

cultivated [84]. The advanced backcross populations (BC2F3 and B3F2) from this cross were utilized for QTL mapping of flowering date, pod weight, pod number, seed number, pod size, seed size, pod maturity, and biomass under well-watered and water-limited treatments, yielding a total of 95 QTL [15]. About half of the QTL positive effects were associated with alleles of the wild parent, highlighting that peanut wild relatives represent a reservoir of useful alleles for peanut breeding. In addition, by using BC4F3 (CSSL) lines from the same cross, the authors also found 42 QTL mapped for plant growth habits, the height of the main stem, plant spread, and flower color [64]. Likewise, using the same recurrent cultivated parent, Fleur 11 with a different synthetic parent ISATR52B, 38 QTL were identified underlying the flowering date, plant architecture, yield-related, pod, and seed morphology traits on 16 chromosomes [61]. They found that almost 50% of the positive QTL effects were associated with alleles of ISATR52B.

Furthermore, 28 QTL explaining 6.7–50.9% PEV linked to 100-seed weight, oleic to linoleic acid ratio, and late leaf spot and rust resistance were identified from the two interspecific populations derived from ICGV 8764 x ISATGR 265-5A and ICGV 91114 and ISATGR 1212 genotyped with SSR and DArT markers spanning 1415.7–1500.8 cM map length [63]. Seven QTL for late leaf spot and rust resistance and one for the oleic to linoleic acid ratio, with phenotypic variance explained up to 9.7 and 14.8 on chromosomes A01, A07, and A08, were found in the population derived from ICGV 91114 X ISATGR 1212. In the population derived from ICGV 8764 XISATGR 265-5A, three QTL were found for each trait for the oleic to linoleic acid ratio and 100-seed weight, up to 47% phenotypic variance, and 14 QTL were found for late leaf spot and rust resistance, up to 50.9% phenotypic variance. Of the 28 QTL, three contributed favorable alleles from wild genomic segments. In an effort to enhance foliar disease resistance, a cross between ICGS 76 and synthetic amphidiploid ISATGR 278-18 has been recently revealed [62]. In this population, 14 and 10 QTL associated with late leaf and rust resistance, up to 38.58% PEV, respectively, were identified.

Overall, synthetic tetraploids are used for alien chromatin introgression into cultivated peanut by resolving polyploidy differences and easing genetic and meiotic analyses and QTL detection for numerous economically important traits despite constraints, such as hybrid fertility and linkage drag.

4.5. Clustering of QTL of Quality and Agronomic Traits

Of the 1261 main-effect QTL reviewed here, we identified relevant QTL related to quality traits that clustered with QTL of agronomic traits. For instance, the QTL of oil content associated with the SSR marker IPAHM103 on chromosome A3 with a PEV of 7.1–10.2% for oil content [41] is the same identified for rust resistance with a PEV of 6.9–55.2% [51] and late leaf spot resistance. Likewise, the SSR marker PM36 mapped for oil content in several studies [18,41] was identified in the QTL regions of pod and seed weight and shelling percentage in chromosome 5 in various studies [15,30,31,33].

As far as the oil quality traits are concerned, QTL (TC6H03–TC11A04, TC5A07–IPAHM395, and TC3A12–PM433) were common for both oleic and linoleic acid, which is also supported by their strong negative correlation [41]. Consistent QTL, ahFAD2A and ahFAD2B, IPAHM372-ahFAD2A, GM1840-ahFAD2B, and GNB377-ahFAD2A linked oleic acid, linoleic acid, and O/L ratio [42]. In addition to the SSR markers linked for oil quality traits, SNP markers Marker2575339 and Marker2379598 in B09 and Marker4391589 and Marker4463600 in A09 were associated with oleic acid, linoleic acid, and the ratio of oleic acid to linoleic acid (O/L) [44]. A major and stable QTL on A05, flanked by the markers bin1572 and bin1573 on 0–0.5 cM was detected and showed a negative additive effect on oil, palmitic, stearic, arachidic, and behenic acid content and positive additive effects on protein, oleic, arachidonic acid [47]. These stable oil-related QTL on A05 is quite common to a stable and major genomic region on A05 that has been reported for pod and seed-related traits in several studies [30,33]. The co-localized interval on A05 was located on 1.3 cM (99.50–99.78 Mb) by the flanking markers Ad05A20262 and AHGA160418 and harbored the major QTL for pod length, pod width, and 100-pod weight with 17.97–43.62% of phenotypic variations [33]. For these traits, from another QTL mapping,

three more major QTL co-located in about 2.47 Mb genomic region of the A05 with 13.75 to 26.68% PVE by the flanking markers A05A1430-A05A1601 [30]. Moreover, three major QTL common for pod length and seed length on A05 with up to 26.11% PEV were identified [29]. The cluster of many major QTL detected on A05 in different studies for oil content and seed or pod-related traits suggests it may harbor important genes controlling these traits, which can be used, simultaneously in marker-assisted breeding. This clustering also suggests linked QTL of these distinct traits or QTL with pleiotropic effects. Thus, breeding for an agronomic trait may indirectly, positively, or negatively affect a quality-related trait.

4.6. Statistical Methods and Limitations of QTL Mapping in Peanut

QTL are, by definition, merely significant statistical associations between genotypic values and phenotypic variability among the segregating progeny. Statistical methods for family-based mapping include (i) single-marker analysis (SMA) used to identify QTL according to the difference between the average phenotypes of different genotype groups without linkage map; (ii) interval mapping (IM) based on maximum-likelihood parameter estimation and regression, which efficiently estimates the effect and position of a QTL within two flanking markers; (iii) composite interval mapping (CIM), to overcome such limitations of the IM method; (iv) inclusive composite interval mapping (ICIM) and (v) multiple interval mapping (MIM), an extension of IM, that tends to be more powerful and precise than CIM in identifying QTL and allows the estimation of multiple QTL with epistasis. A large number of software implementing the above methods are used in peanut, including R/qtl, QTL Cartographer, and ICIM Mapping [15,28,30,33,34].

From most studies reviewed here, QTL mapping technique accuracy depends on several factors, including the statistical method's capacity to locate and estimate the genetic effect of the QTL, the type and size of the mapping population, the genetic and heritability of the trait, the number and contribution of each QTL to the total variance, their interactions and their distribution over the genome. In addition, the ploidy coupled with mixed meiotic behavior is not yet considered in QTL detection and may affect the accuracy of QTL mapping in peanut. Along with these accuracy factors, QTL analysis has limitations like other techniques. Some of these limitations include the inability to detect all loci, the number of QTL detected, their precise position, and their effects are subject to statistical error. Major QTL are often missed and epistatic effects and QTL environmental interactions are found in some cases. QTL mapping is often time-consuming and labor-intensive, requires in-depth knowledge about the function and genomics of the trait of interest, and incurs high costs for genotyping and phenotyping. The large size of QTL and the low resolution of mapping greater than 10 cM in size are some of QTL mapping's population specificities. In many cases, more experiments are needed to confirm the results of QTL mapping. However, by using consistent QTL that have been mapped, it is expected that the next-generation crop varieties could be developed with enhanced quality traits, better yield, and disease resistance.

5. Toward More Effective Use of Marker Assisted Selection (MAS) and QTL in Peanut

One of the current challenges in peanut is to use QTL of interest to accelerate genetic improvement. Considering the constraints of phenotypic selection—labor-intensive, costly, and time-consuming—MAS has emerged as a potential tool to achieve rapid results with the help of molecular markers and QTL of interest in plant breeding [99,100]. There are different molecular approaches used under the umbrella of MAS, such as marker-assisted backcrossing (MABC), gene pyramiding, MARS, and GS [101,102]. Some of the innovative applications of MAS including combined MAS, marker-directed phenotyping, inbred or pure-line enhancement, single large-scale MAS, breeding by design, and Mapping As You Go (MAYG) have been published previously [103,104]. Here, some schemes of MAS have been highlighted and may be useful for peanut breeding.

5.1. Marker-Assisted Backcross Selection (MABC)

MABC is a technique that can be used to incorporate one or more QTL from a donor parent (DP) to a recurrent parent (RP), which is a superior variety but lacking the target trait. Four to six generations of backcrossing are required to introduce the QTL into an elite cultivar and recover the recurrent parent [103,105]. Foreground selection, recombinant selection, and background selection are the three basic steps in marker-assisted backcrossing. In foreground selection, the desired plant is chosen using markers linked to the target QTL. Random markers across the entire genome can be used to screen the recurrent parent genome in the background selection context [103–105]. Recombinant selection is a kind of foreground selection that aims to remove the DP genome flanking the target QTL to avoid the linkage drag brought on by the close linkage of some undesirable traits with the target trait from the DP. MABC has been used in peanut breeding. Two BC3F1 lines, TMG-29 and TMG-46, have shown enhanced resistance over the highly susceptible TMV 2 peanut variety of late leaf spot (LLS) and rust-resistant genotypes using resistant donor "GPBD 4" [106]. According to [107], the backcrossing lines that were created by introducing the two mutant alleles, ahFAD2A and ahFAD2B, into the high oil content breeding line ICGV06100 showed a 97% increase in oleic acid content in comparison to the recurrent parent. A total of 22 BC3F4 and 30 BC2F4 introgression lines for rust and late leaf spot resistance, as well as 46 BC3F4 and 41 BC2F4 for high oleic acid, were created by crossing the donor rust resistance parent, GPBD 4, with three susceptible peanut cultivars (GJG 9, GG 20, and GJGHPS 1) in order to develop rust-resistance and late leaf-spot-resistance and a high oleic acid content genotypes [108]. Recently, the high-oleic-acid BC4F6 line "YH61" was created after four backcrossing of "huayu22" with the donor "KN176" with a high-oleic-acid content [109]. Furthermore, chromosome segment substitution lines [64,65], near-isogenic lines [67], and AB-populations [15,17,18,61–63] use, in such cases, MABC approaches and all facilitate genetic analysis, QTL introgression, and variety development in a simultaneous manner. MABC strategy can be more effectively used for introgression of major-effect QTL controlling different economically important traits to develop improved varieties.

5.2. Marker-Assisted Recurrent Selection (MARS)

In marker-assisted recurrent selection (MARS), plant genotypes are selected with the help of molecular markers that have been linked to the genes or QTL of interest. Once markers that are tightly linked to QTL of interest have been identified breeders use specific DNA marker alleles as a diagnostic tool to identify plants carrying the QTL [86], the chosen individuals that have QTL of interest are then subjected to controlled pollination to create lines that have the best possible complement of QTL from both parents [101,110]. This scheme is less used and may be useful for developing breeding material displaying QTL for targeted breeding traits. MARS is more suitable for introgression of minor-effect QTL controlling different important traits.

5.3. Marker-Assisted QTL Pyramiding (MAQP)

Pyramiding is the simultaneous integration of several QTL from multiple parents into a single genotype to create superior lines and varieties [101,102,111]. Marker-assisted QTL pyramiding can speed up the process by lowering the number of generations that the researchers must evaluate to ensure that they have the desired QTL combination [112]. The QTL pyramiding method consists of two fundamental steps: the QTL fixation step, which aims to fix the target QTL into a homozygous state, and the pyramiding step, which aims to accumulate all the target QTL into a single genotype known as the root genotype [113]. This breeding technique was used in peanut to develop nematode resistance and high oleic gene-containing genotypes [114]. It is expected that the next-generation crop varieties could be developed with enhanced quality traits, disease resistance, and better yield by using MAQP.

5.4. Genomic Selection (GS) by Using Know QTL

GS is a promising method for genetic improvement of complex traits that are regulated by many QTL, each of which has a small or main effect [72]. In addition to promising to address complex traits, the GS strategy offers the benefit of shortening the selection cycle and eliminating lengthy phenotyping by favoring superior lines based on the prediction of the genomic-estimated breeding values (GEBV). When phenotype data and information on markers known to be associated with known QTL were combined to calculate estimated breeding values (EBVs), the gains from selection in plant breeding experiments increase significantly [102]. In the same way, the targeted QTL were accumulated at a much higher frequency when known QTL were included in the GS model as compared to when the standard ridge regression was applied. Several factors, including the size of the training population and its constitution/structure, precision, and quality of phenotyping, marker density, and trait heritability, have an effect on the prediction accuracy of GS [102]. This approach has not yet been applied in peanut and could be helpful for peanut breeding by using both cultivated and wild relatives and known QTL in prediction models.

5.5. Combined MAS and Marker-Directed Phenotyping for Quality Traits

In comparison to MAS or phenotypic screening alone, MAS combined with marker-directed phenotyping increases genetic gain and may help identify undiscovered QTL [115]. This combined selection aids in the selection of traits when phenotyping is more expensive than genotyping. In most cases, there is a low level of recombination between QTL and marker, which means we cannot entirely rely on markers for selecting desirable phenotype traits. However, it will help reduce the number of plants to be evaluated, which reduces the cost of phenotyping. One of the successful examples to explain this scheme is the rice primary QTL sub 1, which controls submergence tolerance [116]. This scheme may be useful, mainly for quality traits in peanut such as fatty and amino-acid acid composition where phenotypic screening is costlier than marker genotyping.

6. Use of Mapped QTL in Peanut Genetic and Breeding

For the over 1261 QTL reviewed here, to the best of our knowledge, only less than 10 have been used in peanut breeding programs. The current challenge is to use validated mapped QTL for peanut breeding for the fast-track development of improved varieties. The common QTL found between different studies and different genetic materials, and those with high positive effects, as reported above, can be mobilized in the breeding program by making sure to identify the beneficial QTL allele and combining the ability of the parents. In addition, using mapped QTL for breeding requires fine mapping, QTL and markers validation, and marker-assisted selection. QTL validation often needs cross-verification of QTL in different populations or/and different environments and fine mapping. Marker verification needs testing of molecular markers in germplasm and identifying polymorphic markers. Polymorphic markers around the validated QTL could be used for an indirect selection to strengthen conventional breeding. We expect that the QTL, once validated, are deployed in molecular breeding programs aimed at enhancing targeted traits in peanut through MAS, genomic selection (GS), or holistic and innovative schemes.

7. Use of Lines with Beneficial QTL Alleles for Fast QTL Introgression and Variety Development

QTL analysis from the above studies identified QTL alleles with favorable effects on peanut breeding. Likewise, AB-QTL analysis identified several introgression lines with good agronomic, oil quality, and disease-resistance traits. In addition, AB-QTL and CSSL lines, in common cases, are designed to map and facilitate QTL introgression from unadapted germplasms such as landraces and wild species into elite lines [117]. There are several advantages such as the simplicity of mapping the population in phenotypes to the recurrent parent and reducing, in the process, deleterious alleles from the donor parent, the possibility of epistasis, and linkage drag. After QTL mapping, only one or

a few generations are needed for identifying QTL-NILs [118]. Several AB-QTL populations [15,17,18,63], CSSL [64,65], and NIL [67] have been developed in peanut. Beneficial QTL alleles, carried by AB-QTL lines or CSSL have been identified and, could be utilized through introgression into the genetic backgrounds of cultivars used by producers. In some cases, AB-QTL, CSSL, or NIL outperformed cultivated varieties and meet market needs. Thus, they could be directly promoted as a new variety. For instance, recently in Senegal, six new varieties, Rafet car, Tosset, Komkom, Jambar, Yakaar, and Raw Gadu with high yield profiles were homologated from CSSL [64], 12CS_031, 12CS_069, 12CS_120, 12CS_068, 12CS_037, 12CS_028, respectively, derived from the cross between "Fleur11" and the wild synthetic tetraploid AiAd (*A. duranensis and A. ipaensis*). The next decade will see heavy use of these kinds of lines for the development of new varieties.

8. Conclusions and Perspectives

The chromosomal or genomic regions known as QTL are responsible for variation in a quantitative phenotype. Finding genomic regions with QTL, estimating the effect of the QTL on the quantitative trait, determining how much of the trait's variation is due to a specific region, and discovering the gene action linked to the QTL are the objectives of QTL mapping. In this paper, we have reviewed 1261 QTL that govern economically important traits for peanut breeding that have recently been mapped through diverse sources of mapping populations including F2, recombinant inbred lines, advanced backcross populations, chromosome segment substitution lines, and nested association mapping (NAM) or multiparent advanced generation inter-crossing (MAGIC) populations, as well as a variety of molecular markers. Protein content, oil content, fatty acid composition, yield, yield component, drought resistance, and pest and disease resistance are considered significant and important quantitative traits. Various limitations of QTL mapping have been discussed, and the solutions proposed to overcome them are the constant development of molecular platforms, new genetic materials such as introgression lines that help in mapping the small effects, and sophisticated bioinformatics that can handle polyploidy issues and mixed meiotic behavior, false-positive results or statistical errors. Introgression of validated mapped QTL alleles, fruitfully associated with preferred traits, into the genetic background of the elite varieties is a current challenge for peanut breeding. Integration of high throughput phenotyping and new-generation phenomics tools with MAS could greatly accelerate progress in peanut genetic improvement. The present review discusses the current status and future scope of using mapped QTL for breeding purposes in peanut, which will cause not only an increase in the rate of developing climate-resilient superior cultivars but also help in providing vegetable oil and proteins to the growing human population worldwide.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes14061176/s1, Table S1: Detailed data of QTL related to quality traits in peanut.

Author Contributions: Data collection and writing, F.C.K.; project administration, funding acquisition and supervision, H.B.N. and J.M.B.; Writing—Review and Editing D.V.M.A., M.K.G., W.G.A., H.-A.T., A.S. and M.S.; Software J.-F.R.; formal analysis and writing, D.F.; Conceptualization, Design, formal analysis, writing and editing J.R.N. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the GENES intra-Africa Academic Mobility scheme of the European Union (EU-GENES: EACEA/2917/2552).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available through the supplementary materials. No new data were created.

Acknowledgments: The authors thank the GENES program of the intra-Africa Academic Mobility scheme of the European Union (EU-GENES: EACEA/2917/2552,) for providing a PhD scholarship.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bertioli, D.; Seijo, G.; Freitas, F.; Valls, J.; Leal-Bertioli, S.; Moretzsohn, M. An overview of peanut and its wild relatives. *Plant Genet. Resour.* **2011**, *9*, 134–149. [CrossRef]
2. Nguepjop, J.R.; Tossim, H.-A.; Bell, J.M.; Rami, J.-F.; Sharma, S.; Courtois, B.; Mallikarjuna, N.; Sane, D.; Fonceka, D. Evidence of Genomic Exchanges between Homeologous Chromosomes in a Cross of Peanut with Newly Synthetized Allotetraploid Hybrids. *Front. Plant Sci.* **2016**, *7*, 1635. [CrossRef]
3. Bertioli, D.J.; Jenkins, J.; Clevenger, J.; Dudchenko, O.; Gao, D.; Seijo, G.; Leal-Bertioli, S.C.M.; Ren, L.; Farmer, A.D.; Pandey, M.K.; et al. The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nat. Genet.* **2019**, *51*, 877–884. [CrossRef]
4. Patil, A.; Salunke, M. Breeding for Quality Traits in Groundnut. *Pop. Kheti* **2014**, *2*, 173–178.
5. FAOSTAT. Available online: https://www.fao.org/faostat/en/#home (accessed on 12 May 2022).
6. Statista. Available online: https://www.statista.com/statistics/263933/production-of-vegetable-oils-worldwide-since-2000/ (accessed on 29 January 2023).
7. Desmae, H.; Janila, P.; Okori, P.; Pandey, M.K.; Monyo, E.; Mponda, O.; Okello, D.; Sako, D.; Echeckwu, C.; Oteng-Frimpong, R.; et al. Genetics, genomics and breeding of groundnut (*Arachis hypogaea* L.). *Plant Breed.* **2019**, *138*, 425–444. [CrossRef] [PubMed]
8. Nawade, B.; Mishra, G.P.; Radhakrishnan, T.; Dodia, S.M.; Ahmad, S.; Kumar, A.; Kumar, A.; Kundu, R. High oleic peanut breeding: Achievements, perspectives, and prospects. *Trends Food Sci. Technol.* **2018**, *78*, 107–119. [CrossRef]
9. De Godoy, I.J.; Dos Santos, J.F.; Michelotto, M.D.; de Moraes, A.R.A.; Bolonhezi, D.; de Freitas, R.S.; de Carvalho, C.R.L.; Finoto, E.L.; Martins, A.L.M. IAC OL 5-New high oleic runner peanut cultivar. *Crop Breed. Appl. Biotechnol.* **2017**, *17*, 295–298. [CrossRef]
10. Chamberlin, K.; Melouk, H.; Madden, R.; Dillwith, J.W.; Bannore, Y.; El Rassi, Z.; Payton, M. Determining the Oleic/linoleic Acid Ratio in a Single Peanut Seed: A Comparison of Two Methods. *Peanut Sci.* **2011**, *38*, 78–84. [CrossRef]
11. Bertioli, D.J.; Cannon, S.B.; Froenicke, L.; Huang, G.; Farmer, A.D.; Cannon, E.K.S.; Liu, X.; Gao, D.; Clevenger, J.; Dash, S.; et al. The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. *Nat. Genet.* **2016**, *48*, 438–446. [CrossRef]
12. Jung, S.; Tate, P.L.; Horn, R.; Kochert, G.; Moore, K.; Abbott, A.G. The Phylogenetic Relationship of Possible Progenitors of the Cultivated Peanut. *J. Hered.* **2003**, *94*, 334–340. [CrossRef]
13. Simpson, C.E. Use of Wild *Arachis* Species/Introgression of Genes into *A. hypogaea* L. *Peanut Sci.* **2001**, *28*, 114–116. [CrossRef]
14. Rami, J.-F.; Leal-Bertioli, S.C.M.; Foncéka, D.; Moretzsohn, M.C.; Bertioli, D.J. Groundnut. In *Alien Gene Transfer in Crop Plants*; Pratap, A., Kumar, J., Eds.; Springer New York: New York, NY, USA, 2014; Volume 2, pp. 253–279. [CrossRef]
15. Fonceka, D.; Tossim, H.-A.; Rivallan, R.; Vignes, H.; Faye, I.; Ndoye, O.; Moretzsohn, M.C.; Bertioli, D.J.; Glaszmann, J.-C.; Courtois, B. Fostered and left behind alleles in peanut: Interspecific QTL mapping reveals footprints of domestication and useful natural variation for breediNg. *BMC Plant Biol.* **2012**, *12*, 26. [CrossRef] [PubMed]
16. Kumari, V.; Gowda, M.V.C.; Tasiwal, V.; Pandey, M.K.; Bhat, R.S.; Mallikarjuna, N.; Upadhyaya, H.D.; Varshney, R.K. Diversification of primary gene pool through introgression of resistance to foliar diseases from synthetic amphidiploids to cultivated groundnut (*Arachis hypogaea* L.). *Crop J.* **2014**, *2*, 110–119. [CrossRef]
17. Burow, M.D.; Starr, J.L.; Park, C.-H.; Simpson, C.E.; Paterson, A.H. Introgression of homeologous quantitative trait loci (QTLs) for resistance to the root-knot nematode [Meloidogyne arenaria (Neal) Chitwood] in an advanced backcross-QTL population of peanut (*Arachis hypogaea* L.). *Mol. Breed.* **2014**, *34*, 393–406. [CrossRef]
18. Wilson, J.N.; Chopra, R.; Baring, M.R.; Selvaraj, M.G.; Simpson, C.E.; Chagoya, J.; Burow, M.D. Advanced backcross quantitative trait loci (QTL) analysis of oil concentration and oil quality traits in peanut (*Arachis hypogaea* L.). *Trop. Plant Biol.* **2017**, *10*, 1–17. [CrossRef]
19. Janila, P.; Nigam, S.N.; Pandey, M.K.; Nagesh, P.; Varshney, R.K. Groundnut improvement: Use of genetic and genomic tools. *Front. Plant Sci.* **2013**, *4*, 23. [CrossRef]
20. Carvalho, P.V.; Márcio, D.C.M.; Brasileiro, A.C.M.; GuimarÃ£es, P.M.; Tânia, D.S.A.C.; Joseane, P.D.S.; Gimenes, M.A. Evidences that polyploidization and hybridization affected resveratrol content in Arachis interspecific hybrids. *J. Plant Breed. Crop Sci.* **2019**, *11*, 265–270. [CrossRef]
21. Simpson, C.E.; Nelson, S.C.; Starr, J.L.; Woodard, K.E.; Smith, O.D. Registration of TxAG-6 and TxAG-7 Peanut Germplasm Lines. *Crop Sci.* **1993**, *33*, 1418. [CrossRef]
22. Fávero, A.P.; Simpson, C.E.; Valls, J.F.M.; Vello, N.A. Study of the Evolution of Cultivated Peanut through Crossability Studies among *Arachis ipaënsis*, *A. duranensis*, and *A. hypogaea*. *Crop Sci.* **2006**, *46*, 1546–1552. [CrossRef]
23. Mallikarjuna, N.; Senapathy, S.; Hoisington, D. Development of new sources of tetraploid Arachis to broaden the genetic base of cultivated groundnut (*Arachis hypogaea* L.). *Genet. Resour. Crop Evol.* **2010**, *58*, 889–907. [CrossRef]
24. Stalker, H.T. Utilizing Wild Species for Peanut Improvement. *Crop Sci.* **2017**, *57*, 1102–1120. [CrossRef]
25. Motagi, B.N.; Bhat, R.S.; Pujer, S.; Nayak, S.N.; Pasupaleti, J.; Pandey, M.K.; Varshney, R.K.; Bera, S.K.; Pal, K.K.; Mondal, S.; et al. Genetic Enhancement of Groundnut: Current Status and Future Prospects. In *Accelerated Plant Breeding, Volume 4: Oil Crops*; Gosal, S.S., Wani, S.H., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 63–110. ISBN 978-3-030-81107-5.

26. Xu, Y.; Li, P.; Yang, Z.; Xu, C. Genetic mapping of quantitative trait loci in crops. *Crop J.* **2017**, *5*, 175–184. [CrossRef]
27. Pandey, M.K.; Guo, B.; Holbrook, C.C.; Janila, P.; Zhang, X.; Bertioli, D.J.; Isobe, S.; Liang, X.; Varshney, R.K. Molecular markers, genetic maps and QTLs for molecular breeding in peanut. In *Genetics, Genomics and Breeding of Peanuts*; CRC Press: Boca Raton, FL, USA, 2014; pp. 61–113.
28. Vishwakarma, M.K.; Nayak, S.N.; Guo, B.; Wan, L.; Liao, B.; Varshney, R.K.; Pandey, M.K. Classical and molecular approaches for mapping of genes and quantitative trait loci in peanut. In *The Peanut Genome*; Varshney, R.K., Pandey, M.K., Puppala, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 93–116. [CrossRef]
29. Chen, W.; Jiao, Y.; Cheng, L.; Huang, L.; Liao, B.; Tang, M.; Ren, X.; Zhou, X.; Chen, Y.; Jiang, H. Quantitative trait locus analysis for pod- and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.). *BMC Genet.* **2016**, *17*, 25. [CrossRef] [PubMed]
30. Luo, H.; Ren, X.; Li, Z.; Xu, Z.; Li, X.; Huang, L.; Zhou, X.; Chen, Y.; Chen, W.; Lei, Y.; et al. Co-localization of major quantitative trait loci for pod size and weight to a 3.7 cM interval on chromosome A05 in cultivated peanut (*Arachis hypogaea* L.). *BMC Genom.* **2017**, *18*, 58. [CrossRef] [PubMed]
31. Chen, Y.; Ren, X.; Zheng, Y.; Zhou, X.; Huang, L.; Yan, L.; Jiao, Y.; Chen, W.; Huang, S.; Wan, L.; et al. Genetic mapping of yield traits using RIL population derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.). *Mol. Breed.* **2017**, *37*, 17. [CrossRef] [PubMed]
32. Khedikar, Y.; Pandey, M.K.; Sujay, V.; Singh, S.; Nayak, S.N.; Klein-Gebbinck, H.W.; Sarvamangala, C.; Mukri, G.; Garg, V.; Upadhyaya, H.D.; et al. Identification of main effect and epistatic quantitative trait loci for morphological and yield-related traits in peanut (*Arachis hypogaea* L.). *Mol. Breed.* **2017**, *38*, 7. [CrossRef]
33. Luo, H.; Guo, J.; Ren, X.; Chen, W.; Huang, L.; Zhou, X.; Chen, Y.; Liu, N.; Xiong, F.; Lei, Y.; et al. Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **2018**, *131*, 267–282. [CrossRef]
34. Liang, Y.; Baring, M.R.; Septiningsih, E.M. Mapping of Quantitative Trait Loci for Yield and Grade Related Traits in Peanut (*Arachis hypogaea* L.) Using High-Resolution SNP Markers. *Plant Breed. Biotechnol.* **2018**, *6*, 454–462. [CrossRef]
35. Chavarro, C.; Chu, Y.; Holbrook, C.; Isleib, T.; Bertioli, D.; Hovav, R.; Butts, C.; Lamb, M.; Sorensen, R.; Jackson, S.A.; et al. Pod and Seed Trait QTL Identification To Assist Breeding for Peanut Market Preferences. *G3 Genes Genomes Genet.* **2020**, *10*, 2297–2315. [CrossRef]
36. Vishwakarma, M.K.; Pandey, M.K.; Shasidhar, Y.; Manohar, S.S.; Nagesh, P.; Janila, P.; Varshney, R.K. Identification of two major quantitative trait locus for fresh seed dormancy using the diversity arrays technology and diversity arrays technology-seq based genetic map in Spanish-type peanuts. *Plant Breed.* **2016**, *135*, 367–375. [CrossRef]
37. Kumar, R.; Janila, P.; Vishwakarma, M.K.; Khan, A.W.; Manohar, S.S.; Gangurde, S.S.; Variath, M.T.; Shasidhar, Y.; Pandey, M.K.; Varshney, R.K. Whole-genome resequencing-based QTL -seq identified candidate genes and molecular markers for fresh seed dormancy in groundnut. *Plant Biotechnol. J.* **2020**, *18*, 992–1003. [CrossRef]
38. Zhang, M.; Zeng, Q.; Liu, H.; Qi, F.; Sun, Z.; Miao, L.; Li, X.; Li, C.; Liu, D.; Guo, J.; et al. Identification of a stable major QTL for fresh-seed germination on chromosome Arahy.04 in cultivated peanut (*Arachis hypogaea* L.). *Crop J.* **2022**, *10*, 1767–1773. [CrossRef]
39. Bomireddy, D.; Gangurde, S.S.; Variath, M.T.; Janila, P.; Manohar, S.S.; Sharma, V.; Parmar, S.; Deshmukh, D.; Reddisekhar, M.; Reddy, D.M.; et al. Discovery of Major Quantitative Trait Loci and Candidate Genes for Fresh Seed Dormancy in Groundnut. *Agronomy* **2022**, *12*, 404. [CrossRef]
40. Wang, M.L.; Wang, H.; Zhao, C.; Tonnis, B.; Tallury, S.; Wang, X.; Clevenger, J.; Guo, B. Identification of QTLs for Seed Dormancy in Cultivated Peanut Using a Recombinant Inbred Line Mapping Population. *Plant Mol. Biol. Rep.* **2022**, *40*, 208–217. [CrossRef]
41. Sarvamangala, C.; Gowda, M.V.C.; Varshney, R.K. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (*Arachis hypogaea* L.). *Field Crops Res.* **2011**, *122*, 49–59. [CrossRef]
42. Pandey, M.K.; Wang, M.L.; Qiao, L.; Feng, S.; Khera, P.; Wang, H.; Tonnis, B.; Barkley, N.A.; Wang, J.; Holbrook, C.C.; et al. Identification of QTLs associated with oil content and mapping FAD2 genes and their relative contribution to oil quality in peanut (*Arachis hypogaea* L.). *BMC Genet.* **2014**, *15*, 133. [CrossRef]
43. Shasidhar, Y.; Vishwakarma, M.K.; Pandey, M.K.; Janila, P.; Variath, M.T.; Manohar, S.S.; Nigam, S.N.; Guo, B.; Varshney, R.K. Molecular Mapping of Oil Content and Fatty Acids Using Dense Genetic Maps in Groundnut (*Arachis hypogaea* L.). *Front. Plant Sci.* **2017**, *8*, 794. [CrossRef] [PubMed]
44. Hu, X.H.; Zhang, S.Z.; Miao, H.R.; Cui, F.G.; Shen, Y.; Yang, W.Q.; Xu, T.T.; Chen, N.; Chi, X.Y.; Zhang, Z.M.; et al. High-Density Genetic Map Construction and Identification of QTLs Controlling Oleic and Linoleic Acid in Peanut using SLAF-seq and SSRs. *Sci. Rep.* **2018**, *8*, 5479. [CrossRef] [PubMed]
45. Liu, N.; Guo, J.; Zhou, X.; Wu, B.; Huang, L.; Luo, H.; Chen, Y.; Chen, W.; Lei, Y.; Huang, Y.; et al. High-resolution mapping of a major and consensus quantitative trait locus for oil content to a ~0.8-Mb region on chromosome A08 in peanut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **2020**, *133*, 37–49. [CrossRef]
46. Guo, J.; Liu, N.; Li, W.; Wu, B.; Chen, H.; Huang, L.; Chen, W.; Luo, H.; Zhou, X.; Jiang, H. Identification of two major loci and linked marker for oil content in peanut (*Arachis hypogaea* L.). *Euphytica* **2021**, *217*, 29. [CrossRef]
47. Sun, Z.; Qi, F.; Liu, H.; Qin, L.; Xu, J.; Shi, L.; Zhang, Z.; Miao, L.; Huang, B.; Dong, W.; et al. QTL Mapping of Quality Related Traits in Peanut Using Whole-Genome Resequencing. *Res. Sq.* **2021**, preprint. [CrossRef]

48. Ravi, K.; Vadez, V.; Isobe, S.; Mir, R.R.; Guo, Y.; Nigam, S.N.; Gowda, M.V.C.; Radhakrishnan, T.; Bertioli, D.J.; Knapp, S.J.; et al. Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **2011**, *122*, 1119–1132. [CrossRef] [PubMed]
49. Gautami, B.; Pandey, M.K.; Vadez, V.; Nigam, S.N.; Ratnakumar, P.; Krishnamurthy, L.; Radhakrishnan, T.; Gowda, M.V.C.; Narasu, M.L.; Hoisington, D.A.; et al. Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Mol. Breed.* **2012**, *30*, 757–772. [CrossRef]
50. Pandey, M.K.; Gangurde, S.S.; Sharma, V.; Pattanashetti, S.K.; Naidu, G.K.; Faye, I.; Hamidou, F.; Desmae, H.; Kane, N.A.; Yuan, M.; et al. Improved Genetic Map Identified Major QTLs for Drought Tolerance- and Iron Deficiency Tolerance-Related Traits in Groundnut. *Genes* **2020**, *12*, 37. [CrossRef] [PubMed]
51. Khedikar, Y.P.; Gowda, M.V.C.; Sarvamangala, C.; Patgar, K.V.; Upadhyaya, H.D.; Varshney, R.K. A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **2010**, *121*, 971–984. [CrossRef]
52. Leal-Bertioli, S.C.M.; Cavalcante, U.; Gouvea, E.G.; Ballén-Taborda, C.; Shirasawa, K.; Guimarães, P.M.; Jackson, S.A.; Bertioli, D.J.; Moretzsohn, M.C. Identification of QTLs for Rust Resistance in the Peanut Wild Species Arachis magna and the Development of KASP Markers for Marker-Assisted Selection. *G3 Genes Genomes Genet.* **2015**, *5*, 1403–1413. [CrossRef]
53. Han, S.; Yuan, M.; Clevenger, J.P.; Li, C.; Hagan, A.; Zhang, X.; Chen, C.; He, G. A SNP-Based Linkage Map Revealed QTLs for Resistance to Early and Late Leaf Spot Diseases in Peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* **2018**, *9*, 1012. [CrossRef] [PubMed]
54. Chu, Y.; Chee, P.; Culbreath, A.; Isleib, T.; Holbrook, C.; Ozias-Akins, P. Major QTLs for Resistance to Early and Late Leaf Spot Diseases Are Identified on Chromosomes 3 and 5 in Peanut (*Arachis hypogaea*). *Front. Plant Sci.* **2019**, *10*, 883. [CrossRef]
55. Ahmad, S.; Nawade, B.; Sangh, C.; Mishra, G.P.; Bosamia, T.C.; Kumar, N.; Dobaria, J.R.; Gajera, H.P. Identification of novel QTLs for late leaf spot resistance and validation of a major rust QTL in peanut (*Arachis hypogaea* L.). *3 Biotech* **2020**, *10*, 458. [CrossRef]
56. de Blas, F.J.; Bruno, C.I.; Arias, R.S.; Ballén-Taborda, C.; Mamani, E.; Oddino, C.; Rosso, M.; Costero, B.P.; Bressano, M.; Soave, J.H.; et al. Genetic mapping and QTL analysis for peanut smut resistance. *BMC Plant Biol.* **2021**, *21*, 312. [CrossRef]
57. Tomar, R.S. Introduction to QTL mapping in plants. *Ann. Plant Sci.* **2015**, *4*, 1072–1079.
58. Zhao, Y.; Zhang, C.; Chen, H.; Yuan, M.; Nipper, R.; Prakash, C.S.; Zhuang, W.; He, G. QTL mapping for bacterial wilt resistance in peanut (*Arachis hypogaea* L.). *Mol. Breed.* **2016**, *36*, 13. [CrossRef] [PubMed]
59. Burow, M.D.; Simpson, C.E.; Starr, J.L.; Paterson, A.H. Transmission Genetics of Chromatin From a Synthetic Amphidiploid to Cultivated Peanut (*Arachis hypogaea* L.): Broadening the Gene Pool of a Monophyletic Polyploid Species. *Genetics* **2001**, *159*, 823–837. [CrossRef]
60. Fonceka, D.; Tossim, H.A.; Rivallan, R.; Faye, I.; Leal-Bertioli, S.; Valls, J.F.; Glaszmann, J.C.; Courtois, B.; Rami, J.F. Interspecific Introgression and Trait Dissection in Peanut (*Arachis hypogaea*). In Proceedings of the Plant and Animal Genome Conference, San Diego, CA, USA, 12–16 January 2013.
61. Sambou, A. *Genetic Studies and Mapping of Beneficial Alleles from Wild Species to Improve Agronomic Traits in Cultivated Groundnut (Arachis hypogaea L.)*; University of Legon: Accra, Ghana, 2017.
62. Kumari, V.; Gowda, M.V.C.; Yeri, S.B. Utilization of Advanced Backcross Population Derived from Synthetic Amphidiploid for Dissecting Resistance to Late Leaf Spot in Peanut (*Arachis hypogaea* L.). *Trop. Plant Biol.* **2020**, *13*, 50–61. [CrossRef]
63. Khera, P.; Pandey, M.K.; Mallikarjuna, N.; Sriswathi, M.; Roorkiwal, M.; Janila, P.; Sharma, S.; Shilpa, K.; Sudini, H.; Guo, B.; et al. Genetic imprints of domestication for disease resistance, oil quality, and yield component traits in groundnut (*Arachis hypogaea* L.). *Mol. Genet. Genom.* **2018**, *294*, 365–378. [CrossRef]
64. Fonceka, D.; Tossim, H.-A.; Rivallan, R.; Vignes, H.; Lacut, E.; de Bellis, F.; Faye, I.; Ndoye, O.; Leal-Bertioli, S.C.M.; Valls, J.F.M.; et al. Construction of Chromosome Segment Substitution Lines in Peanut (*Arachis hypogaea* L.) Using a Wild Synthetic and QTL Mapping for Plant Morphology. *PLoS ONE* **2012**, *7*, e48642. [CrossRef] [PubMed]
65. Tossim, H.-A.; Nguepjop, J.R.; Diatta, C.; Sambou, A.; Seye, M.; Sane, D.; Rami, J.-F.; Fonceka, D. Assessment of 16 Peanut (*Arachis hypogaea* L.) CSSLs Derived from an Interspecific Cross for Yield and Yield Component Traits: QTL Validation. *Agronomy* **2020**, *10*, 583. [CrossRef]
66. Essandoh, D.A.; Odong, T.; Okello, D.K.; Fonceka, D.; Nguepjop, J.; Sambou, A.; Ballén-Taborda, C.; Chavarro, C.; Bertioli, D.J.; Leal-Bertioli, S.C.M. Quantitative Trait Analysis Shows the Potential for Alleles from the Wild Species *Arachis batizocoi* and *A. duranensis* to Improve Groundnut Disease Resistance and Yield in East Africa. *Agronomy* **2022**, *12*, 2202. [CrossRef]
67. Alyr, M.H.; Pallu, J.; Sambou, A.; Nguepjop, J.R.; Seye, M.; Tossim, H.-A.; Djiboune, Y.R.; Sane, D.; Rami, J.-F.; Fonceka, D. Fine-Mapping of a Wild Genomic Region Involved in Pod and Seed Size Reduction on Chromosome A07 in Peanut (*Arachis hypogaea* L.). *Genes* **2020**, *11*, 1402. [CrossRef]
68. Scott, M.F.; Ladejobi, O.; Amer, S.; Bentley, A.R.; Biernaskie, J.; Boden, S.A.; Clark, M.; Dell'acqua, M.; Dixon, L.E.; Filippi, C.V.; et al. Multi-parent populations in crops: A toolbox integrating genomics and genetic mapping with breeding. *Heredity* **2020**, *125*, 396–416. [CrossRef]
69. Pandey, M.K.; Roorkiwal, M.; Singh, V.K.; Ramalingam, A.; Kudapa, H.; Thudi, M.; Chitikineni, A.; Rathore, A.; Varshney, R.K. Emerging Genomic Tools for Legume Breeding: Current Status and Future Prospects. *Front. Plant Sci.* **2016**, *7*, 455. [CrossRef] [PubMed]

70. Chu, Y.; Holbrook, C.; Isleib, T.; Burow, M.; Culbreath, A.K.; Tillman, B.; Chen, J.; Clevenger, J.; Ozias-Akins, P. Phenotyping and genotyping parents of sixteen recombinant inbred peanut populations. *Peanut Sci.* **2018**, *45*, 1–11. [CrossRef]
71. Wankhade, A.P.; Kadirimangalam, S.R.; Viswanatha, K.P.; Deshmukh, M.P.; Shinde, V.S.; Deshmukh, D.B.; Pasupuleti, J. Variability and Trait Association Studies for Late Leaf Spot Resistance in a Groundnut MAGIC Population. *Agronomy* **2021**, *11*, 2193. [CrossRef]
72. Pandey, M.K.; Pandey, A.K.; Kumar, R.; Nwosu, C.V.; Guo, B.; Wright, G.C.; Bhat, R.S.; Chen, X.; Bera, S.K.; Yuan, M.; et al. Translational genomics for achieving higher genetic gains in groundnut. *Theor. Appl. Genet.* **2020**, *133*, 1679–1702. [CrossRef]
73. Gangurde, S.S.; Wang, H.; Yaduru, S.; Pandey, M.K.; Fountain, J.C.; Chu, Y.; Isleib, T.; Holbrook, C.C.; Xavier, A.; Culbreath, A.K.; et al. Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnol. J.* **2019**, *18*, 1457–1471. [CrossRef]
74. Wang, M.L.; Sukumaran, S.; Barkley, N.A.; Chen, Z.; Chen, C.Y.; Guo, B.; Pittman, R.N.; Stalker, H.T.; Holbrook, C.C.; Pederson, G.A.; et al. Population structure and marker–trait association analysis of the US peanut (*Arachis hypogaea* L.) mini-core collection. *Theor. Appl. Genet.* **2011**, *123*, 1307–1317. [CrossRef]
75. Pandey, M.K.; Upadhyaya, H.D.; Rathore, A.; Vadez, V.; Sheshshayee, M.S.; Sriswathi, M.; Govil, M.; Kumar, A.; Gowda, M.V.C.; Sharma, S.; et al. Genomewide Association Studies for 50 Agronomic Traits in Peanut Using the 'Reference Set' Comprising 300 Genotypes from 48 Countries of the Semi-Arid Tropics of the World. *PLoS ONE* **2014**, *9*, e105228. [CrossRef] [PubMed]
76. Zhang, X.; Zhang, J.; He, X.; Wang, Y.; Ma, X.; Yin, D. Genome-Wide Association Study of Major Agronomic Traits Related to Domestication in Peanut. *Front. Plant Sci.* **2017**, *8*, 1611. [CrossRef]
77. Zhang, H.; Wang, M.L.; Dang, P.; Jiang, T.; Zhao, S.; Lamb, M.; Chen, C. Identification of potential QTLs and genes associated with seed composition traits in peanut (*Arachis hypogaea* L.) using GWAS and RNA-Seq analysis. *Gene* **2020**, *769*, 145215. [CrossRef]
78. Liu, N.; Huang, L.; Chen, W.; Wu, B.; Pandey, M.K.; Luo, H.; Zhou, X.; Guo, J.; Chen, H.; Huai, D.; et al. Dissection of the genetic basis of oil content in Chinese peanut cultivars through association mapping. *BMC Genet.* **2020**, *21*, 60. [CrossRef]
79. Zhou, X.; Guo, J.; Pandey, M.K.; Varshney, R.K.; Huang, L.; Luo, H.; Liu, N.; Chen, W.; Lei, Y.; Liao, B.; et al. Dissection of the Genetic Basis of Yield-Related Traits in the Chinese Peanut Mini-Core Collection Through Genome-Wide Association Studies. *Front. Plant Sci.* **2021**, *12*, 637284. [CrossRef]
80. Zhou, X.; Luo, H.; Yu, B.; Huang, L.; Liu, N.; Chen, W.; Liao, B.; Lei, Y.; Huai, D.; Guo, P.; et al. Genetic dissection of fatty acid components in the Chinese peanut (*Arachis hypogaea* L.) mini-core collection under multi-environments. *PLoS ONE* **2022**, *17*, e0279650. [CrossRef] [PubMed]
81. Halward, T.; Stalker, H.T.; Kochert, G. Development of an RFLP linkage map in diploid peanut species. *Theor. Appl. Genet.* **1993**, *87*, 379–384. [CrossRef] [PubMed]
82. Herselman, L.; Thwaites, R.; Kimmins, F.M.; Courtois, B.; van der Merwe, P.J.A.; Seal, S.E. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. *Theor. Appl. Genet.* **2004**, *109*, 1426–1433. [CrossRef] [PubMed]
83. Varshney, R.K.; Bertioli, D.J.; Moretzsohn, M.C.; Vadez, V.; Krishnamurthy, L.; Aruna, R.; Nigam, S.N.; Moss, B.J.; Seetha, K.; Ravi, K.; et al. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **2008**, *118*, 729–739. [CrossRef]
84. Foncéka, D.; Hodo-Abalo, T.; Rivallan, R.; Faye, I.; Sall, M.N.; Ndoye, O.; Fávero, A.P.; Bertioli, D.J.; Glaszmann, J.-C.; Courtois, B.; et al. Genetic mapping of wild introgressions into cultivated peanut: A way toward enlarging the genetic basis of a recent allotetraploid. *BMC Plant Biol.* **2009**, *9*, 103. [CrossRef] [PubMed]
85. Zhou, X.; Xia, Y.; Ren, X.; Chen, Y.; Huang, L.; Huang, S.; Liao, B.; Lei, Y.; Yan, L.; Jiang, H. Construction of a SNP-based genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). *BMC Genom.* **2014**, *15*, 351. [CrossRef]
86. Collard, B.C.Y.; Jahufer, M.Z.Z.; Brouwer, J.B.; Pang, E.C.K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* **2005**, *142*, 169–196. [CrossRef]
87. Andersen, S.B.; Torp, A.M. QTL Mapping in Crop Plants. In *Molecular Techniques in Crop Improvement*; Jain, S.M., Brar, D.S., Ahloowalia, B.S., Eds.; Springer: Dordrecht, The Netherlands, 2002; pp. 541–562. [CrossRef]
88. Vassiliou, E.K.; Gonzalez, A.; Garcia, C.; Tadros, J.H.; Chakraborty, G.; Toney, J.H. Oleic acid and peanut oil high in oleic acid reverse the inhibitory effect of insulin production of the inflammatory cytokine TNF-α both in vitro and in vivo systems. *Lipids Health Dis.* **2009**, *8*, 10–25. [CrossRef] [PubMed]
89. Huang, L.; He, H.; Chen, W.; Ren, X.; Chen, Y.; Zhou, X.; Xia, Y.; Wang, X.; Jiang, X.; Liao, B.; et al. Quantitative trait locus analysis of agronomic and quality-related traits in cultivated peanut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **2015**, *128*, 1103–1115. [CrossRef]
90. Wang, M.L.; Khera, P.; Pandey, M.; Wang, H.; Qiao, L.; Feng, S.; Tonnis, B.; Barkley, N.A.; Pinnow, D.; Holbrook, C.C.; et al. Genetic Mapping of QTLs Controlling Fatty Acids Provided Insights into the Genetic Control of Fatty Acid Synthesis Pathway in Peanut (*Arachis hypogaea* L.). *PLoS ONE* **2015**, *10*, e0119454. [CrossRef] [PubMed]
91. Zhao, S.; Li, A.; Li, C.; Xia, H.; Zhao, C.; Zhang, Y.; Hou, L.; Wang, X. Development and application of KASP marker for high throughput detection of AhFAD2 mutation in peanut. *Electron. J. Biotechnol.* **2016**, *25*, 9–12. [CrossRef]

92. Faye, I.; Pandey, M.; Hamidou, F.; Rathore, A.; Ndoye, O.; Vadez, V.; Varshney, R.K. Identification of quantitative trait loci for yield and yield related traits in groundnut (*Arachis hypogaea* L.) under different water regimes in Niger and Senegal. *Euphytica* **2015**, *206*, 631–647. [CrossRef] [PubMed]
93. Chen, Y.; Wang, Z.; Ren, X.; Huang, L.; Guo, J.; Zhao, J.; Zhou, X.; Yan, L.; Luo, H.; Liu, N.; et al. Identification of major QTL for seed number per pod on chromosome A05 of tetraploid peanut (*Arachis hypogaea* L.). *Crop J.* **2019**, *7*, 238–248. [CrossRef]
94. Wang, L.; Yang, X.-L.; Cui, S.-L.; Wang, J.-H.; Hou, M.-Y.; Mu, G.-J.; Li, Z.-C.; Liu, L.-F. Identification of main effect and epistatic QTLs controlling initial flowering date in cultivated peanut (*Arachis hypogaea* L.). *J. Integr. Agric.* **2020**, *19*, 2383–2393. [CrossRef]
95. Lv, J.; Liu, N.; Guo, J.; Xu, Z.; Li, X.; Li, Z.; Luo, H.; Ren, X.; Huang, L.; Zhou, X.; et al. Stable QTLs for Plant Height on Chromosome A09 Identified from Two Mapping Populations in Peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* **2018**, *9*, 684. [CrossRef]
96. Li, L.; Yang, X.; Cui, S.; Meng, X.; Mu, G.; Hou, M.; He, M.; Zhang, H.; Liu, L.; Chen, C.Y. Construction of High-Density Genetic Map and Mapping Quantitative Trait Loci for Growth Habit-Related Traits of Peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* **2019**, *10*, 745. [CrossRef]
97. Liang, Y.; Baring, M.; Wang, S.; Septiningsih, E.M. Mapping QTLs for Leafspot Resistance in Peanut Using SNP-Based Next-Generation Sequencing Markers. *Plant Breed. Biotechnol.* **2017**, *5*, 115–122. [CrossRef]
98. Tanksley, S.D.; Nelson, J.C. Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* **1996**, *92*, 191–203. [CrossRef]
99. Tabor, H.K.; Risch, N.J.; Myers, R.M. Candidate-gene approaches for studying complex genetic traits: Practical considerations. *Nat. Rev. Genet.* **2002**, *3*, 391–397. [CrossRef]
100. Francia, E.; Tacconi, G.; Crosatti, C.; Barabaschi, D.; Bulgarelli, D.; Dall'Aglio, E.; Valè, G. Marker assisted selection in crop plants. *Plant Cell Tissue Organ Cult.* **2005**, *82*, 317–342. [CrossRef]
101. Nadeem, M.A.; Nawaz, M.A.; Shahid, M.Q.; Doğan, Y.; Comertpay, G.; Yıldız, M.; Hatipoğlu, R.; Ahmad, F.; Alsaleh, A.; Labhane, N.; et al. DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. *Biotechnol. Biotechnol. Equip.* **2018**, *32*, 261–285. [CrossRef]
102. Vishwakarma, M.; Yadav, P.; Rai, V.; Kumar, U.; Joshi, A. *Molecular Markers and Genomics Assisted Breeding for Improving Crop Plants*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 303–334. [CrossRef]
103. Singh, B.D.; Singh, A.K. *Marker-Assisted Plant Breeding: Principles and Practices*; Springer: New Delhi, India, 2015. [CrossRef]
104. Kumawat, G.; Kumawat, C.K.; Chandra, K.; Pandey, S.; Chand, S.; Mishra, U.N.; Lenka, D.; Sharma, R. Insights into Marker Assisted Selection and Its Applications in Plant Breeding. In *Plant Breeding-Current and Future Views*; IntechOpen: London, UK, 2021; ISBN 978-1-83968-311-4. [CrossRef]
105. Nayak, S.N.; Singh, V.K.; Varshney, R.K. Marker-Assisted Selection. In *Encyclopedia of Applied Plant Sciences*; Thomas, B., Murray, B.G., Murphy, D.J., Eds.; Elsevier: Amsterdam, The Netherlands; Academic Press: Cambridge, MA, USA, 2017; Volume 2, pp. 183–197. [CrossRef]
106. Kolekar, R.M.; Sukruth, M.; Shirasawa, K.; Nadaf, H.L.; Motagi, B.N.; Lingaraju, S.; Patil, P.V.; Bhat, R.S. Marker-assisted backcrossing to develop foliar disease-resistant genotypes in TMV 2 variety of peanut (*Arachis hypogaea* L.). *Plant Breed.* **2017**, *136*, 948–953. [CrossRef]
107. Bera, S.K.; Kamdar, J.H.; Kasundra, S.V.; Patel, S.V.; Jasani, M.; Maurya, A.K.; Dash, P.; Chandrashekar, A.B.; Rani, K.; Manivannan, N.; et al. Steady expression of high oleic acid in peanut bred by marker-assisted backcrossing for fatty acid desaturase mutant alleles and its effect on seed germination along with other seedling traits. *PLoS ONE* **2019**, *14*, e0226252. [CrossRef]
108. Shasidhar, Y.; Variath, M.T.; Vishwakarma, M.K.; Manohar, S.S.; Gangurde, S.S.; Sriswathi, M.; Sudini, H.K.; Dobariya, K.L.; Bera, S.K.; Radhakrishnan, T.; et al. Improvement of three popular Indian groundnut varieties for foliar disease resistance and high oleic acid using SSR markers and SNP array in marker-assisted backcrossing. *Crop J.* **2020**, *8*, 1–15. [CrossRef]
109. Tang, Y.; Qiu, X.; Hu, C.; Li, J.; Wu, L.; Wang, W.; Li, X.; Li, X.; Zhu, H.; Sui, J.; et al. Breeding of a new variety of peanut with high-oleic-acid content and high-yield by marker-assisted backcrossing. *Mol. Breed.* **2022**, *42*, 42. [CrossRef]
110. Gokidi, Y.; Bhanu, A.; Singh, M.N. Marker Assisted Recurrent Selection: An Overview. *Adv. Life Sci.* **2016**, *5*, 6493–6499.
111. Chukwu, S.C.; Rafii, M.Y.; Ramlee, S.I.; Ismail, S.I.; Oladosu, Y.; Okporie, E.; Onyishi, G.; Utobo, E.; Ekwu, L.; Swaray, S.; et al. Marker-assisted selection and gene pyramiding for resistance to bacterial leaf blight disease of rice (*Oryza sativa* L.). *Biotechnol. Biotechnol. Equip.* **2019**, *33*, 440–455. [CrossRef]
112. Xu, L.Y.; Zhao, F.P.; Sheng, X.H.; Ren, H.X.; Zhang, L.; Wei, C.H.; Du, L.X. Optimal Design for Marker-assisted Gene Pyramiding in Cross Population. *Asian-Australas. J. Anim. Sci.* **2012**, *25*, 772–784. [CrossRef]
113. Servin, B.; Martin, O.; Mézard, M.; Hospital, F. Toward a Theory of Marker-Assisted Gene Pyramiding. *Genetics* **2004**, *168*, 513–523. [CrossRef]
114. Chu, Y.; Wu, C.L.; Holbrook, C.C.; Tillman, B.L.; Person, G.; Ozias-Akins, P. Marker-Assisted Selection to Pyramid Nematode Resistance and the High Oleic Trait in Peanut. *Plant Genome* **2011**, *4*, 110–117. [CrossRef]
115. Lande, R.; Thompson, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **1990**, *124*, 743–756. [CrossRef] [PubMed]

116. Mackill, D.J.; Collard, B.C.Y.; Neeraja, C.N.; Rodriguez, R.M.; Heuer, S.; Ismail, A.M. QTLs in rice breeding: Examples for abiotic stresses. In Proceedings of the Fifth International Rice Genetics Symposium, Manila, The Philippines, 19–23 November 2005. [CrossRef]
117. Pillen, K.; Zacharias, A.; Léon, J. Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *107*, 340–352. [CrossRef] [PubMed]
118. Bhanu, A.; Godiyi, Y.; Singh, M.N. Advanced Backcross QTL Method: A Brief Overview. *Trends Biosci.* **2017**, *10*, 20–25.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.