META-ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA

By

RESHMA PARVEEN J.

(2014-09-104)

THESIS

Submitted in partial fulfilment of the

requirement for the degree of

B. Sc. - M. Sc. (INTEGRATED) BIOTECHNOLOGY

Faculty of Agriculture

Kerala Agricultural University, Thrissur

DEPARTMENT OF PLANT BIOTECHNOLOGY
COLLEGE OF AGRICULTURE
VELLAYANI, THIRUVANANTHAPURAM-695 522
KERALA, INDIA

2019

DECLARATION

I hereby declare that the thesis entitled **"META-ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA"** is a bonafide record of research work done by me during the course of research and that the thesis has not previously formed the basis for the award of any degree, diploma, associateship, fellowship or other similar title, of any other University or Society.

Place: Vellayani

Date: 29-11-2019

Reshma Parveen J.

2014-09-104

भा.कृ.अनु.प- केंद्रीय कन्द फसल अनुसंधान संस्थान

(भारतीय कृषि अनुसंधान परिषद, कृषि और किसान कल्याण मंत्रालय, भारत सरकार)
श्रीकार्यम, तिरुवनन्तपुरम-695 017, केरल, भारत

ICAR- CENTRAL TUBER CROPS RESEARCH INSTITUTE

(Indian Council of Agriculture Research, Ministry of Agriculture and Farmers Welfare, Govt. of India)
Sreekariyam, Thiruvananthapuram-695 017, Kerala, India

ISO 9001:2008

CERTIFICATE

Certified that this thesis entitled **"META-ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA"** is a record of research work done independently by **Ms. RESHMA PARVEEN J. (2014-09-104)** under my guidance and supervision and this has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to her.

Place: Sreekariyam
Date: 29-11-2019

Dr. J. Sreekumar
(Chairman, Advisory Committee)
Principal Scientist (Agrl. Statistics),
Section of Extension and Social Sciences,
ICAR-CTCRI, Sreekariyam
Thiruvanathapuram-695 017

Phone : +91-471-2598551 to 2598554
Director (Per) : +91-471-2598431
(Res) : +91-471-2597211
Sr. Admn. Officer : +91-471-2598193
Fax : +91-471-2590063
E-mail : director.ctcri@icar.gov.in
ctcritvm@gmail.com
Web : http://www.ctcri.org

CERTIFICATE

We, the undersigned members of the advisory committee of Ms. Reshma Parveen J. (2014-09-104) a candidate for the degree of B. Sc. - M. Sc. (Integrated) Biotechnology, agree that the thesis entitled **"META - ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA"** may be submitted by Ms. Reshma Parveen J. in partial fulfillment of the requirement for the degree.

Dr. J. Sreekumar
Chairman, Advisory committee
Principal Scientist (Agrl. Statistics),
Section of Extension and
Social Sciences.
ICAR-CTCRI, Sreekariyam,
Thiruvananthapuram - 695 017

Dr. K. B. Soni
(Member, Advisory Committee)
Professor and Head
Department of Plant Biotechnology
College of Agriculture, Vellayani
Thiruvananthapuram- 695 552

Dr. C. Mohan
(Member, Advisory Committee)
Principal Scientist (Plant Breeding)
Division of Crop improvement
ICAR- CTCRI
Sreekariyam,
Thiruvananthapuram- 695 017

Dr. Swapna Alex
(Member, Advisory Committee)
Professor and Course Director
B. Sc. - M. Sc. (Integrated)
Biotechnology Course
Department of Plant Biotechnology
College of Agriculture, Vellayani
Thiruvananthapuram- 695 552

Dr. M. K. Rajesh
(External Examiner)
Principal Scientist (Biotechnology)
Division of Crop improvement
ICAR- CPCRI
Kasaragod - 671124

ACKNOWLEDGEMENT

First and foremost, I would like to thank God Almighty for giving me the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it satisfactorily. Without his blessings, this achievement would not have been possible. Gratitude comes from within and with joy and gratitude acknowledges the help of those who were involved in the successful completion of the endeavor.

It is with my heartfelt feelings, I wish to express my deep sense of gratitude to my beloved advisor Dr. J. Sreekumar, Principal Scientist, ICAR-CTCRI, Thiruvananthapuram for his exemplary guidance, patience, and constant encouragement. In addition to his support for completing the project, his advice was valuable for me in every aspect.

My special thanks to the Director of ICAR-CTCRI Dr. Archana Mukherjee and Dr. Sheela Immanual, , Head, Section of Extension and Social Sciences, ICAR-CTCRI for permitting me to do my project work and for her support.

I express my profound gratitude to Dr. Swapna Alex, Dr. K. B. Soni, Dr. C. Mohan, not only for their insightful comments and encouragement, but also for their valuable discussions and critical suggestions related to the project.

My special thanks to my BSL mates Pattu and Raju for being with me in my all stressfull situations, loved, supported and helped me get through this agonizing period in the most positive way

My sincere gratitude is towards Ambu Chettan for giving valuable advice and helping me throughout my project.

My very sincere gratitude towards Dr. Senthil, Scientist and Prakash ettan, Technical Assistant, ICAR-CTCRI who helped me with the technical aspects of my research work including the handling of various instrument.

I am very thankful to Dr. Makeshkumar T., Dr. Sangeetha, Jayakrishnan chettan, Sumayya chechi, Divya chechi, Iype chettan and merlin chechi of Division of crop protection, ICAR-CTCRI for their valuable timely help during the difficult situation of the work.

My acknowledgement would be lacking if I don't mention my gratitude to my beloved friends, Hasu, Ambu, Jo, Adi, Keeru, Vishnu, Muhzi, Neethu, Ami, Akku, Lachu, Sree,

Alif and Linu for their selfless help and support throughout my work.

I want to express my gratitude and deepest appreciation to my friends M. J. Jithu, Rahul and Amal for being with me and supporting me.

I deeply indebted to all the scientists and staff members of ICAR- CTCRI, teachers in college, my seniors and juniors for their timely support.

I acknowledge the favour of numerous persons who, though not been individually mentioned here, who have all directly or indirectly contributed to this work.

My acknowledgement would be incomplete without thanking the biggest source of my strength, my family: my parents and to my sisters for supporting me spiritually throughout writing this thesis and my life in general.

Reshma Parveen J.

DEDICATED TO MY PARENTS AND SISTERS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

Sl. No.	Title	Appendix No.
1	CTAB extraction buffer	I
2	TBE Buffer (10 X)	II
3	70% ethanol	III

LIST OF ABBREVIATIONS

%	Percentage
A260	Absorbance at 260 nmwavelength
A280	Absorbance at 280 nmwavelength
AIC	Akaike information criterion
AFLP	Amplified fragment length polymorphism
BLAST	Basic local alignment search tool
bp	Basepair
CAD	Cassava anthracnose disease
CBB	Cassava bacterial blight
CBSD	Cassava brown streak disease
CBSD-RN	Cassava brown streak disease-Root necrosis
CG	Candidate genes
CI	Confidence interval
CIM	Composite interval mapping
cM	Centi morgan
CMD	Cassava mosaic disease
DNA	Deoxyribonucleic acid
et al.	et alia
F	Forwardprimer
GO	Gene ontology
IM	Interval mapping
LG	Linkage group
LOD	Logarithm of odds
MAS	Marker assisted selection
min	Minute
MIM	Multiple interval mapping
MQTL	Meta-QTL
Mb	Mega basepair
PCR	Polymerase chain reaction
PFAM	Protein family
PVE	Phenotypic variance explained

QTL	Quantitative Trait Loci
R	Reverse primer
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	revolutions per minute
RT	Reverse transcriptase
s	second
SIM	Simple interval mapping
SNP	Single nucleiotide polymorphism
SSR	Simple sequence repeats
Tm	Meltingtemperature
µl	Microlitre

INTRODUCTION

1. INTRODUCTION

Manihot esculenta Crantz , commonly known as cassava, is a tropical root crop and perennial shrub in the family *Euphorbiaceae,* which is indigenous to South America. For its edible starchy tuberous root, a major source of carbohydrates, cassava is extensively grown as an annual crop in tropical and subtropical regions. After rice and maize, cassava is the third largest source of food carbohydrates in the tropics, providing the staple food of around 800 million individuals globally (Save and Grow: Cassava. 2013).Globally Nigeria ranks first in terms of cassava production, 20.6% global production of Cassava production is from them and as per FAO 2017 their net production was about 59.58 million tonnes. As per FAO 2018 the worldwide cassava production was about 277.1 million tonnes. In South Asia, cassava plays a role in food security in India, particularly in the major growing states of Kerala (consumption of fresh roots) and Tamil Nadu (starch for food manufacturing). Combined, the two states account for 98 percent of national output (FAO 2017).

Cassava is susceptible to pests and diseases that can cause heavy yield losses. Cassava Mosaic Disease (CMD), and Cassava Brown Streak Disease (CBSD) are the major diseases affecting cassava. The Cassava Mosaic Disease (CMD) is caused by several distinct Begomovirus species and the strains are transmitted by whitefly (*Bemisia tabaci*) (Brown *et al.,* 1995). Cassava brown streak disease (CBSD), the most serious threat to cassava cultivation and is caused by two virus species, *Cassava Brown Streak Virus* (CBSV) and *Ugandan Cassava Brown Streak Virus* (UCBSV) (Winter *et al.,* 2010).

The most significant economic traits in plants, including disease resistance are classified into complex traits whose features are determined by both genetic and environmental factors and such traits are regulated by QTLs (Wu *et al.,* 2016). A Quantitative Trait Loci (QTL) is a specific chromosome fragment that correlates with the variability of certain phenotypes.

Several QTLs for disease resistance in cassava have been recognized and reported by QTL analysis. The molecular markers associated with these QTLs can be used for marker assisted selection (MAS) to develop disease-resistant varieties. Selection based on molecular markers that define a quantitative trait locus (QTL) can efficiently improve the heritability of the associated trait by negating environmental influence (Masumba *et al.*, 2017).

The successful use of QTL in marker-assisted selection relies on their impacts and consistency across different genetic backgrounds and environments. Identification of major and stable QTLs from a single study is so difficult because it is very time-consuming, laborious, and costly to build large mapping populations, collect the phenotypic data and genotyping the mapping population to construct a high-density genetic map, and evaluate QTLs in multiple environments or different years (Swami *et al.*, 2011). Additionally, the QTL identified from one genetic background may not be transferrable to other backgrounds because of unfavourable epistatic interactions resulting in reduced or even no effects in a new genetic background (Collins *et al.*, 2008), so it is difficult to predict the usefulness of QTL for MAS based only on their performance in an individual genetic background in any particular study.

A more efficient way to select QTL for MAS is to compare the identified QTLs studies for their consistency of location and effect across genetic backgrounds and environments. Consistently identified QTLs at the same chromosomal location, explaining high phenotypic variance and having a major effect on a trait, can be effectively used in MAS. Integration of results from multiple QTL studies will enhance our knowledge of the genetic basis for dissection of complex traits (Veyrieras *et al.*, 2007).

Meta-analysis is a statistical powerful tool that can be used for QTL detection and precise estimation of their genetic effects by integrating information from multiple QTL studies (Wu *et al.*, 2016).The meta-analysis method is stronger than those of individual studies and can give greater insight into the

genetic architecture of complex traits and this method is simple, labour-saving, and also saves money (Wu and Hu 2012). In QTL meta-analysis, information about previously reported QTLs is required. Meta-QTL (MQTL) with the smallest Confidence Interval (CI) and having a consistent and large effect on a trait identified by meta-analysis can be used for MAS. In plants, the concept of meta-analysis has been applied to the analysis of QTL/genes for rice panicle related traits (Wu *et al.,* 2016), seed iron and zinc concentration and content in common bean (Izquierdo *et al.,* 2018), seedling-stage salt tolerance in rice (Islam *et al.,* 2019).

QTL regions consist of many genes; among them, a few key genes could be more important in the regulation of a complex trait. MQTLs with precise and narrow physical intervals obtained from the meta-analysis are useful in mining and listing the genes underlying the QTLs i.e. the candidate genes. The digital candidate gene approach *i.e.*, based on the computer is primarily applied to identify potential candidate genes which may be associated with the traits. By meta -analysis method the genes in the MQTL region which may contribute to the trait under study can be identified.

The present study conducted with an objective to construct a consensus map from different QTL mapping studies associated with cassava mosaic disease resistance by meta-analysis, to identify Meta-QTLs regions and to identify the candidate genes present in these MQTL regions and then it is followed by validation of these QTLs using real time-PCR.

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

2.1 CASSAVA

Cassava (*Manihot esculenta* Crantz) is a perennial tropical crop cultivated for its starch containing tuberous roots. It is an outbreeding species possessing $2n = 36$ chromosomes. Cassava is commonly cultivated in tropical Africa, Asia, and Latin America and is an important food security crop for many tropical and subtropical nations. It is primarily grown by small-scale farmers and consumed by an approximately 500 million individuals every day. The starchy roots are primarily used as human food, fresh when low in cyanogens, or in many processed forms and products, mostly starch, flour, and for animal feed (Balyejusa Kizito. 2006.). Cassava ranks fourth as a food crop in developing nations, after rice, corn, and wheat, according to the Food and Agriculture Organization of the United Nations (FAO). Approximately 70 million people in Africa are dependent on cassava as a main food source (FAO, 2019). In South Asia (India) and Southeast Asia (Indonesia and the Philippines), cassava is widely consumed as food (FAO, 2017). Cassava is the basis of various products, including food, flour, animal feed, alcohol, starches for sizing paper and textiles, sweeteners, prepared foods and bio-degradable products. Cassava can grow and generate high yields in areas where maize and other crops will not grow or perform well. It can tolerate drought and can be cultivated on low nutrient soils, but it responds well to irrigation or greater rainfall regions and to the use of fertilizers. Cassava is highly flexible in its requirements for management and has the potential to produce high-energy per unit area of land. The crop has long been used as a hunger reserve and food safety crop. Because cassava has no definite maturation point, harvesting can be postponed until market, processing or other conditions are more favourable, this flexibility implies that cassava can be field stored for several months, therefore Cassava is highly acceptable in rural areas.

2.2 PESTS AND DISEASES IN CASSAVA

Like other crops, cassava is vulnerable to pests and diseases that can cause heavy yield losses. It includes viral, bacterial and fungal diseases. The major pests of cassava are the cassava green mite, thrips, mealybugs and white flies. The main diseases affecting cassava are Cassava Mosaic Disease (CMD), Cassava Brown Streak Disease (CBSD), cassava bacterial blight, cassava anthracnose disease, and root rot.

2.2.1. Cassava bacterial blight

Cassava Bacterial Blight (CBB) is caused by a Gram-negative bacteria *Xanthomonas axonopodis phaseoli v. manihotis* (Xam). It is one of the most severe diseases of cassava in several countries where the crop plays an important dietary and economic role. It causes various symptoms like canker with exudates on stems, wilting of leaves etc. Under favourable ecological conditions, wilting of leaves and leaf fall due to CBB can be high. This disease also causes losses of fresh roots and also of planting material. Epidemics occur during the rainy season when high humidity and warm temperature favour the movement of bacteria and symptoms development. As CBB affects systemically cassava stems, this leads to shortages in the supply of healthy (bacteria-free) planting materials. (Fanou *et al.*, 2018).

2.2.2. Cassava anthracnose disease

Cassava anthracnose *(Colletotrichum* spp.) is an important stem disease-causing canker, die-back, and defoliation. It especially damages the grassland Savanna regions of Central Africa, where soils are in fertile and acidic and populations of the insect *Pseudo theraptus devastans,* which is associated with infection of the disease, are high.(Hahn *et al.*, 1985).

2.2.3. Cassava Mosaic Disease

Cassava Mosaic Disease (CMD) is caused by Cassava Mosaic Virus of the genus *Begomovirus* in the family Geminiviridae that are transmitted by the whitefly *Bemisia tabaci* (Pita *et al.,* 2001). They can cause yield reductions of up to 90 percent, CMD is prevalent and causes severe losses in all the important cassava growing areas of Sub Saharan Africa. It is widespread throughout Africa and India. CMD produces a variety of foliar symptoms that include mosaic, mottling, misshapen and twisted leaflets, and an overall reduction in size of leaves and plants (Alabi *et al.,* 2011).Additional begomovirus species identified includes East African Cassava Mosaic Cameroon Virus (EACMCV), East Africa Cassava Mosaic Malawi Virus (EACMMV) and East African Cassava Mosaic Zanzibar Virus (EACMZV) (Bull *et al.,* 2006). Two *Cassava Mosaic Geminivirus* (CMG) reported from Indian sub-continents are *Indian Cassava Mosaic Virus* (ICMV) and *Sri Lankan Cassava Mosaic Virus* (SLCMV). (Saunders *et al.,* 2002).

2.2.4 Cassava Brown Streak Disease

Cassava Brown Streak Disease (CBSD) is a devastating viral disease that causes severe damage to cassava roots causing significant crop losses to cassava production in affected areas. CBSD is caused by *Cassava Brown Streak Virus* (CBSV), a mono partite single-stranded RNA encapsulated into flexuous filamentous particles and has been placed in the genus *Ipomovirus,* family Potyrividae (Monger *et al.,* 2001). CBSD was first described and recorded by Storey at the Amani research institute at the foot of the Usambara mountains of Tanzania in the 1930s (Storey *et al.,* 1936). Due to brown, dry necrotic rot in the storage roots, that renders the roots useless, it is considered as the major threat to food security in the worst affected areas of coastal, eastern and southern Africa. Because of the associated symptoms of root necrosis, the disease has a direct impact on both yield and root. (Kulembeka *et al*., 2010). Work on pathological studies by Storey (1939) led to suspicions that whitefly (*Bemisia* spp) was the

vector of the virus in the field. Vector transmission of CBSV by whitefly (*Bemisia tabaci* Gennadius) was confirmed by (Maruthi *et al.*, 2005).

2.2.5 PESTS

Cassava pests represent a wide range of arthropods; approximately 200 species have been recorded. Although many are minor pests, causing little or no economic losses, several are classified as major pests. These include mites, thrips, mealybug, and whiteflies (Bellotti *et al.*, 1978). The pests that co-evolved with the crop cause more economic damage or yield loss.

2.2.5.1. Whiteflies

In Africa, the whitefly *Bemisia tabaci* is presently considered to be the major pest of cassava because it is the vector of Cassava Mosaic Disease (CMD). Whiteflies feed on phloem of cassava leaves and can cause direct damage to cassava inducing leaf curling, chlorosis and defoliation.

2.2.5.2. Mealybugs

They feed on cassava and inject a toxin that causes leaf withering. The origin of mealybug is probably in northern South America, where it was found in cassava growing regions of Colombia and Venezuela. It was first reported in northeast Brazil during the mid-1970s, where high populations caused considerable yield losses. Due to this pest cassava production decreased in Brazil during the 1980s. In Africa, yield losses due to Mealybugs were around 80 %. (Bellotti, A.C., 2008).

2.2.5.3. Cassava mites

They are major pest in all cassava-producing regions. The introduction of green mites from Latin America devastated African cassava production in the early 1970s. They were brought under control by a predatory mite introduced from Brazil, which reduced substantially the damage caused by the pest. Cassava

mites can also be controlled using resistant or tolerant varieties, and by fertilizing the crop to improve plant vigour (FAO2016).

2.2.5.4 Thrips

The most important thrips species is *Frankliniella williamsi.* Thrips larvae and adults feed on the growing points and young leaves of cassava and inhibiting their development. Leaflets affected by pests are deformed and show irregular chlorotic spots. Yield reductions induced by *F. williamsi* range from 5–28%, depending on varietal susceptibility. The average reduction for eight varieties in Colombia was 17.2%. (Bellotti. 2008).

2.3. MOLECULAR GENETICS IN CASSAVA

The introduction of molecular DNA markers, and plant genetic transformation promises to provide methods for overcoming breeding obstacles of plants with long growth cycle because selection can be made earlier in the growth cycle, even at the seedling stage. The identification of naturally occurring DNA sequence polymorphisms in different individuals within a species or group forms the basis for an application of DNA markers. In contrast to traditional breeding methods that depend on the direct selection by phenotypic effect only, DNA markers use indirect selection by identification of desirable genotypes for quantitative traits faster and earlier than the time such traits may be assessed phenotypically. Genetic markers have become basic tools for understanding the inheritance and diversity of natural variation. The genetic markers like morphological and biochemical markers, such as isozymes have been used in cassava. Over the last decade, a number of DNA markers have been developed and used in the study of genes, the cassava genome and genetic diversity in cassava. Marker systems such as Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNAs (RAPDs), Sequence Tagged Sites (STS), Expressed Sequence Tags (EST), Amplified Fragment Length Polymorphism (AFLPs), Single nucleotide polymorphisms (SNPs) and Simple

Sequence Repeats (SSRs) have been developed and applied for breeding (Balyejusa Kizito. 2007). RFLP, AFLP, and SSR markers stand out as most effective in detecting polymorphism in cassava (Weising *et al.,* 2005). However, given a large amount of DNA required for RFLP detection and the difficulties in automating RFLP analysis and the dominance of AFLP markers and their requirement of high-quality DNA to ensure complete restriction (Weising *et al.,* 2005) makes SSRs the markers of choice. DNA markers have greatly contributed to cassava breeding and genetics in the understanding of the phylogenetic relationships in the genus (Fregene *et al.,* 1994) , assessment of genetic diversity (Mkumbira *et al.,* 2003) and development of genetic maps and identification of quantitative trait loci (QTL) for some traits of importance (Okogbenin and Fregene 2002).

2.4. CASSAVA DISEASE RESISTANCE BREEDING THROUGH MARKER ASSISTED SELECTION (MAS)

The potential benefits of using markers linked to genes of interest in breeding programmes, thus moving from phenotype based towards genotype-based selection, have been obvious for many decades. MAS is a method whereby a phenotype was selected based on the genotype of a marker (Collard *et al.*, 2005). Identifying resistance genes using molecular markers was the basic pre-requisite for performing MAS in resistance breeding programmes (Miah *et al.*, 2013). The two main strategies used to identify molecular markers associated with traits of interest were QTL mapping and Bulk Segregant Analysis (Giovannoni *et al.*, 1991). Genes of agronomic and scientific importance can be isolated especially on the basis of their position on the genetic map by using molecular marker technologies. Identified tightly linked markers for resistance genes helps in detecting plants carrying these genes simultaneously without subjecting them to the pathogen or insect attack in early generations. Little amount of DNA is required from each of the individual plants to be tested without destroying the plant. Marker-assisted selection can be used to pyramid the major resistance genes, to produce varieties with more desirable characters and it also help

breeders to conduct many rounds of selection in a year (Mohan *et al.*, 1997). Using MAS resistance for CMD and Cassava Green Mite (CGM) was developed by International Centre for Tropical Agriculture (CIAT) and National breeding programs. MAS has helped in the breeding of *CMD2*-meditated resistance in Latin America and in Africa, where the disease is most prevalent (Blair *et al.*, 2007). The markers *RME1* and *NS158* was identifies as an excellent predictors of CMD resistance (Okogbenin *et al.*, 2007). CMD resistance was introgressed into improved elite CIAT lines using these markers. These are now referred to as the CRseries (CR families). Two markers (NS1009 and NS346) associated with CGM resistance have been used in MAS (Okogbenin *et al.*, 2007) Combining CMD and CGM resistance, the markers are being used to transfer CMD resistance into desirable genetic backgrounds of East African farmer preferred varieties with CBSD tolerance in order to combine resistance to both viral diseases. Results from MAS-bred CGM genotypes indicate variation in response to the pest. Progenies selected with the markers for CGM resistance tended to show good resistance to the pest in East Africa in contrast to the moderate tolerance observed for CGM in West Africa (Ferguson *et al.*, 2012).

2.5 CASSAVA DISEASE RESISTANCE BREEDING THROUGH QTL MAPPING

Many valuable agricultural features such as yield, quality and some sort of resistance to disease are regulated by many genes (polygenes) and are defined as quantitative traits. A QTL is a map position on the chromosome localized relative to the locus of the genetic marker and is identified by significant correlations between segregation at a specific locus of the genetic marker and variation in the value of quantitative (trait) resistance (Pandey *et al.*, 2006). A basic principle of QTL mapping is to identify the genomic region comprising few or more genes regulating these complex traits. Through marker-assisted selection and molecular breeding, the information obtained on the QTL analysis can be used to improve crops. The markers are used to divide the mapping population into distinct genotypic groups at the marker locus based on genotypes and perform correlative

statistics to determine whether the individual of one genotype varies considerably from the individuals of another genotype in relation to the trait being studied. Due to recombination a significant p-value obtained for the differences between the marker and QTL. The lesser the distance between marker and QTL, the lower the chance of recombination occurring between them, therefore, the QTL and marker will be usually be inherited together in the progeny, and the mean of the group with the tightly-linked marker will be significantly different ($p<0.05$) to the mean of the group without the marker. Unlinked markers are located far apart or on different chromosomes to the QTL and they are randomly inherited with the QTL; therefore, no significant differences between means of the genotype groups will be detected. QTL mapping includes the construction of genetic linkage maps. The simplest method for QTL mapping is an analysis of variance (ANOVA) at the marker loci, sometimes called marker regression or single-marker analysis. Other methods are interval mapping (Lander and Botstein, 1989), composite interval mapping (Jansen *et al.*, 1993, Zeng *et al.*, 1993, 1994) and multiple interval mapping. The first step in a QTL mapping experiment is usually to construct a mapping population (often at the F_2 level) that originates from parents that differentiate as far as possible for the specific trait of interest. The second step is to look for associations between genotypes and phenotypes in the (F_2) mapping population. If the parents carry different alleles for the QTL controlling a given trait the trait values in the segregating mapping population will be associated with the alleles of the markers that are closely linked to the QTL. By scanning the markers on a linkage map for association with trait values, likely map positions for QTL can be detected, which is an important step towards understanding the inheritance and genetic basis of the traits (Kizito *et al.*, 2007).

2.5.1. Disease resistance QTLs in Cassava

Lokko *et al.* in 2005 identified molecular markers associated with resistance to CMD using F_1 progenies derived from a cross between the CMD resistant landrace TME7 and the susceptible line TMS30555, their QTL analysis further established associations between resistance to CMD and three markers,

SSRY28-180, SSRY106-270, and E-ACC/M-CTC-225. The three markers found to be donated by the resistant parent. TME7, SSRY28-180 accounted for 57.41% of total phenotypic variation for resistance. CMD1 is a quantitative source of resistance identified in the Amani derived inter specific variety TMS 30572 by Fregene *et al.*, 2000.

Akano *et al.* in 2002 identified a gene, designated as *CMD2*, which contributes to CMD resistance in an F1 progeny from a cross between a resistant variety TME 3 (Nigerian landrace) and a susceptible improved line (TMS 30555). CMD2 was flanked by the SSR and RFLP markers. The marker, SSRY28, associated with CMD2 is found to located on linkage group R of the male-parent-derived molecular genetic map.

A study was conducted by Okogbenin *et al.* (2012) using two Nigerian cassava cultivars (TMS 97/2205 and TMS 98/0505) and found that the *CMD2* gene was associated with CMD resistance in TMS 97/2205.They also identified a novel marker linked to the QTL (*CMD3*) for CMD resistance located at NS198 marker loci explaining11% of the phenotypic variance. This SSR marker was located on the same linkage group as the *CMD2* gene but was 36 cM away from the *CMD2* marker loci (SSRY28, SSRY158, and NS169). They concluded that the combined effect of this QTL and *CMD2* may account for the high level of resistance of TMS 97/2205.

Lo'-pez *et al.* (2007) found two QTLs associated with resistance to Cassava Bacterial Blight pathogen *Xanthomonas axonopodis phaseoli* v. *manihotis (Xam)* in F1 progenies developed from a cross between TMS 30572 (female parent) and CM2177-2 (male parent). One QTL was associated with marker rGY75 and accounted for 21.4% of the phenotypic variance of resistance to *Xam* strain CIO121, the other QTL was inherited from the male parent and accounted for 61% of resistance to strain CIO151.

A total of three QTL regions associated with Cassava Anthracnose Disease (CAD) infection were detected by a study, conducted by Boonchanawiwat *et al.*

(2016) on three linkage groups. In 2010, two QTLs, CAD_10R1, and CAD_10R2 were detected and it was located on linkage group 2 and 3, respectively. QTL, CAD_11R3 located on linkage group 9 was identified in 2011.

Mohan *et al.* (2013) identified four markers associated with CMD resistance NS136, SSRY44 SSRY235, SSRY28 using mapping population generated from CO2/MNga-1. They found that SSRY28 and SSRY235 are strongly associated with QTL contributing resistance to CMD.

SU.E *et al.,* in 2016 identified a single locus for CMD resistance with a strong peak LOD and explaining 60% of the phenotypic variation identified on LG 15 of the linkage map using F1 progenies derived from a cross between the CMD resistant landrace TMS961089A and the susceptible line TMEB117.

Soto *et al.* (2016) identified two novel QTL associated with CBB resistance. One QTL was located in linkage groups 4 and named as *QLB-4* with LOD 2.5 and explains the 12.6% the field resistance to CBB. The interval flanking markers of *QLB-4* were MB_21980 and MB_2536. The second QTL was located on the linkage group 8, explained 10.9 % of CBB resistance and was named as *QLB-8*. whit a peak marker matching to the SNP MB_8500 7.

Genetic mapping of quantitative trait loci (QTL) for resistance to cassava brown streak disease (CBSD), cassava mosaic disease (CMD), and cassava green mite (CGM) was performed using an F1 cross developed between the Tanzanian landrace, Kiroba, and a breeding line, AR37-80 by Nzuki *et al.* in 2017. Fifteen significant QTLs were identified; three QTL associated with CBSD root necrosis, namely qCBSDRNc5K, qCBSDRNFc11K,and qCBSDRNc12K, detected on chromosomes V, XI, and XII, with maximum LOD-values of 6.20, 13.45, and 11.05, respectively, explaining up to 10.1% of the phenotypic variation (PVE%). Seven QTL associated with CBSD foliar symptoms only, namely qCBSDFc4KL, qCBSDFc4R,qCBSDFc6KRa and b, qCBSDFc17K, qCBSDFc18Ka, and b, was detected on chromosomes IV, VI, XVII and XVIII, respectively. They have maximum LODs of 2.78, 60.67, 54.75, 20.92, 27.01, 23.72 and 23.08,

respectively, and explain up to 8.45% of the variation. Two QTL, namely qCMDc12Ar and qCMDc14Ar, associated with CMD resistance was detected on chromosomes XII and XIV with maximum LODs of 13.20. They found candidate genes within QTLs which encodes several protein domains that have been reported to be involved in disease resistance in plants. The F-box protein domains found on chromosome XVIII are said to contain Leucine-Rich Repeat (LRR) domains associated with pathogen responses.

A study was conducted by Masumba *et al.* in 2017 for the identification and characterization of QTL associated with CBSD and CMD resistance in the Tanzanian landraces Namikonga and Albert. They identified two QTL with consistent flanking markers across seasons on chromosomes XI and II, qCBSDRNc11Nm and qCBSDRNFc2Nm, respectively and a third putative QTL associated with CBSD root necrosis resistance. In addition, for the first time, two QTL which co-locates with the earlier identified CMD2 locus, namely qCMDc12.1A and qCMDc12.2A, have been identified. The QTL qCBSDRNc11Nm was found to be associated with resistance to CBSD root necrosis. The QTL qCBSDRNc11.2Nm on chromosome 11 associated with resistance to CBSD root necrosis had the highest LOD score of 7.50, explaining 17.3% of phenotypic variance. A number of candidate resistance genes were identified within these QTL regions, including two LRR genes and a gene encoding a signal recognition protein.

In a study conducted by Tappiban *et al.* in 2018, 12 quantitative trait loci (QTL) associated with CBB infection were identified in the F1 progenies of a cross between the BHuay Bong 60 and BHanatee cassava cultivars. Among all identified QTL, CBB14_10dai_1, CBB14_10dai_2, and CBB14_12dai showed the most significant ($p< 0.0001$) associations with CBB infection, and explained 21.3, 13.8, and 26.5% of phenotypic variation, respectively. Candidate genes underlying QTLs were identified and they were found to exhibit significant differences in expression between resistant and susceptible lines. Genes including Brassinosteroid-1 associated receptor kinase (Manes.04G059100), cyclic

nucleotide-gated ion channel 2 (Manes.02G051100), and autophagy-related protein 8a-related (Manes.17G026600), and regulator of nonsense transcripts 1 homolog (Manes.17G021900) was identified. The expression pattern of all genes showed higher levels in resistant (B82, B32, B20, and B70) as compared to susceptible (HB60, B100, B95, and B47) plants. Their results showed that the resistant plants expressed genes involved with signalling, ion channels induce ion influx, mRNA control, and protein degradation process at higher levels than susceptible plants.

2.6. META-ANALYSIS OF QTLs

Meta-analysis is a method of integrating information from several studies and analyzing it as a single one. QTL meta-analysis involves combining results of individual QTL studies and detecting major and constant QTLs. The pooling of results from several studies allows greater statistical power for QTL detection and a more precise estimation of their genetic effects. The QTLs identified by a meta-analysis from a stack of QTLs at a confidence interval of 95% are called meta-QTLs (MQTLs), which further require validation using a set of germplasm or breeding lines. The MQTLs, with small CI, consistency, and large effect on a trait are useful for marker-assisted selection (Islam *et al.*, 2019), in addition, a meta-analysis can assess the amount of variability between studies, identify study characteristics associated with particular QTLs (Baird *et al.,* 2008). Goffinet and Gerber proposed a meta-analysis strategy to resolve the number of QTL and their locations based on model selection. Their method provides a modified Akaike criterion that can be used to decide how many QTL are actually represented by the QTL detected in different experiments. This criterion is computed to choose between models with one, two, three, etc., QTLs. QTL meta-analysis reduces the confidence interval of QTL location and only actual or real QTL will be present. Comparative analysis of QTL between species reveals the existence of homologous QTL for plant height and maturity within the Poaceae (sorghum, maize, rice, wheat, and barley; Lin *et al.,* 1995). Comparing species is also a means to find new QTLs, increasing their potential for plant breeding as in the

tomato (Fulton *et al.*, 1997). Moreover, the existence of small common(QTL) on linkage maps between taxa that diverged a long time ago may provide the opportunity to extend results obtained in one species and permit the cross-utilization of resources that have been developed for a given species (Kowalski *et al.*, 1994).

In meta-analysis, only previously reported QTL information is required; candidate genes can be mined from the MQTLs (Meta - QTL) with small genetic and physical intervals. Integrated QTL analysis can facilitate the identification of real QTL and has attracted a great deal of attention. In a given species, the locations of QTLs for the same trait identified in one population generally correspond to those in another (Wang *et al.*, 2006). Thus, meta-analyses can be quite useful and beneficial for QTL analysis and candidate gene mining. However, the reliability of these results is not clear. The premise to gain reliable results are quality controls of data and a large enough total samples. For QTL meta-analysis, collecting data as much as possible is necessary as well as building a QTL mapping population, which may be useful for candidate gene identification in the follow-up study. Meta-analyses are performed to validate and refine confidence intervals of QTLs from those QTLs projected onto the reference map. The major, consistent MQTL detected can be used further for Marker assisted selection, to develop improved varieties.

Candidate gene identification through meta-analysis method is DigiCGA, (Digital Candidate Gene Approach) which also named in *silico* candidate gene approach or computer facilitated candidate gene approach, it is a novel web resource-based candidate gene identification approach. DigiCGA is as an approach that objectively extract, filter, (re)assemble, or (re)analyze all possible resources available derived from the public web databases mainly in accordance with the principles of biological ontology (e.g., anatomy ontology, cell & tissue ontology, developmental ontology, gene ontology, and phenotype & trait ontology) and complex statistical methods to make computational identification of

the potential candidate genes of specific interest, which is generally followed a subsequent validation of actual association analysis.

The meta-analysis approach was used to identify QTL in several plant species. Meta- analysis has been successfully done for QTLs associated with Fusarium Head Blight Resistance in Wheat (Liu *et al.,* 2009), Disease resistance traits of *Theobroma cacao* (Lanaud *et al.,* 2009), Silage quality of maize (Truntzler *et al.*, 2010), Grain yield in grass (Swamy *et al.*, 2011), Abiotic stresses in barley (Li *et al.*, 2013), Low-phosphorus tolerance in maize (Zhang *et al.*, 2014). A meta- QTL analysis of the 37 QTL was conducted for partial resistance to white mold common bean using the genetic linkage map of Stampede x Red Hawk population as the reference map by Vasconcellos *et al.,* in 2017. Nine meta-QTL with confidence intervals ranging from 0.65 to 9.41 Mb obtained by Meta-analysis. Candidate genes were shown to express under *Sclerotenia sclerotiorum* infection. Genes including cell wall receptor kinase, *COI1*, ethylene responsive transcription factor, peroxidases, and transcription factor, were found within the confidence interval for five of the meta-QTL.

Meta-analyses of QTL associated with protein and oil composition of seed in soybean by Van *et al.,* in 2017 yielded 55 MQTLs. Sixty nine functional candidates including, transcription factors and genes related to metabolic activities that may influence seed protein and oil biosynthesis and accumulation were identified within these identified 55 MQTLs.

The Meta-QTL analysis was conducted for a few salt tolerance-related traits in rice by Islam *et al.,* in 2018. 11 meta-QTLs on chromosomes 1 and 2 were identified. Candidates genes like cytochrome P450, expressed protein, hypothetical protein, zinc finger protein, protein kinase, protein phosphatase 2C, receptor-like protein kinase precursor, transcription factor-like protein, phytosulfokine receptor, transporter-like protein, transferases like protein, F-box, hydrolase-like protein, leucine zipper protein, and DNA domain containing protein was detected within this MQTL.

Meta-QTL analysis of seed iron and zinc concentration and content in common bean by Izquierdo *et al.,* 2018 resulted in twelve meta-QTL for seed Fe and Zn concentration. This MQTL was obtained from 87 QTLs originating from seven population grown in 16 field trials. This meta-QTL included 2 specific to iron, 2 specific to zinc and 8 that co-localize for iron and zinc concentrations and within five of the meta-QTL, candidate genes were identified from six gene families that have been associated with the transport of iron and zinc in plants.

2.7. SOFTWARE USED FOR META-ANALYSIS OF QTL

QTL meta-analysis is made easy by developing automated software. Biomercator is the major software used for meta-analysis. Several studies have been done using this software including Meta-Analysis of 100-Seed Weight QTLs in Soybean (Qi *et al.*, 2011), Meta-analysis of QTL associated with tolerance to abiotic stresses in barley (Li *et al.,* 2013), Meta–QTL analysis of disease resistance traits of *Theobroma cacao* (Lanaud *et al*., 2009) etc.

2.7.1.Biomercator :

(http://moulon.inra.fr/index.php/fr/seminairedoc/doc_download/64-biomercator-v4)

BioMercator is an automated map compilation and visualization tool that has been used for QTL meta-analysis to integrate information from multiple QTL studies, detect MQTLs and estimate their genetic effects precisely (Arcade *et al.,* 2004; Sosnowski *et al,* 2012). A number of studies have reported on QTL meta-analysis using BioMercator software (Delourme *et al.,* 2013). It has several versions like V2, V3, V4. BioMercator V4 is a genetic map compilation and QTL meta-analysis software providing tools to integrate the data with genome structural and functional annotation. It offers wizards to run analyses and display input and output maps through a user-friendly graphical user interface (GUI). It is a unique software and database to mine, integrate, display and query QTLs and meta-QTLs altogether with genome structural and functional annotation (Arcade *et al.,* 2004).

2.8. FUNCTIONAL ANNOTATION DATABASE

2.8.1. Phytozome (http://www.phytozome.net)

Phytozome is a comparative hub for plant genome and gene family data and analysis. Currently, it provides information on about 25 plant families. BLAST and BLAT searches of organism genomes and BLAST searches of proteomes and gene family consensus sequences can be used to find the genomic regions, gene transcripts, peptides and gene families most similar to a given query sequence. All gene and gene family attributes such as names, symbols, synonyms, external database identifiers, and functional annotation IDs are searchable (Goodstein *et al.*, 2011).

MATERIALS AND METHODS

3. MATERIALS AND METHODS

The study entitled "Meta-analysis of QTLs associated with pest and disease resistance genes in cassava" was conducted at the Central Tuber Crop Research Institute during 2018-2019. In this chapter, details regarding the experimental materials used and methodology adopted are disclosed.

3.1 DATA COLLECTION AND INPUT FILE PREPARATION

3.1.1. QTLs and Map files

Six published QTL mapping studies from 2004 to 2017 for disease resistance in cassava were selected (Table 1), in which detailed information of the genetic maps, including parents, types of mapping population, population size, number and type of marker and genetic distances in the linkage groups was available. A total of 61 QTLs were used in this study (Lokko *et al.*,2004, Rabbi *et al.*,2014, Vidhya *et al.*,2017, Mohan *et al.*,2013, Masumba *et al.*,2017, Kulembeka *et al.*,2010) in which 39 QTLs were associated with resistance to Cassava Mosaic Disease (CMD) and 22 QTLs were associated with resistance to Cassava Brown Streak Disease (CBSD). For each QTL, name, trait, experiment location, year of experiment, names of chromosome and linkage group, LOD score, R_2 value (proportion of phenotypic variance explained), most likely position of QTL (in cM) and confidence interval (CI, in cM) was collected. Input data files in XML format were prepared from each study according to the instruction manual of BioMercator V.4.2. The input files consists information about QTLs and their linkage maps. The genetic linkage map files consist the information about the map name, mapping cross type, cross size, genus name, species name, mapping unit, chromosome, linkage group, marker name and marker location. The QTL files consists the details about QTL name, trait name, Position in cM, Confidence Interval (cM), phenotypic variance and LOD score.

Table1. The QTL mapping studies used for meta-QTL analysis for QTLs associated with disease resistance in Cassava.

No.	Traits[a]	No.of QTLs	Parents	Location	Year of experiment	Population type	Population size	Markers used	Analysis method[b]	References
1	CMD	5	TMSI30572×TME17	Nigeria	2003	F1	132	SSR	SMA	Lokko *et al.*, 2004
2	CMD	2	ITA-TMS-011412× IITA-TMS-41425	Nigeria	2011&2012	F1	180	SNP	CIM	Rabbi *et al.*, 2014
3	CMD	2	MNga-1 × CI-732	ICAR-CTCRI, TRIVANDRUM	2017	F1	114	SSR	MIM	Vidhya *et al.*, 2017
4	CMD	2	MNga×CO2	TNAU, Coimbatore	2004	F1	300	SSR	SMA	Mohan *et al.*, 2013
5	CMD	28	Namikonga × Albert	Tanzania	2013&2014	F1	252	SNP	IM&ICIM	Masumba *et al.*, 2017
6	CBSD-RN	9	Namikonga × Albert	Tanzania	2013&2014	F1	252	SNP	IM&ICIM	Masumba *et al.*, 2017
7	CBSD - RN	13	Namikonga × Albert	Tanzania	2010	DH	60	SSR	SIM	Kulembeka *et al.*, 2010

[a] CMD, Cassava Mosaic Disease ; CBSD-RN, Cassava Brown Streak Disease-Root Necrosis

[b] SMA, Single Marker Analysis; CIM ,Composite Interval Mapping; MIM ,Multiple Interval Mapping; IM, Interval Mapping ;ICIM, Inclusive Composite Interval Mapping

38

3.1.2. Reference Map

A composite genetic linkage map for cassava harbouring 22,403 markers with map length of 2,412cM from International Cassava Genetic Map Consortium (ICGMC) (2014) developed by International Institute of Tropical Agriculture (IITA) was used as reference map for anchoring all individual genetic linkage maps used in this study. This composite map consists of 18 LGs with a marker density of 1.95 non-redundant markers per cM. The composite map was constructed by merging 10 maps (one S1 and nine F1 populations) derived from diverse parents from African cassava breeding projects. Markers were generated via Genotyping by Sequencing (GBS) and a map was constructed from each of the ten crosses with JoinMap and these maps were merged with LP merge to generate a integrated genetic linkage map of Cassava. About 71.9% of the genome assembly of cassava was anchored to this genetic map. The mapping populations used for constructing composite map were associated with traits such as resistance to Cassava Brown Streak Disease (CBSD) and green mite, Cassava Mosaic Disease (CMD) resistance, Traits related to Starch, dry matter content and Root carotenoid content.

3.2. DATA SUBMISSION TO BIOMERCATOR V.4.2

BioMercator V4.2 is a genetic map compilation and QTL meta-analysis software. It is used to integrate information from multiple QTL studies, it detects MQTLs and estimate their genetic effects precisely. BioMercator enables compilation and visualization of a large number of genetic maps, from different sources (literature, database, own experiment, etc.). Meta-analysis is run at the chromosome level and QTLs from related traits can be jointly subjected to a single meta-analysis by grouping corresponding traits into meta-traits. Biomercator V.4.2. was downloaded from URGI platform and is installed. A new project named "Cassava" was created in this software and the input files including

genetic map file, QTL file and reference map file were loaded to this project for meta-analysis. The workflow is given in Figure 1.

3.3. PROJECTION OF INDIVIDUAL MAPS TO REFERENCE MAP AND CONSTRUCTION OF CONSENSUS MAP USING CONSMAP TOOL

The ICGMC map was used as reference map and linkage maps from 6 studies were projected to this reference map .The consensus map construction is based on the markers shared by QTL maps and this reference map, this analysis was done in single one step.

3.3.1 Method:

1. On the "Analysis" option in menu "Map compilation" was selected
2. Within this map compilation "ConsMap" option was selected.
3. On the dialogue box appeared the input maps i.e. all the individual maps from the project "Cassava" were selected and the ICGMC map was given as reference map.
4. Then the analysis on the name " CONSENSUS_MAP " was launched.
5. After the analysis the resulting consensus map is visualized on the drawing panel.

3.4 QTL PROJECTION TO CONSENSUS MAP USING 'QTL PROJ'

The 61 QTLs identified QTLs from the 6 studies were projected onto the consensus map for meta-analysis based on the position, LOD score, Confidence Interval CI and Phenotypic Variance (R_2) explained by each of the QTL. This computation is based only on loci position data. QTLs are projected based on map projection algorithm of Biomercator V.4.2. The algorithm is such that for each pair of homologous chromosomes, common loci are listed. A specific distance ratio is then computed for each interval between two common loci.

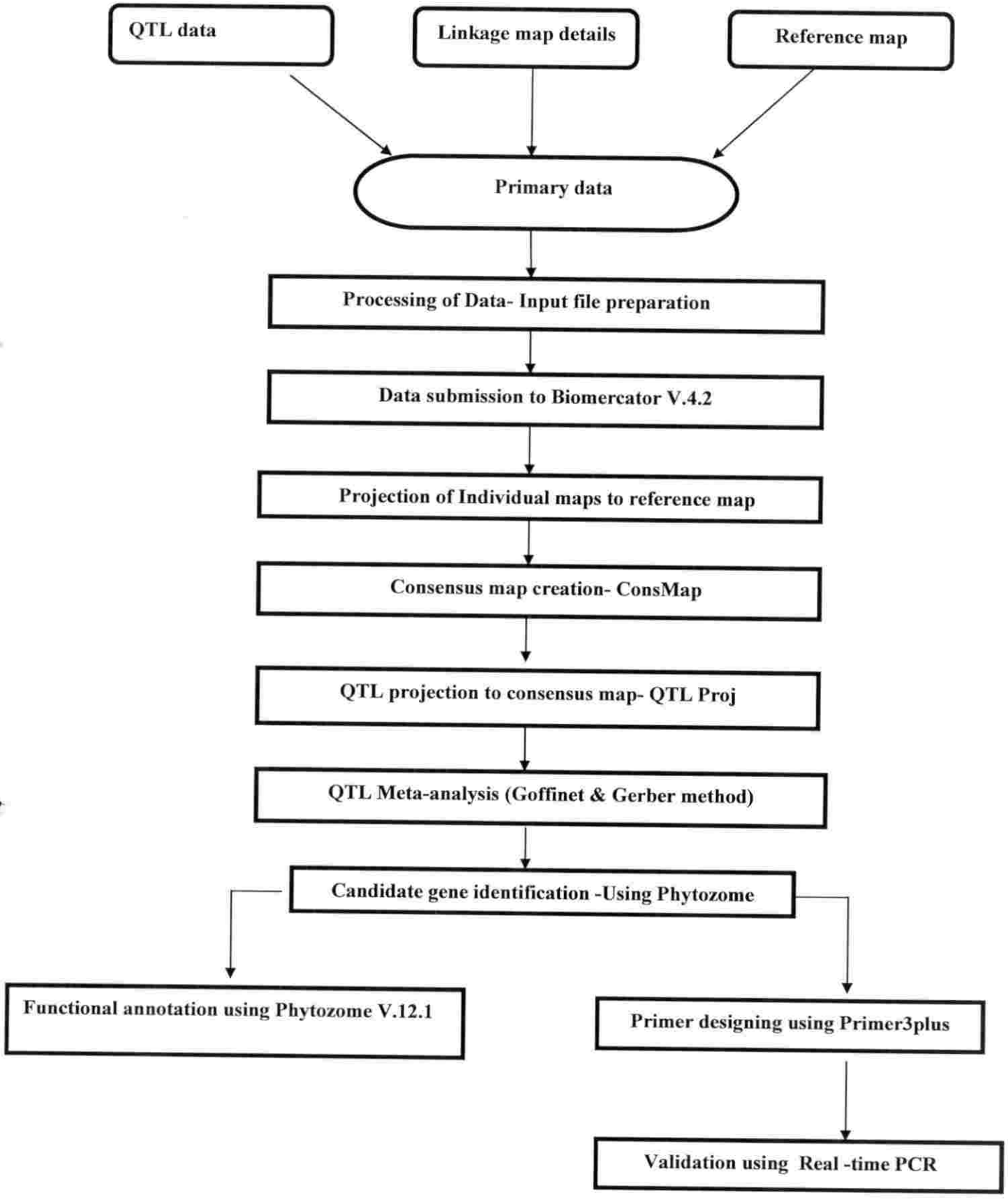

Figure1. The workflow for Meta-analysis of QTLs associated with Pest and Disease resistance genes in Cassava

Inverted pairs of loci can be automatically discarded from the list of common loci. Finally, QTL and/or remaining loci positions on the target map are computed by application of the appropriate distance ratio through a homothetic projection process.

3.4. Method:

1. From the "Analysis" option "Map compilations" was selected.
2. Within map compilation "QTLProj" option was selected.
3. In the appearing next window, in the left explorer, all maps from the "Cassava" project were selected.
4. In the right explorer, the "CONSENSUS_MAP" from the "Cassava" project was selected. Default parameters for ratio and p-value was given
5. 0.25 was applied as the minimum value of the ratio of the flanking marker interval distance.
6. 0.5 was applied as the minimum p-value threshold for testing homogeneity of the flanking marker interval distances between original and consensus maps.
7. The resulting map name was set as "QTL PROJ" and the analysis was launched.

3.5. META-ANALYSIS OF QTLS

Meta-analyses were performed to validate and refine confidence intervals of QTLs from those QTLs projected onto the reference map. Meta-analysis computing is based on the position of the QTL and on the variance of this position, which is assessed through confidence interval values [deduced from LOD curves, or from population size and proportion of variance explained by the QTL] . QTL meta-analysis using Goffinet and Gerber (2000) method was done here.This meta-analysis is a two steps analysis; the first one calculates and estimates the best models (i.e. number of Meta QTLs), so we can choose the one

to show in the second. This meta-analysis algorithm developed by Goffinet and Gerber can help to determine if N QTLs linked to a same trait or related ones, detected in independent experiments and located in a same region, are consistent with 1-, 2-, 3-, 4- or N- QTL models (N-QTLs model being the case where there are as much "real" QTLs as input QTLs). For each of those five models, the most likely QTL arrangement, assuming a Gaussian distribution, is determined by means of the maximum likelihood method. Then, an Akaike-type statistical criterion indicates the best model among the five ones. For each model, consensus QTL positions are determined as the mean of QTL distribution maximizing the likelihood. The lowest AIC value was used to select the best QTL model for each chromosome, which was considered significant to identify a number of meta-QTLs. QTL presented by the optimum model are regarded as the Meta-QTL (MQTLs).

3.5.1. Method:

1. From the "Analyses" option "QTL Meta-analysis" was selected.
2. From QTL Meta-analysis, Meta-analysis (Gerber & Goffinet) was selected.
3. In the appearing next window, the Meta-analysis name was given, Project "Cassava" selected, Map "QTL PROJ" was selected.
4. From the "QTL PROJ" map Chromosomes and linkage groups were selected.
5. Meta-analysis was done with QTLs related to CMD and CBSD resistance separately.
6. From the 'QTL choice' option QTLs associated with CMD/CBSD were selected and Meta-analysis was done.
7. From the Meta-analysis results the model with lowest AIC criterion was selected for visualization.
8. From the MQTLs obtained their position, Flanking markers and Confidence Interval (CI) were identified.

3.6. IDENTIFICATION OF CANDIDATE GENES

The search for candidate genes was performed based on the physical positions of the Meta-QTL regions. The MQTLs with less confidence interval were selected. In order to identify the positional candidate genes (CG; i.e., closely linked genes localized within QTL regions), genetic and physical map integration was performed based on the positions of markers defining the boundaries of QTL confidence intervals (or the closest to them) in the genome. The most recent annotated version of the *Manihot esculenta* (Cassava) reference genome V.6.1 in Phytozome was used to identify the physical position of the meta-QTL and genes contained in these regions. The physical position (in bp) of flanking markers were used to identify the genomic region of MQTL. From Phytozome database the nucleotide sequence on Cassava genome V.6.1. assembly (Whole genome shortgun sequence) present in between these flanking markers was obtained. From the Phytozome Cassava genome V.6.1.assembly, the candidate genes present in between the base pair regions were identified.

3.7 FUNCTIONAL ANNOTATION OF CANDIDATE GENES USING PHYTOMINE

Functional annotation of the retrieved candidate genes was done using PHYTOMINE. Phytomine is an intermine interface to data from phytozome. From phytomine detailed functional informations of genes including descriptions for PFAM, KEGG, Gene Ontology, PANTHER, and KOG classifications assigned to this gene will be obtained. Gene ID of each candidate genes were used to search against Phytomine and the detailed functional annotations of genes were obtained.

3.8 EXPERIMENTAL VALIDATION

The experimental validations of MQTLs identified were conducted by randomly choosing two QTLs associated with CMD resistance. Primer sequences specific to candidate genes associated with these QTLs were synthesized. RT-qPCR were performed as described below using total RNA isolated from cassava leaves samples of resistant and susceptible varieties of cassava available at ICAR-CTCRI.

3.8.1 Selected Cassava varieties

- MNga-1 : CMD Resistant variety
- H-165 : CMD Susceptible variety

3.8.2 Primer designing using Primer3plus

Primer3plus is a widely used program for designing Polymerase Chain Reaction (PCR) primers. Primer3 can also design hybridization probes and sequencing primers. Primer3 picks primers for PCR reactions, Considering certain criteria such as oligonucleotide melting temperature, size, GC content, Primer dimer possibilities, PCR product size, Positional constraints within the source (template) sequence, possibilities for ectopic priming (amplifying the wrong sequence) and many other constraints. The parameters considered in primer designing are:

Primer Length:

It is generally accepted that the optimal length of PCR primers is 18-22bp. This length is long enough for adequate specificity and short enough to bind easily to the template at the annealing temperature.

Primer Melting Temperature:

Primer Melting Temperature (T_m) by definition is the temperature at which one half of the DNA duplex will dissociate to become single stranded and indicates the duplex stability. Primers with melting temperatures in the range of 52-58^{0}C generally produce the best results.

GC Content:

The GC content (The number of G's and C's in the primer as a percentage of total bases) of primer should be 40-60%.

GC Clamp:

The presence of G or C bases within the last five bases from the 3' end of the primers (GC clamp) helps promote specific binding at the 3'end due to the stronger bonding of G and C bases. More than 3 G's or C's should be avoided in the last 5 bases at the 3'end of the primer.

3.8.3. RNA isolation

RNA was extracted from young leaves of resistant variety (MNga) and susceptible variety (H-165) of Cassava available at ICAR-CTCRI using RNeasy plant mini kit in accordance with manufacturer's protocol, TRI reagent method and Lithium chloride method.

$LiCl_2$ method: 100 mg leaf grinded with liquid nitrogen and 1mL of CTAB buffer (pre-warmed at 65°C for 10 minutes) was added to it. The extract is transferred to a fresh centrifuge tube and is centrifuged at 15000 rpm for 15 minutes. The supernatant was transferred to fresh tube and equal volume of chloroform: isoamyl (24:1) alcohol was added and centrifuged at 20,000g for 10 minutes at 4°C. The Chroroform:Isoamyl step was repeated . The upper layer was transferred to fresh tube and 0.25V of ice cold 10M lithium chloride was added, mixed well and incubated overnight at 20°C. On the next day it was centrifuged at 30,000g for 30 minutes at 4°C. The pellet was washed with 75% ethanol by centrifuging at 10,000 rpm at 4°C for 10 minutes. The washing step was repeated. RNA pellet was then dried at 37°C for 30 minutes and dissolved in 20µL RNase free water. The isolated RNA was stored at -80°C.

3.8.4. RNA quantification

The concentration of RNA was determined by using a Nano-drop (using 1 OD260=40µg RNA). A 260/ A280 ratios were also calculated for each sample.

3.8.5. cDNA synthesis

cDNA from the isolated RNA was prepared using RevertAid First Strand cDNA Synthesis Kit of Thermofisher scientific in accordance with manufacturer's protocol.

3.8.6. RT- qPCR

Real-time quantitative polymerase chain reaction (RT-qPCR) is a sensitive technique for gene expression studies. The qPCR reaction was performed with forward and reverse primer (specific to annotated candidate genes associated with QTLs) using cDNA samples from the young leaves of resistant and susceptible variety of cassava collected from ICAR-CTCRI.

Table 2. RT-qPCR reaction profile

Components	Volume (µl)
Diluted cDNA	1.5 µl
Forward primer (F)	1 µl
Reverse primer (R)	1 µl
2X H-eff qPCR master mix, Rox (GCC Biotech)	5 µl
Double distilled water	1.5 µl

3.8.6.1. Thermal profile

Initial denaturation - 95°C

Denaturation - 95°C

Annealing - 55°C

Extension - 72°C

Number of cycles - 35-45 cycles, step 2-4

After the completion of the real time PCR reactions, the threshold cycle (Ct) was recorded and gene expression level was calculated using comparative Ct method or delta-delta Ct method.

The relative gene expression level of resistant and susceptible varieties is represented as $2^{-\Delta\Delta CT}$ method.

ΔCt= Ct (target gene) - Ct (reference gene)

$\Delta\Delta Ct$= ΔCt (sample) - Δ Ct (Control)

RESULTS

4. RESULTS

The results of the study "META-ANALYSIS OF QTLS ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA" carried out at the Section of Extension and Social Sciences, ICAR-Central Tuber Crops Research Institute, Sreekariyam, Thiruvananthapuram during 2018-2019 are presented in this chapter.

4.1. DATA COLLECTION AND INPUT FILE PREPARATION

4.1.1 QTL and Map data

A total of six QTL mapping studies for disease resistance in cassava were used in this study which reported a total of 61 QTLs (Table 3), in which 39 QTLs was associated with resistance to Cassava Mosaic Disease (CMD), 22 QTLs was associated with resistance to Cassava Brown Streak Disease (CBSD). The mapping population size ranged from 60 to 300 lines. The markers used these studies includes SSR, EST-SSR and SNPs. These QTL studies have been conducted from different locations including African countries like Nigeria, Tanzania and at Indian institutes including Tamil Nadu Agricultural University (TNAU) and ICAR-Central Tuber Crops Research Institute, Thiruvanathapuram. These QTLs were distributed on 13 chromosomes with a range of 1-13% QTLs per chromosomes (Figure 2). Chromosomes 8 had the lowest number of QTLs and chromosome 12 had the highest number of QTLs. The distribution of QTLs showed that QTLs for CMD and CBSD resistance were located on 8 chromosomes i.e. on chromosomes 1, 2,5,11,12,16,17 and 18. Chromosomes 3, 7 and 10 had QTLs for CMD resistance, and chromosomes 4 and 8 had QTLs for CBSD resistance. The phenotypic variance explained by the initial QTLs varied from 0.031% to 50.37% and the confidence interval (CI) of markers varied from 1.5 to 60.4 cM. These QTLs has an average CI of about 9.56 cM .These QTLs had LOD scores between 1.76 and 45.59.

Table 3. List of QTLs used for meta-QTL analysis for QTLs associated with disease resistance in cassava

Sl. No.	QTL name	Trait[a]	Year	Location	Linkage group	LOD	Phenotypic variance (R^2)	Position (cM)	Confidence Interval (cM)	References
1	CMD-qtl	CMD	2003	Nigeria	17	3.0	10.47	19.9	14-40.1	Lokko *et al.*, 2004
2	CMD-qt2	CMD	2003	Nigeria	18	3.0	11.30	15.2	6-21.6	Lokko *et al.*, 2004
3	CMD-qt3	CMD	2003	Nigeria	11	3.0	12.15	10.9	0-12.8	Lokko *et al.*, 2004
4	CMD-qt4	CMD	2003	Nigeria	17	3.0	10.45	40.1	19.9-60.2	Lokko *et al.*, 2004
5	CMD-qt5	CMD	2003	Nigeria	11	3.0	11.55	6.3	0-10.9	Lokko *et al.*, 2004
6	CQTL1	CMD	2004	TNAU,Coimbatore	5	3.0	0.031	1.8	0.70-2.2	Mohan *et al.*, 2013
7	CQTL2	CMD	2004	TNAU,Coimbatore	11	3.0	0.037	50.18	47.6-51.19	Mohan *et al.*, 2013
8	CMD1S	CMD	2011&2012	Nigeria	16	45.59	0.708	70.00	68.875-70.74	Rabbi *et al.*, 2014
9	CMD3S	CMD	2011&2012	Nigeria	16	43.20	0.696	69.31	68.392 70.74	Rabbi *et al.*, 2014
10	qCMDc1-1A-I	CMD	2014	Tanzania	1	3.07	7.66	66	64.5-66.5	Masumba *et al.*, 2017
11	qCMDc1-2A-I	CMD	2014	Tanzania	1	3.64	7.97	72	70.5-73.5	Masumba *et al.*, 2017
12	qCMDc1-3A-I	CMD	2014	Tanzania	1	3.25	5.18	77	75.5-79.5	Masumba *et al.*, 2017
13	qCMDc2-1A	CMD	2013	Tanzania	2	3.68	7.03	31	29.5-31.5	Masumba *et al.*, 2017

14	qCMDc2-2A	CMD	2013	Tanzania	2	3.41	5.47	85	84.5-85.5	Masumba *et al.*, 2017
15	qCMDc3ACA	CMD	2014	Tanzania	3	7.92	14.95	65	61.5-70.5	Masumba *et al.*, 2017
16	qCMDc3ACB	CMD	2014	Tanzania	3	4.10	4.10	67	61.5-73.5	Masumba *et al.*, 2017
17	qCMDc5-1A-VN1	CMD	2014	Tanzania	5	3.70	11.17	40	39.5-40.5	Masumba *et al.*, 2017
18	qCMDc5-1A-VN2	CMD	2014	Tanzania	5	5.39	7.60	40	39.5-40.5	Masumba *et al.*, 2017
19	qCMDc5-2A-V	CMD	2013	Tanzania	5	3.11	4.44	54	52.5-58.5	Masumba *et al.*, 2017
20	qCMDc5-3A	CMD	2014	Tanzania	5	9.71	14.05	71	68.5-73.5	Masumba *et al.*, 2017
21	qCMDc10-1A	CMD	2014	Tanzania	10	3.12	6.56	1	0-1.5	Masumba *et al.*, 2017
22	qCMDc10-2A	CMD	2013	Tanzania	10	5.93	10.15	12	10.5-13.5	Masumba *et al.*, 2017
23	qCMDc10-3A	CMD	2014	Tanzania	10	4.30	17.35	47	44.5-48.5	Masumba *et al.*, 2017
24	qCMDc10-3AN2	CMD	2014	Tanzania	10	3.53	23.90	48	46.5-49.5	Masumba *et al.*, 2017
25	qCMDc10-4A	CMD	2013	Tanzania	10	3.57	5.93	52	50.5-52.5	Masumba *et al.*, 2017
26	qCMDc12-1AN1	CMD	2013	Tanzania	12	12.68	23.84	39	35.5-42.5	Masumba *et al.*, 2017
27	qCMDc12-1AN2	CMD	2013	Tanzania	12	15.92	27.01	41	38.5-44.5	Masumba *et al.*, 2017
28	qCMDc12-1AC1	CMD	2013	Tanzania	12	9.39	15	39	35.5-45.5	Masumba *et al.*, 2017

29	qCMDc12-1AC2	CMD	2013	Tanzania	12	15.29	27.33	39	35.5-43.5	Masumba *et al.*, 2017
30	qCMDc12-1ACA	CMD	2013	Tanzania	12	11.90	19.01	40	34.5-44.5	Masumba *et al.*, 2017
31	qCMDc12-2AN1	CMD	2013	Tanzania	12	10.83	20.38	38	33.5-41.5	Masumba *et al.*, 2017
32	qCMDc12-2AN2	CMD	2013	Tanzania	12	8.20	16.32	35	32.5-44.5	Masumba *et al.*, 2017
33	qCMDc12-2AC1	CMD	2013	Tanzania	12	10.95	14.32	38	33.5-42.5	Masumba *et al.*, 2017
34	qCMDc12-2AC2	CMD	2013	Tanzania	12	4.90	4.46	38	32.5-48.5	Masumba *et al.*, 2017
35	qCMDc12-2ANA	CMD	2013	Tanzania	12	10.32	20.61	38	31.5-45.5	Masumba *et al.*, 2017
36	qCMDc12-2ANB	CMD	2013	Tanzania	12	6.03	14.04	37	32.5-37.5	Masumba *et al.*, 2017
37	qCMDc12-3ACB	CMD	2014	Tanzania	12	16.00	50.37	56	54.5-56.5	Masumba *et al.*, 2017
38	QTL1	CMD	2015	ICAR-CTCRI, Trivandrum	7	5.32	12.8	0.1001	0.00-60.4	Vidhya *et al.*, 2017
39	QTL2	CMD	2015	ICAR-CTCRI, Trivandrum	7	3.70	2.20	60.336	45.4-60.4	Vidhya *et al.*, 2017
40	cbsd-rn-n4-chz08	CBSD-RN	2008	Tanzania	4	1.76	12.10	15.32	0.00-26.4	Kulembeka *et al.*, 2010
41	cbsd_rn_c5_ch007	CBSD-RN	2007	Tanzania	5	2.61	17.40	70.81	62.0-72.2	Kulembeka *et al.*, 2010
42	cbsd_rn_c5a_ch07	CBSD-	2007	Tanzania	5	2.63	17.50	72.23	70.8-77.6	Kulembeka *et al.*, 2010

		RN								
43	cbsd_rn_c5b_ch07	CBSD-RN	2007	Tanzania	5	2.69	17.80	77.63	70.8-88.3	Kulembeka *et al.*, 2010
44	cbsd-rn-chz08-qtl	CBSD-RN	2008	Tanzania	8	2.21	18.10	7.120	4.3-10.4	Kulembeka *et al.*, 2010
45	cbsd_rn_c16_ch07	CBSD-RN	2007	Tanzania	16	2.69	17.80	15.79	0.00-31.6	Kulembeka *et al.*, 2010
46	cbsd_rn-c18_ch07	CBSD-RN	2007	Tanzania	18	2.59	17.30	10.67	0.00-17.4	Kulembeka *et al.*, 2010
47	cbsd_rnc18a_ch07	CBSD-RN	2007	Tanzania	18	2.80	18.50	17.40	10.7-30.1	Kulembeka *et al.*, 2010
48	cbsd-rn-S1-chz080	CBSD-RN	2008	Tanzania	1	2.68	21.50	14.62	8.0-17.9	Kulembeka *et al.*, 2010
49	cbsd-rn-S4a-chz07	CBSD-RN	2007	Tanzania	4	2.33	19.00	7.880	2-60.2	Kulembeka *et al.*, 2010
50	cbsd-rn-S4b-chz07	CBSD-RN	2007	Tanzania	4	2.06	17.00	2.160	0.0-60.2	Kulembeka *et al.*, 2010
51	cbsd-rn-S12-chz08	CBSD-RN	2008	Tanzania	12	2.75	22.00	12.20	7.7-29.4	Kulembeka *et al.*, 2010
52	cbsd-rn-S17-ndl08	CBSD-RN	2008	Tanzania	17	2.10	17.90	8.530	6.6-9.3	Kulembeka *et al.*, 2010

53	qCBSDRNFc2Nm	CBSD-RN	2013	Tanzania	2	4.76	47.38	23	22.1-23.7	Masumba *et al.*, 2017
54	qCBSDRNFc2Nma	CBSD-RN	2014	Tanzania	2	3.24	42.46	23	22.1-23.7	Masumba *et al.*,2017
55	qCBSDRNc11-1Nm	CBSD-RN	2014	Tanzania	11	7.50	17.40	53	47.8-54.0	Masumba *et al.*, 2017
56	qCBSDRNc11-2Nm	CBSD-RN	2014	Tanzania	11	3.81	7.54	45	36.3-47.8	Masumba *et al.*, 2017
57	qCBSDRNc11-2Nma	CBSD-RN	2013	Tanzania	11	3.60	5.32	55	47.8-59.32	Masumba *et al.*, 2017
58	qCBSDRNc11-3Nm	CBSD-RN	2013	Tanzania	11	3.35	5.19	41	36.26-47.79	Masumba *et al.*, 2017
59	qCBSDRNc18-1Nm	CBSD-RN	2014	Tanzania	18	5.10	36.83	74	70.0-75.9	Masumba *et al.*, 2017
60	qCBSDRNc18-2Nm	CBSD-RN	2014	Tanzania	18	3.99	16.10	25	23.8-27.8	Masumba *et al.*, 2017
61	qCBSDRNc18-3Nm	CBSD-RN	2014	Tanzania	18	3.31	37.00	39	38.0-41.4 Average = 9.56	Masumba *et al.*, 2017

[a] CMD, Cassava Mosaic Disease ; CBSD-RN, Cassava Brown Streak Disease-Root Necrosis

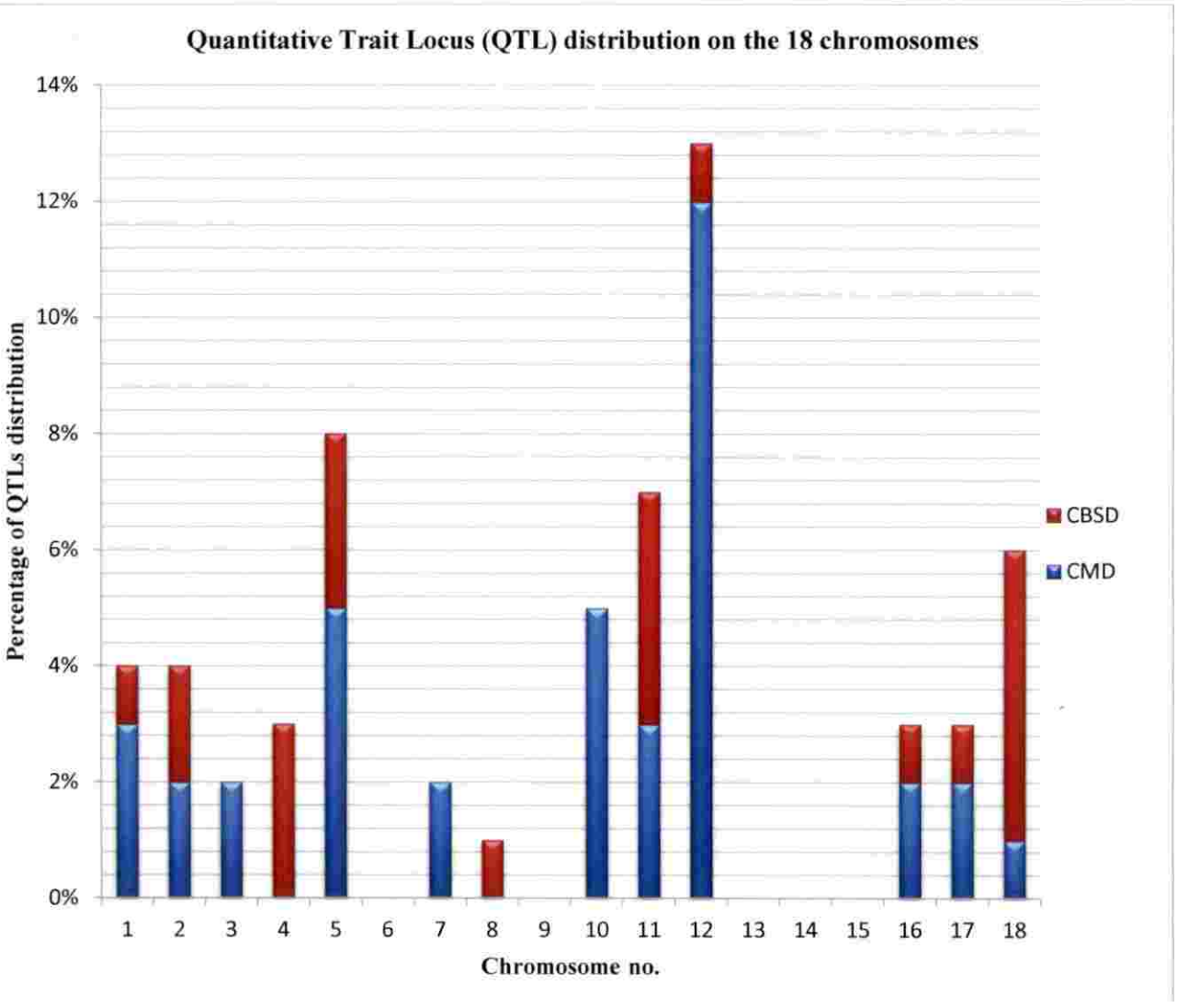

Figure 2. Distribution of QTLs associated with disease resistance genes in cassava on 18 chromosomes

CBSD- Cassava Brown Streak Disease, CMD- Cassava Mosaic Disease

4.1.2 Reference Map

International Cassava Genetic Map Consortium (ICGMC) (2014) with 22,403 SNP markers was used as reference map for anchoring all individual genetic linkage maps used in this study (Figure 3). This map consists of 18 LGs with map length of 2,412cM and marker density of 1.95 non-redundant markers per cM.

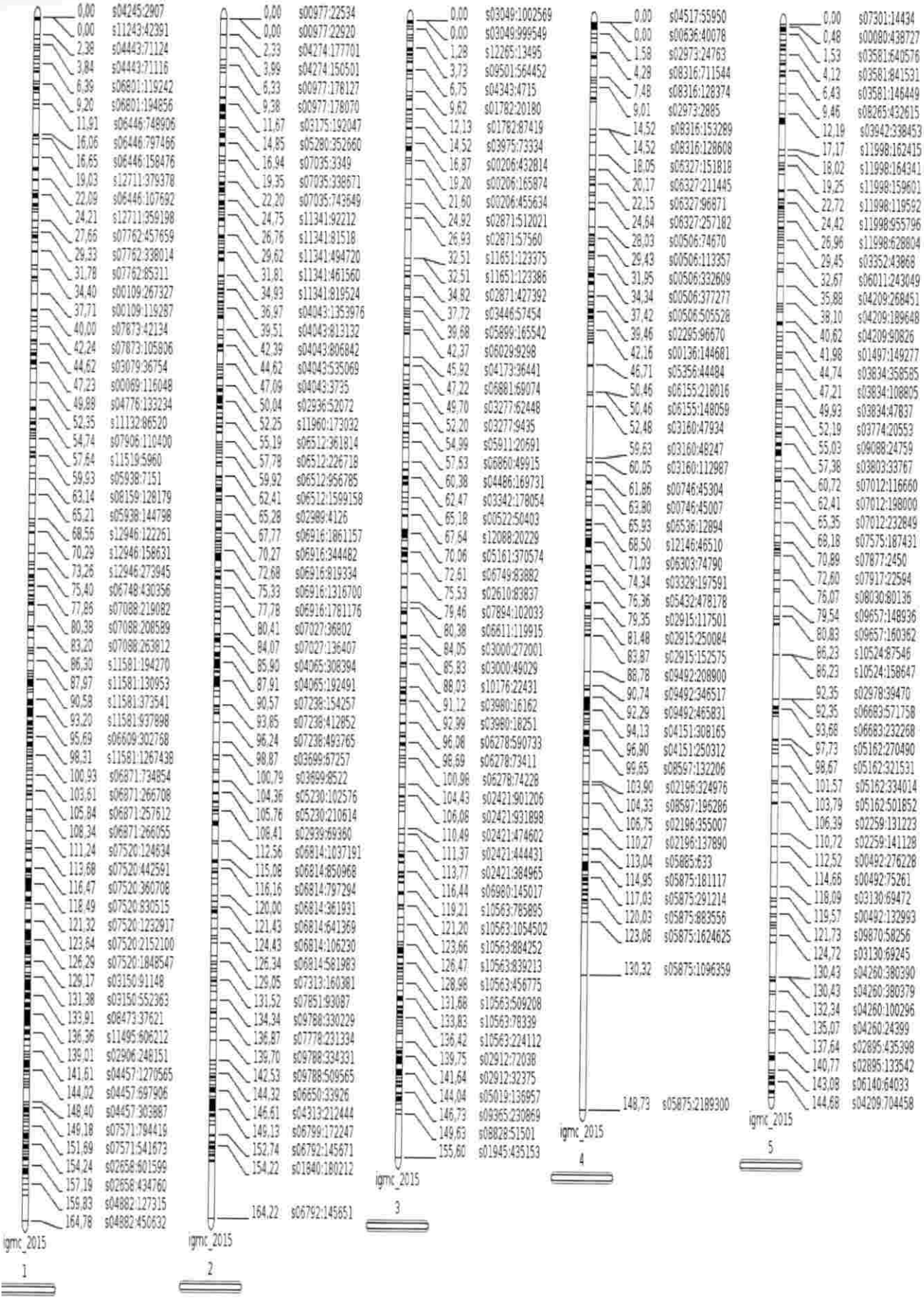

Figure 3. Reference Map used for this study- Integrated Cassava Genetic Map Consortium (ICGMC_2015)

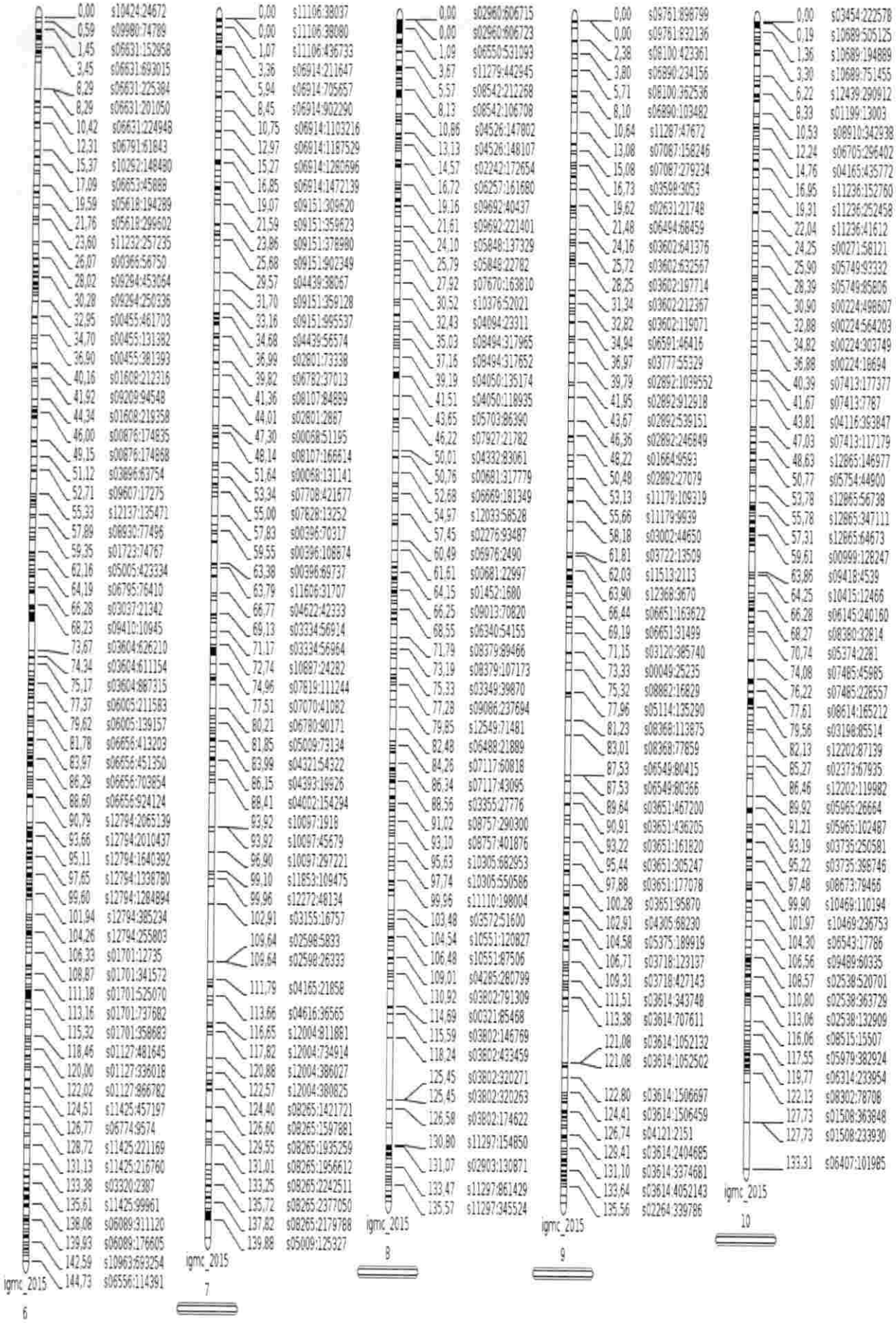

Figure 3. Reference Map used for this study- Integrated Cassava Genetic Map Consortium (ICGMC_2015)

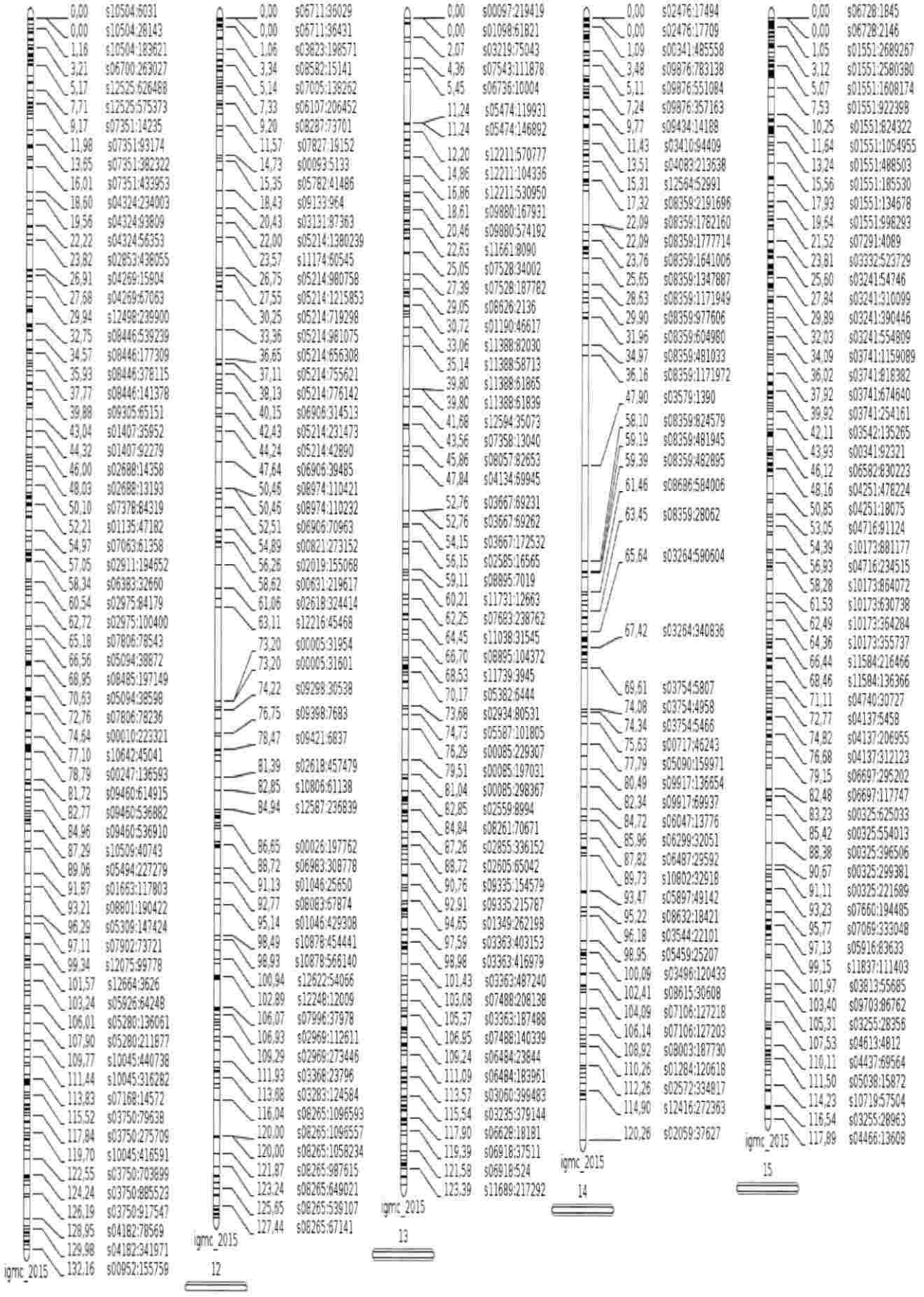

Figure 3. Reference Map used for this study- Integrated Cassava Genetic Map Consortium (ICGMC 2015)

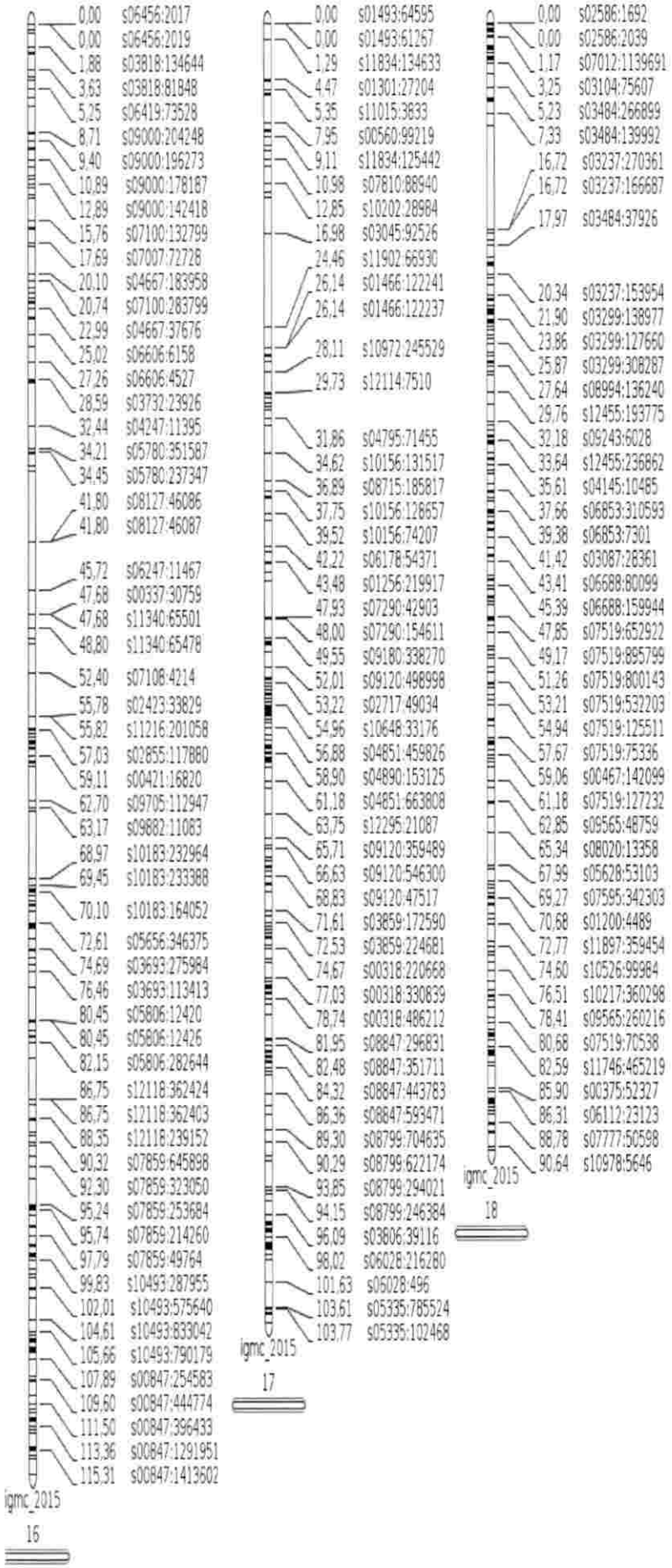

Figure 3. Reference Map used for this study- Integrated Cassava Genetic Map Consortium (ICGMC_2015)

4.2. DATA SUBMISSION TO BIOMERCATOR V.4.2

Biomercator V.4.2 was installed from URGI platform (https://urgi.versailles.inra.fr/Tools/BioMercator-V4). A new project in the name of 'CASSAVA' was created within the software. All the six QTL data, Map data and Reference map were loaded into this software.

4.3 PROJECTION OF INDIVIDUAL MAPS TO REFERENCE MAP AND CONSTRUCTION OF CONSENSUS MAP USING CONSMAP TOOL

A consensus map in the name of 'CMAP' was constructed using ICGMC map as reference map. The consensus map created consists of 18 LG with 22,621 markers (Figure 4). The length of Consensus map was 2,412.35cM. Markers from genetic linkage map of individual QTL studies were incorporated with reference map. In the consensus map chromosome 1 harbours more no. of markers i.e about 2,387 markers and Chromosome 7 has lesser no of markers i.e. about 310 markers. Map length of chromosome 1 was highest among all (164.78 cM) and chromosome 18 has the lowest map length (90.99 cM). The summary of consensus map is given on Table 4.

Table 4. Summary of integrated consensus map

Chromosome no.	**No. of markers**	**Length (cM)**
1	2387	164.78
2	1450	164.22
3	1368	155.60
4	1470	148.73
5	1392	146.87
6	1468	144.73
7	310	141.96
8	1217	137.48
9	1252	137.35
10	1221	133.31
11	1362	132.16
12	845	128.85
13	892	125.09
14	1362	120.26
15	1597	117.89
16	1127	117.13
17	837	104.95
18	1064	90.99
Total	**22,621**	**2,412.35**

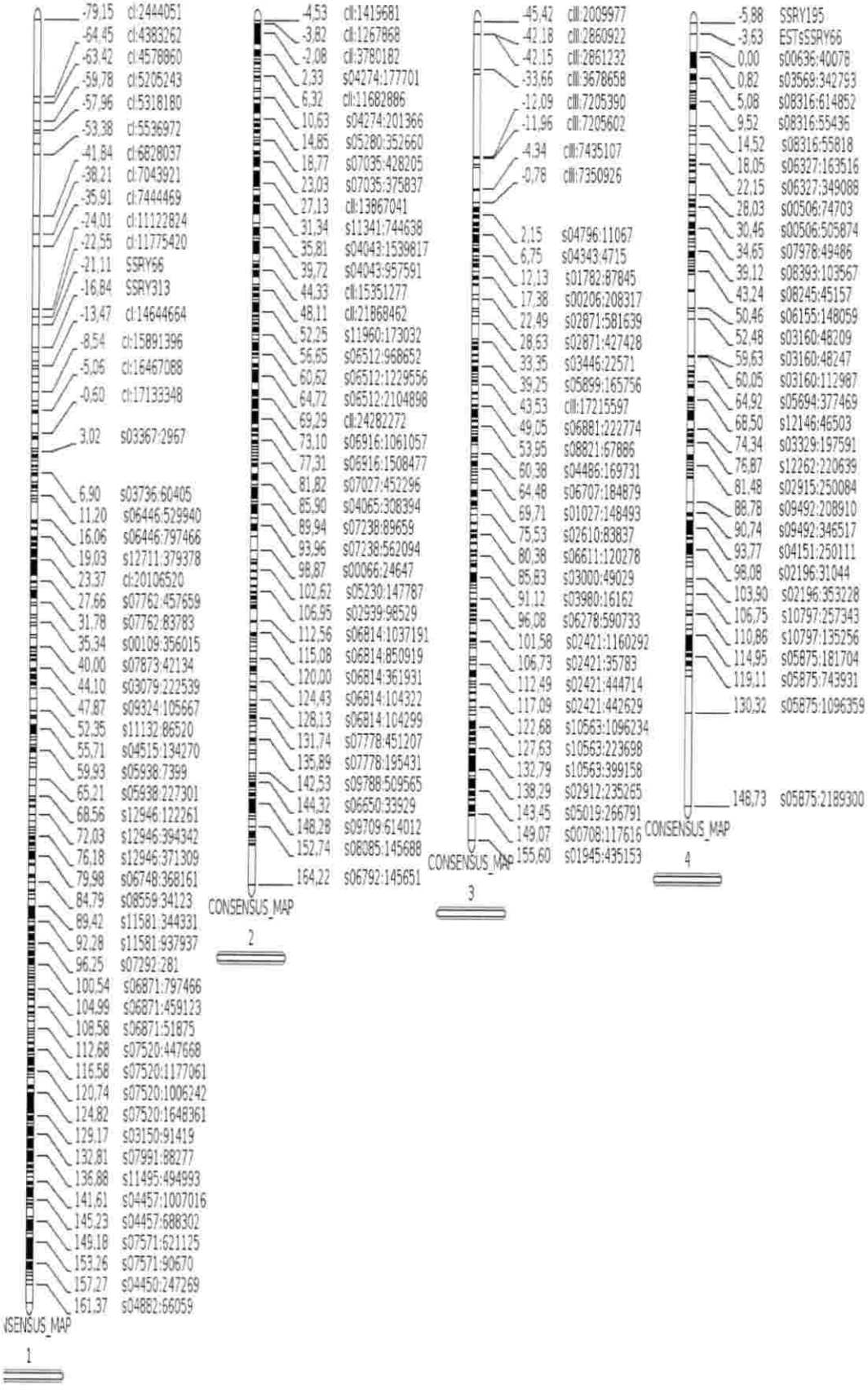

Figure 4. Consensus map created using Cons_Map tool

Marker name on the right of the chromosomes along with their distance in Centi Morgan

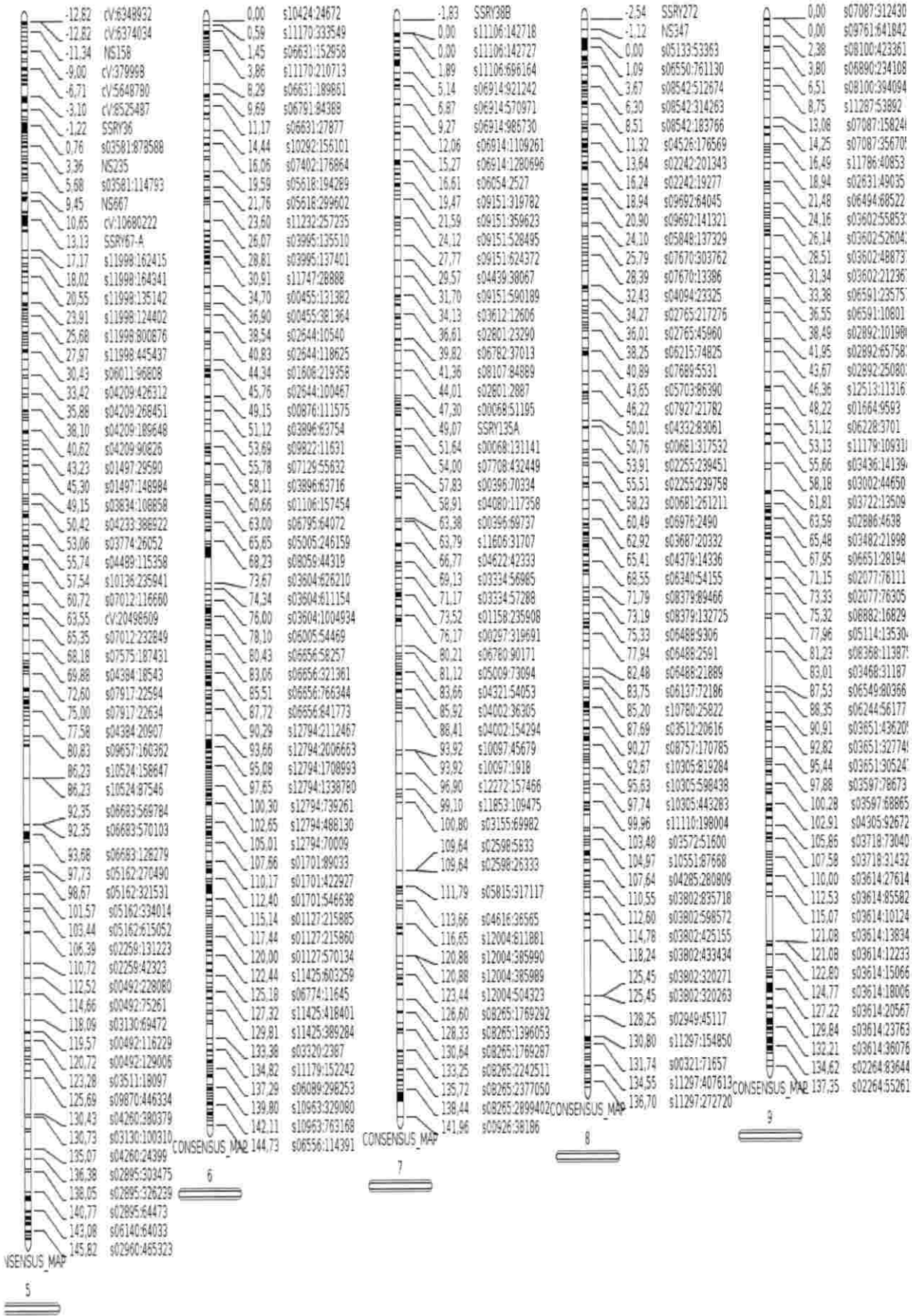

Figure 4. Consensus map created using Cons_Map tool

Marker name on the right of the chromosomes along with their distance in Centi Morgan

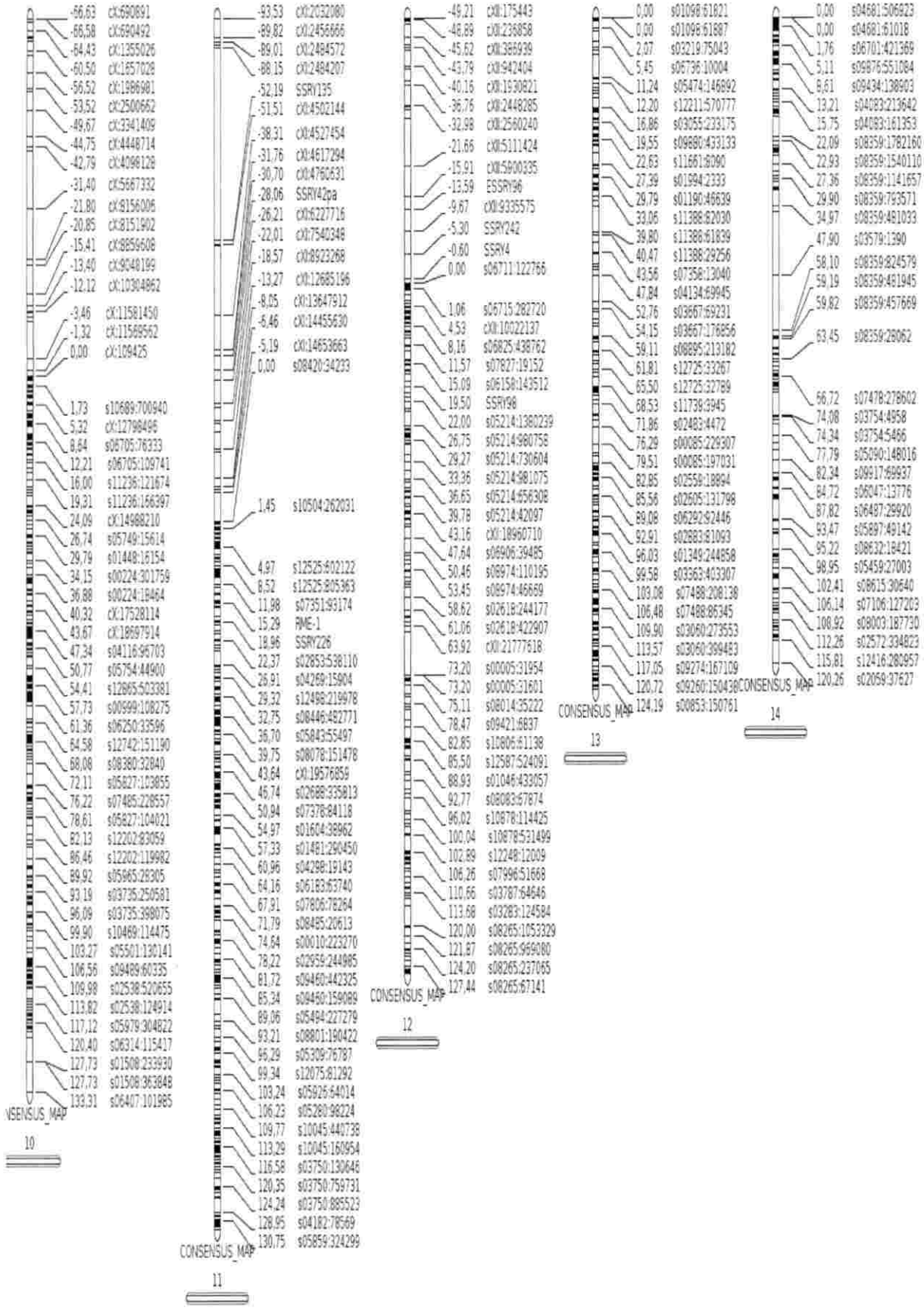

Figure 4. Consensus map created using Cons_Map tool

Marker name on the right of the chromosomes along with their distance in Centi Morgan

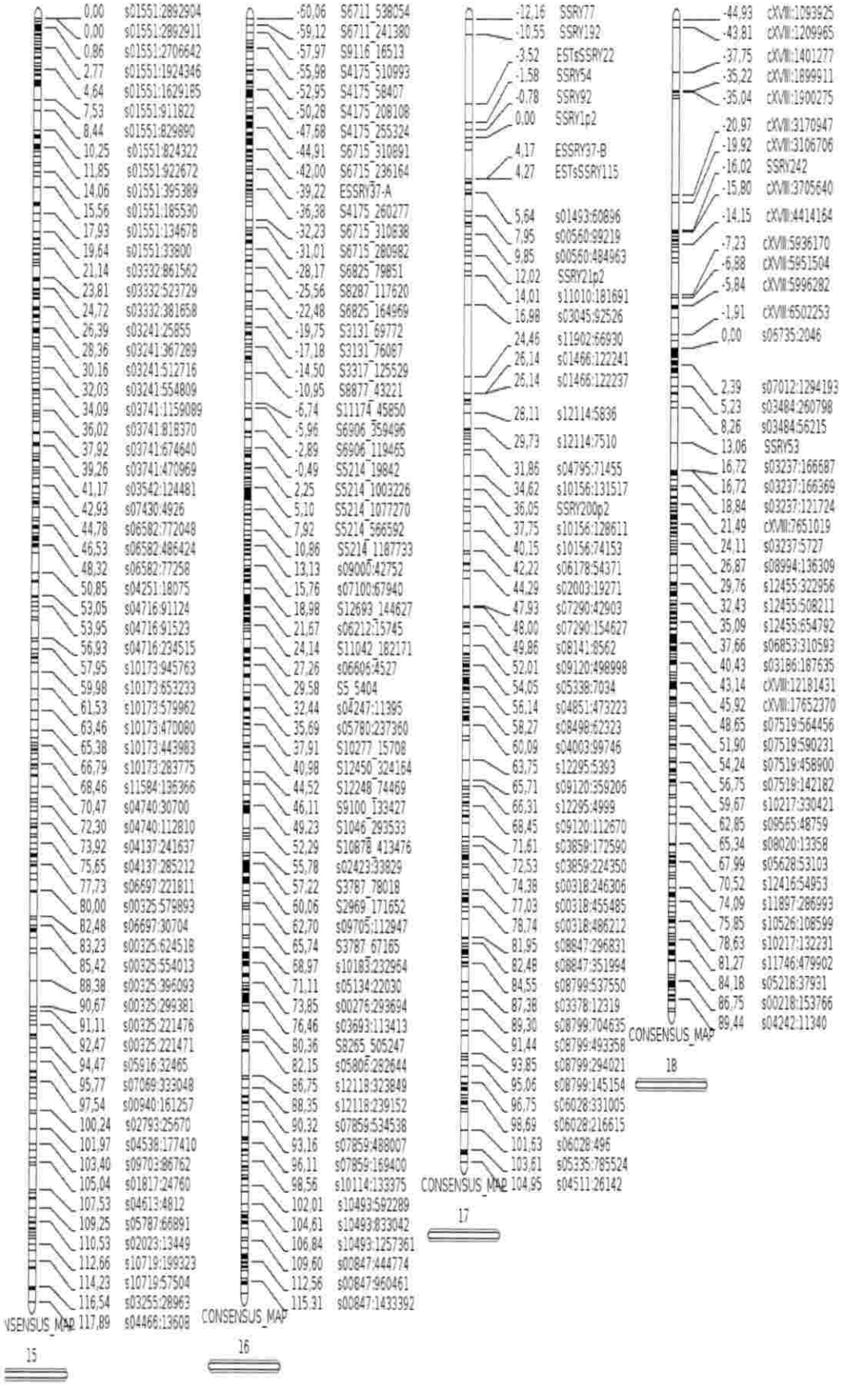

Figure 4. Consensus map created using Cons_Map tool

Marker name on the right of the chromosomes along with their distance in Centi Morgan

4.4 QTL PROJECTION TO CONSENSUS MAP USING 'QTL PROJ'

To the constructed consensus map 58 QTLs from 61 initial QTLs were projected (Figure 5). QTLs associated with CMD and CBSD were projected to 13 chromosomes in the consensus map. Chromosome 12 consists of more number of projected QTLs i.e. 13 QTLs and only 1 QTLs were projected on chromosome 4, 7 and 8. All initial QTLs within the chromosomes 1, 2, 3,5,8,10,11,12,16,17 and 18 were projected after QTL Proj analysis. In chromosome 4 only 1 QTL out of 3 initial QTLs was projected, and in chromosome 7, 1 QTL out of 2 initial QTLs was projected. The failure of projection of these QTLs was because of their confidence interval. The QTLs which consist of large confidence interval were not projected to the consensus map. The QTLs which are not projected on chromosome 4 are cbsd-rn-S4a-chz07 and cbsd-rn-S4b-chz07 and their Confidence interval was about 58.2 and 60.2 respectively. The QTL which is not projected on chromosome 7 was QTL1 and it's confidence interval was about 60.4. The summary of QTLs projected after QTL Proj analyses is given on table 5. In chromosomes 1, 2,5,11,12,16,17 and 18 QTLs associated with CMD and CBSD-RN resistance were projected. In chromosomes 5, 7 and 2 QTLs associated with CMD resistance were projected and in chromosomes 4 and 8 QTLs related to CBSD-RN resistance were projected. Details about projected QTLs are given on Table 6.

4.5 META-ANALYSIS OF QTLS

Meta-Analysis of 58 QTLs projected in the consensus map was performed in Biomercator V.4 software using Goffinet and Gerber method. Meta - analysis of CMD and CBSD-RN resistance QTLs were done separately. After Meta -analysis QTLs were fixed into 4 models, the model with lowest AIC criterion was selected as best model and it was used for visualization. 11 Meta-QTLs (MQTLs) were obtained after meta-analysis of CMD resistance QTLs (Figure 6) and 10 MQTLs were obtained for CBSD-RN resistance (Figure 7). The average CI of CMD

resistance MQTLs was found to be 3.877 whereas the average CI of projected CMD QTLs was 7.79, so CI of QTLs was reduced to half after meta-analysis. In the case of CBSD-RN resistance QTLs the CI of projected QTLs were about 11.36 and after meta- analysis it was reduced to 6.43.

CMD resistance MQTLs:

The MQTLs obtained for CMD resistance QTLs were MQTLCMD1.1, MQTLCMD2.1, MQTLCMD3.1, MQTLCMD5.1, MQTLCMD7.1, MQTLCMD10.1, MQTLCMD11.1, MQTLCMD12.1, MQTLCMD16.1, MQTLCMD17.1, MQTLCMD18.1 in the chromosomes 1,2,3,5,7,10,11,12,16,17, and 18 respectively. A total of 38 QTLs associated with CMD resistance were present within these MQTLs (Table 7). From the obtained MQTLs the QTLs with CI less than 2cM were selected for candidate gene mining. The QTLs with less genetic interval i.e. CI are potential targets for identifying genes. Nine QTLs were found to be within the CI of 2cM and they were selected for integrating with genomic information. One QTL each from MQTLCMD1.1 and MQTLCMD2.1 and MQTLCMD18.1, two QTLs each from MQTLCMD3.1, MQTLCMD5.1 and MQTLCMD16.1 are identified as QTLs within 2cM CI. From MQTLs MQTLCMD7.1, MQTLCMD10.1 MQTLCMD11.1, MQTLCMD12.1 and MQTLCMD17.1 no QTLs was found within the CI of 2cM so they were not selected.

CBSD resistance MQTLs:

The MQTLs obtained for CBSD-RN resistance QTLs were MQTLCBSD1.1, MQTLCBSD2.1, MQTLCBSD4.1, MQTLCBSD5.1, MQTLCBSD8.1, MQTLCBSD11.1, MQTLCBSD12.1, MQTLCBSD16.1, MQTLCBSD17.1, and MQTLCBSD18.1 in the chromosomes 1,2,4,5,8,11,12,16,17 and 18 respectively. A total of 20 QTLs were found within these MQTLs (Table 8). The QTLs within the CI of 2cM are selected for identifying candidate genes. Seven QTLs were found to be within less CI length and they were selected. From the selected 7 QTLs two were from MQTLCBSD2.1, one each from MQTLCBSD11.1 and MQTLCBSD17.1, and 3 QTLs fromMQTLCBSD18.1 .No QTLs were found within 2cM CI of all other MQTLs.

Table 5. Summary of QTLs projected after QTL Proj analysis

Chromosome no.	**Initial no. of QTLs**	**No. of QTLs projected**	**Traits[a]**
1	4	4	CMD & CBSD-RN
2	4	4	CMD & CBSD-RN
3	2	2	CMD
4	3	1	CBSD-RN
5	8	8	CMD & CBSD-RN
7	2	1	CMD
8	1	1	CBSD-RN
10	5	5	CMD
11	7	7	CMD & CBSD-RN
12	13	13	CMD & CBSD-RN
16	3	3	CMD & CBSD-RN
17	3	3	CMD & CBSD-RN
18	6	6	CMD & CBSD-RN
Total	**61**	**58**	

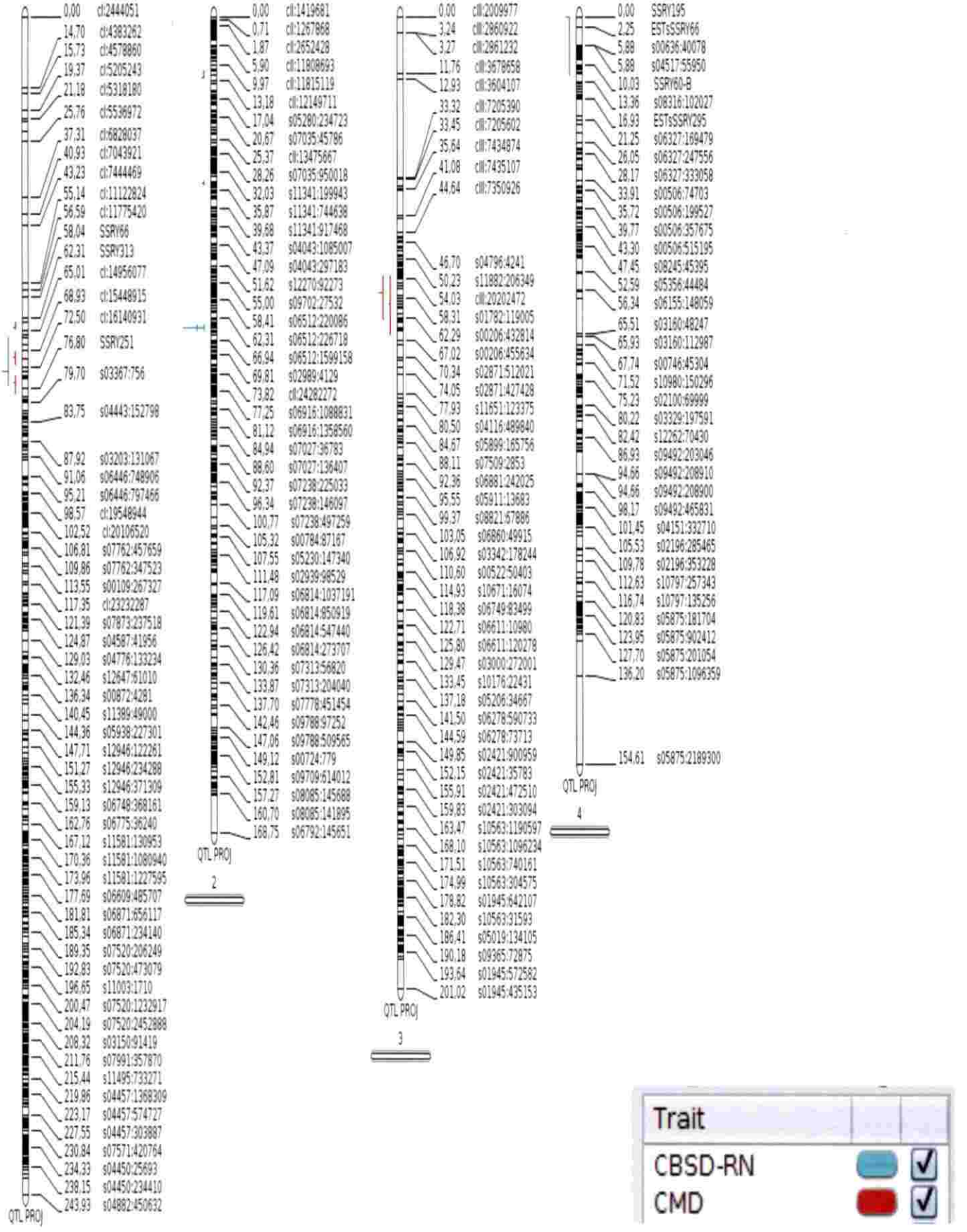

Figure 5. QTLs Projected on Consensus map after QTL Proj analysis

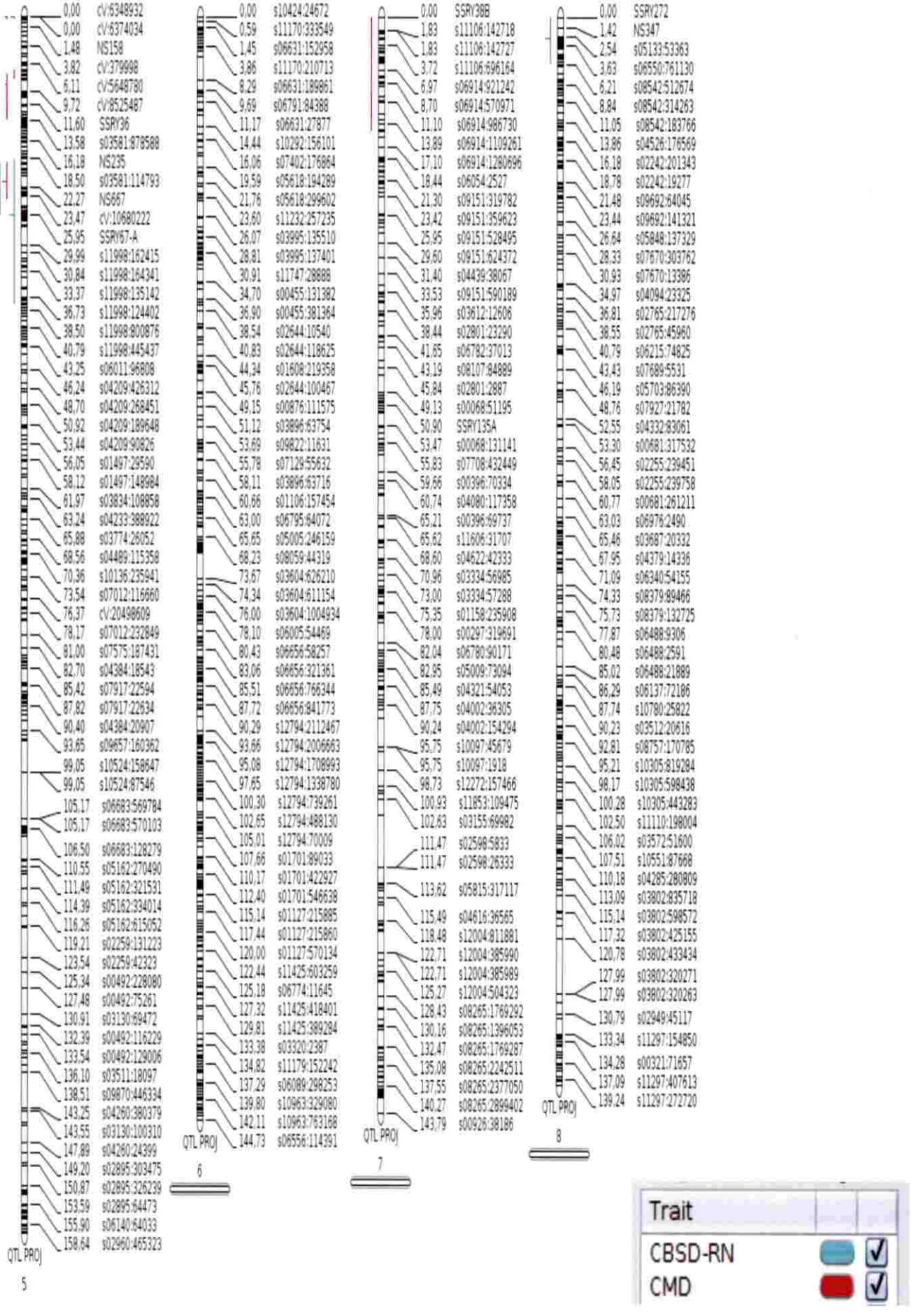

Figure 5. QTLs Projected on Consensus map after QTL Proj analysis

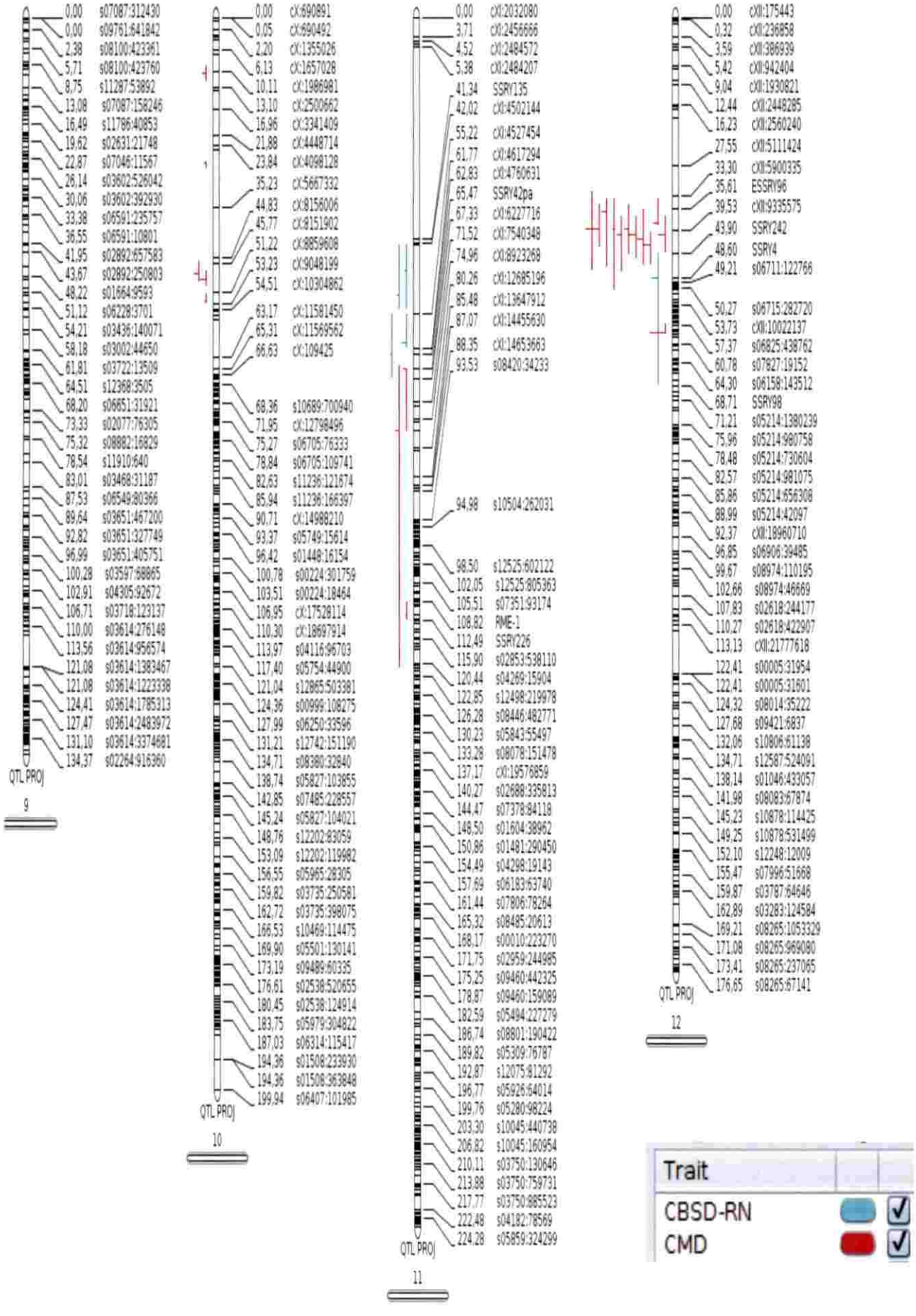

Figure 5. QTLs Projected on Consensus map after QTL Proj analysis

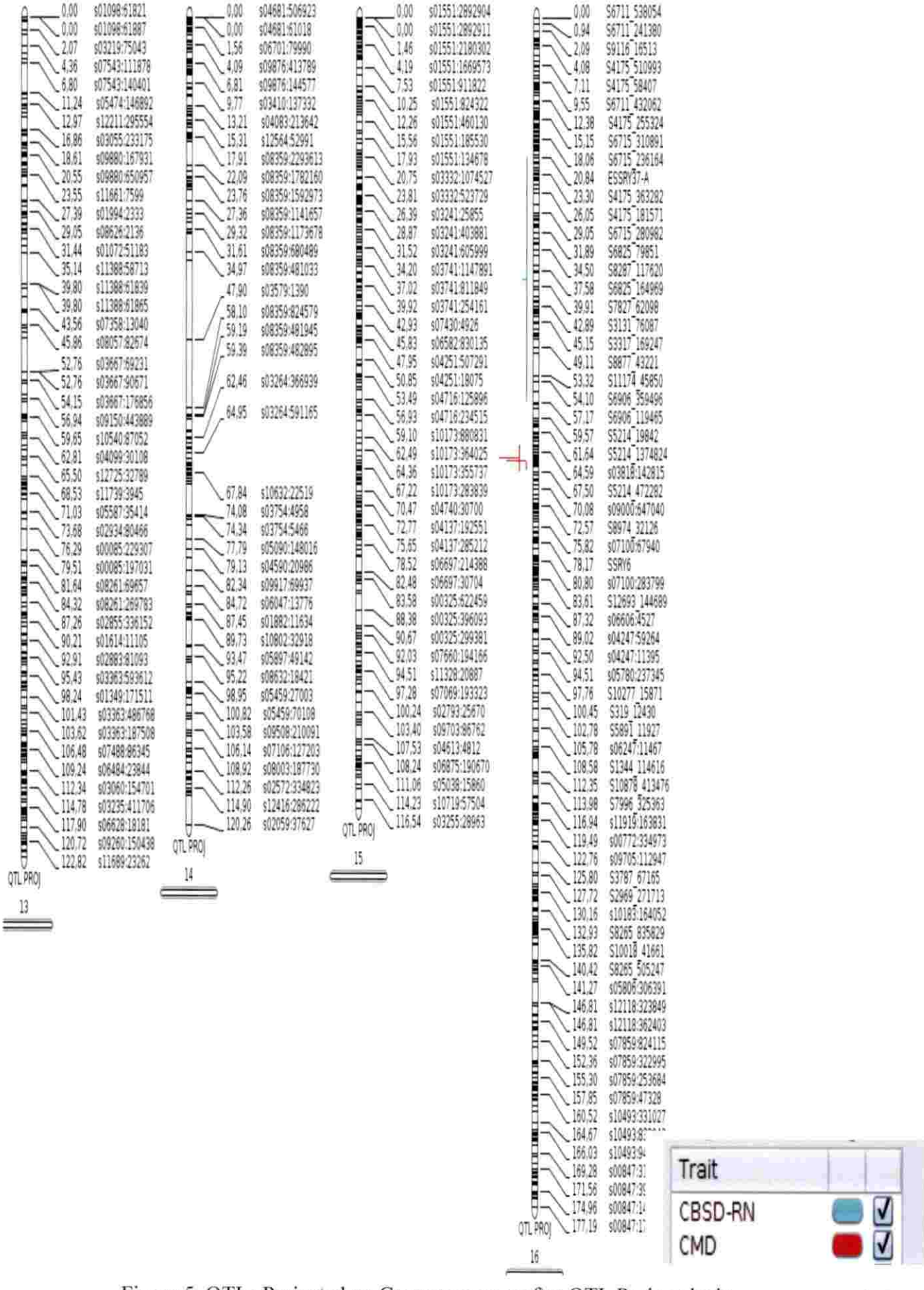

Figure 5. QTLs Projected on Consensus map after QTL Proj analysis

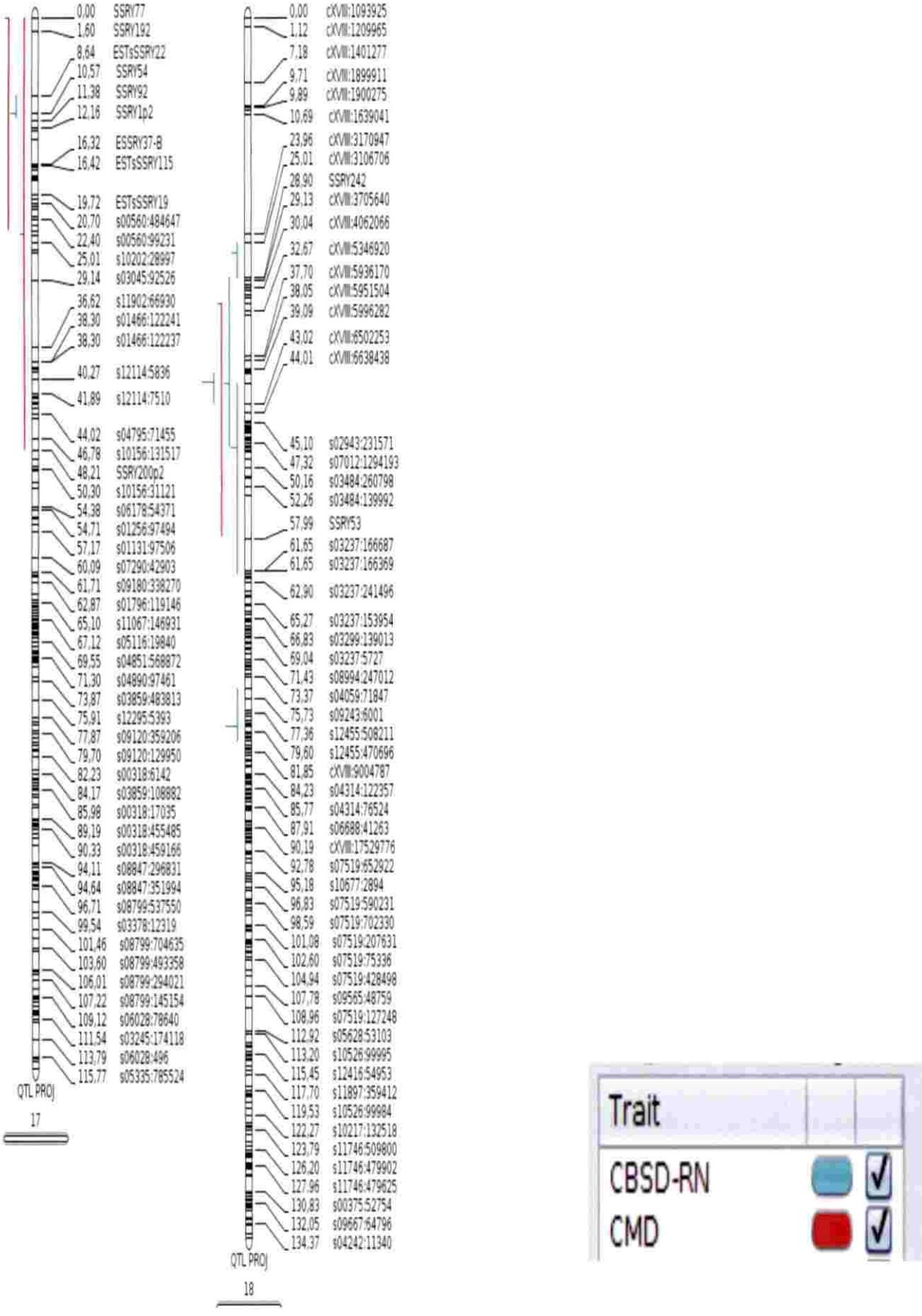

Figure 5. QTLs Projected on Consensus map after QTL Proj analysis

Table 6. Details of projected QTLs

Chromosome no.	No. of projected QTLs	Traits	Name of QTLs	Confidence Interval (cM)
1	4	CMD CBSD-RN	1. qCMDc1-1A-I 2. qCMDc1-2A-I 3. qCMDc1-3A-I 4. cbsd-rn-S1-chz080	2 3 4 9.9
2	4	CMD CBSD-RN	1. qCMDc2-1A 2. qCMDc2-2A 3. qCBSDRNFc-2Nm 4. qCBSDRNFc-2Nma	2 1 1.6 1.6
3	2	CMD	1. qCMDc3ACA 2. qCMDc3ACB	9 12
4	1	CBSD-RN	1. cbsd-rn-c4-ch008	26.4
5	8	CMD CBSD-RN	1. qCMDc5-1A-VN1 2. qCMDc5-1A-VN2 3. qCMDc5-2A-V 4. qCMDc5-3A 5. CQTL1 6. cbsd_rn_c5_ch007 7. cbsd_rn_c5a_ch07 8. cbsd_rn_c5b_ch07	1 1 6 5 2.4 10 8.2 18.5
7	1	CMD	1. QTL2	15
8	1	CBSD-RN	1. cbsd-rn-ch08-qt1	6.2
10	5	CMD	1. qCMDc10-1A 2. qCMDc10-2A 3. qCMDc10-3A 4. qCMDc10-3AN2 5. qCMDc10-4A	1.5 3 4 3 2

11	7	CMD CBSD-RN	1. CMD-qt5 2. CMD-qt3 3. CQTL2 4. qCBSDRNc11-2Nma 5. qCBSDRNc11-2Nm 6. qCBSDRNc11-3Nm 7. qCBSDRNc11-1Nm	10.9 12.8 3.59 11.4 11.5 11.53 6.2
12	13	CMD CBSD-RN	1. qCMDc12-1AN1 2. qCMDc12-1AN2 3. qCMDc12-1AC1 4. qCMDc12-1AC2 5. qCMDc12-1ACA 6. qCMDc12-2AN1 7. qCMDc12-2AN2 8. qCMDc12-2AC1 9. qCMDc12-2AC2 10. qCMDc12-2ANA 11. qCMDc12-2ANB 12. qCMDc12-3ACB 13. cbsd-rn-S12-chz08	7 6 10 8 10 8 8 9 16 14 5 2 21.7
16	3	CMD CBSD-RN	1. CMD1S 2. CMD3S 3. cbsd_rn_c16_ch07	1.865 2.348 31.6
17	3	CMD CBSD-RN	1. CMD-qtl 2. CMD-qt4 3. cbsd-rn-S17-ndl08	26 40.3 2.7
18	6	CMD CBSD-RN	1. CMD-qt2 2. cbsd_rn-c18_ch07 3. qCBSDRNc18-2Nm 4. qCBSDRNc18-3Nm 5. cbsd_rnc18a_ch07 6. qCBSDRNc18-1Nm	15.6 17.4 4 3.1 19.4 5.9
				Average : CMD - 7.79 CBSD - 11.36

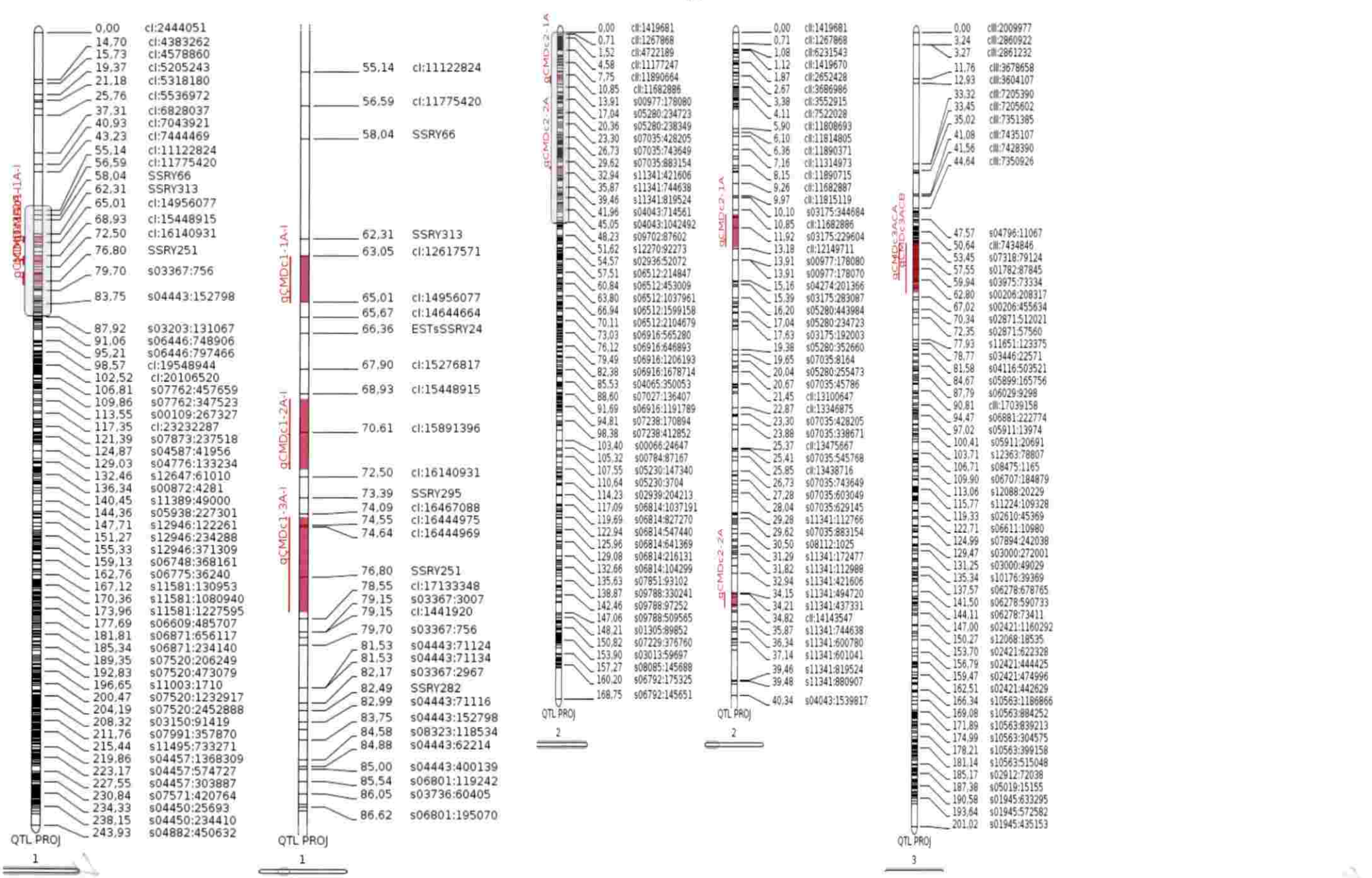

Figure 6. Meta-QTLs for CMD resistance QTLs in chromosomes 1,2 and 3. Vertical lines on the left of the chromosomes show the confidence interval of each QTLs.

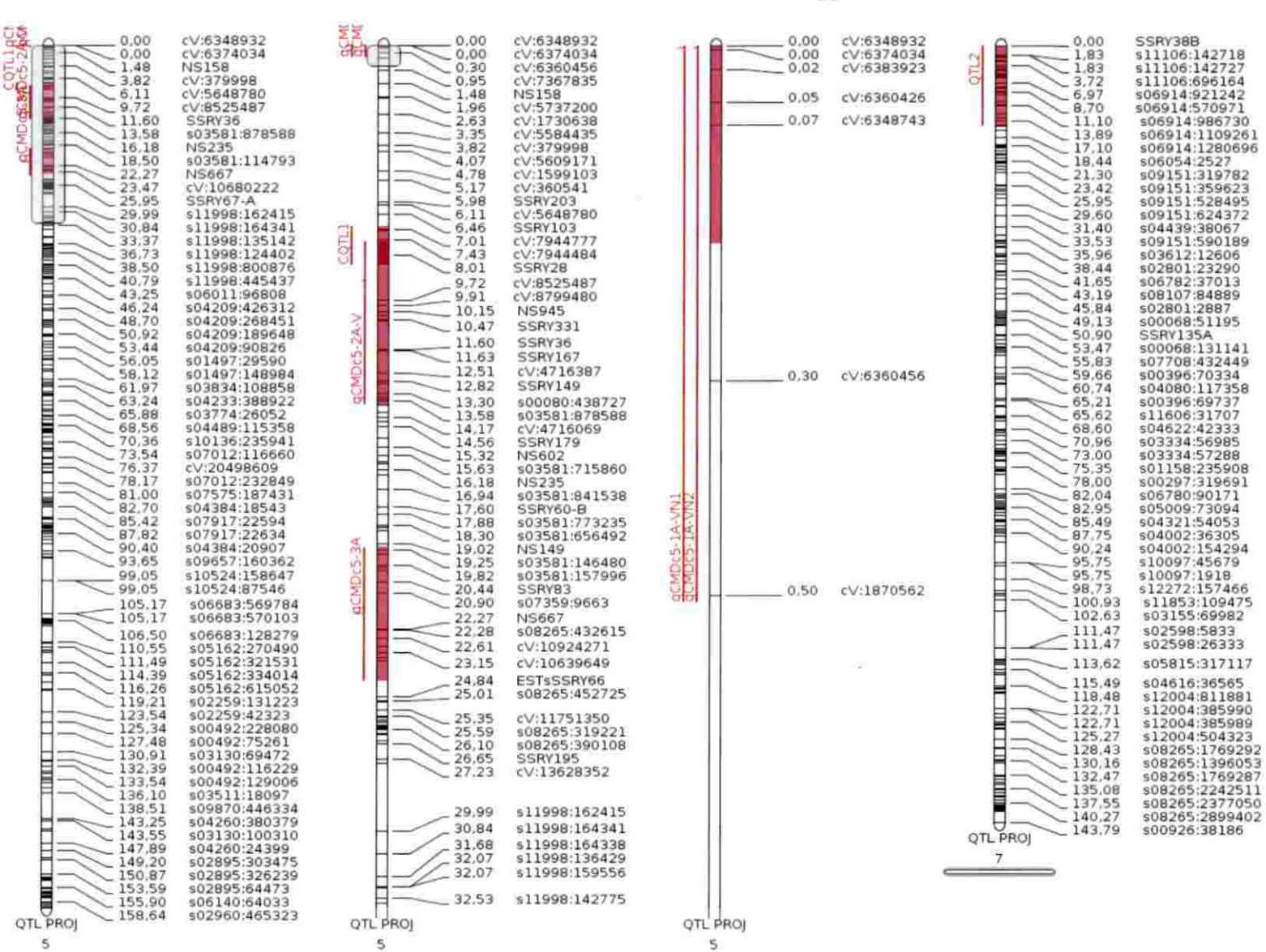

Figure 6. Meta-QTLs for CMD resistance QTLs in chromosomes 5 and 7. Vertical lines on the left of the chromosomes show the confidence interval of each QTL.

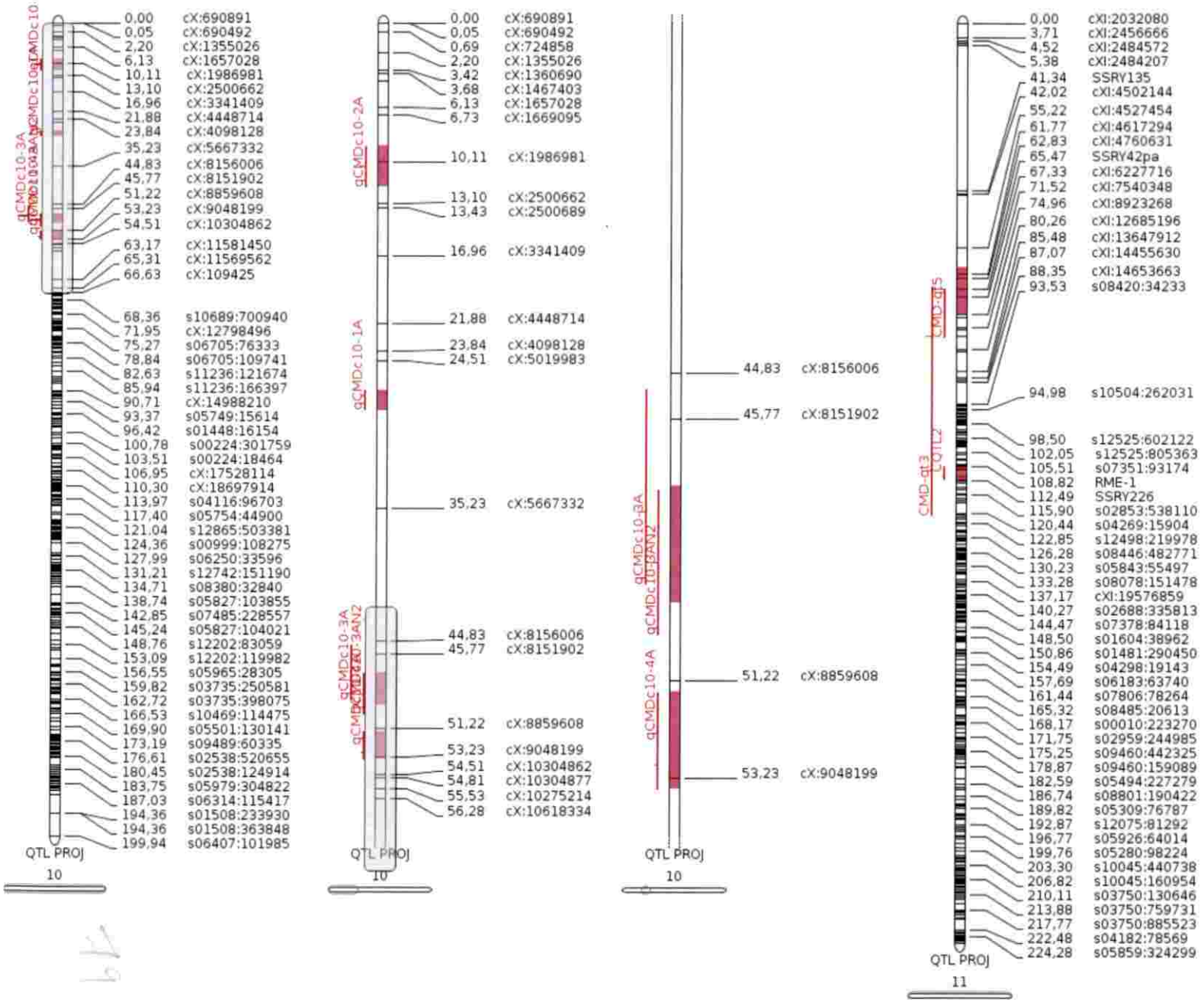

Figure 6. Meta-QTLs for CMD resistance QTLs in chromosomes 10 and 11. Vertical lines on the left of the chromosomes show the confidence interval of each QTLs.

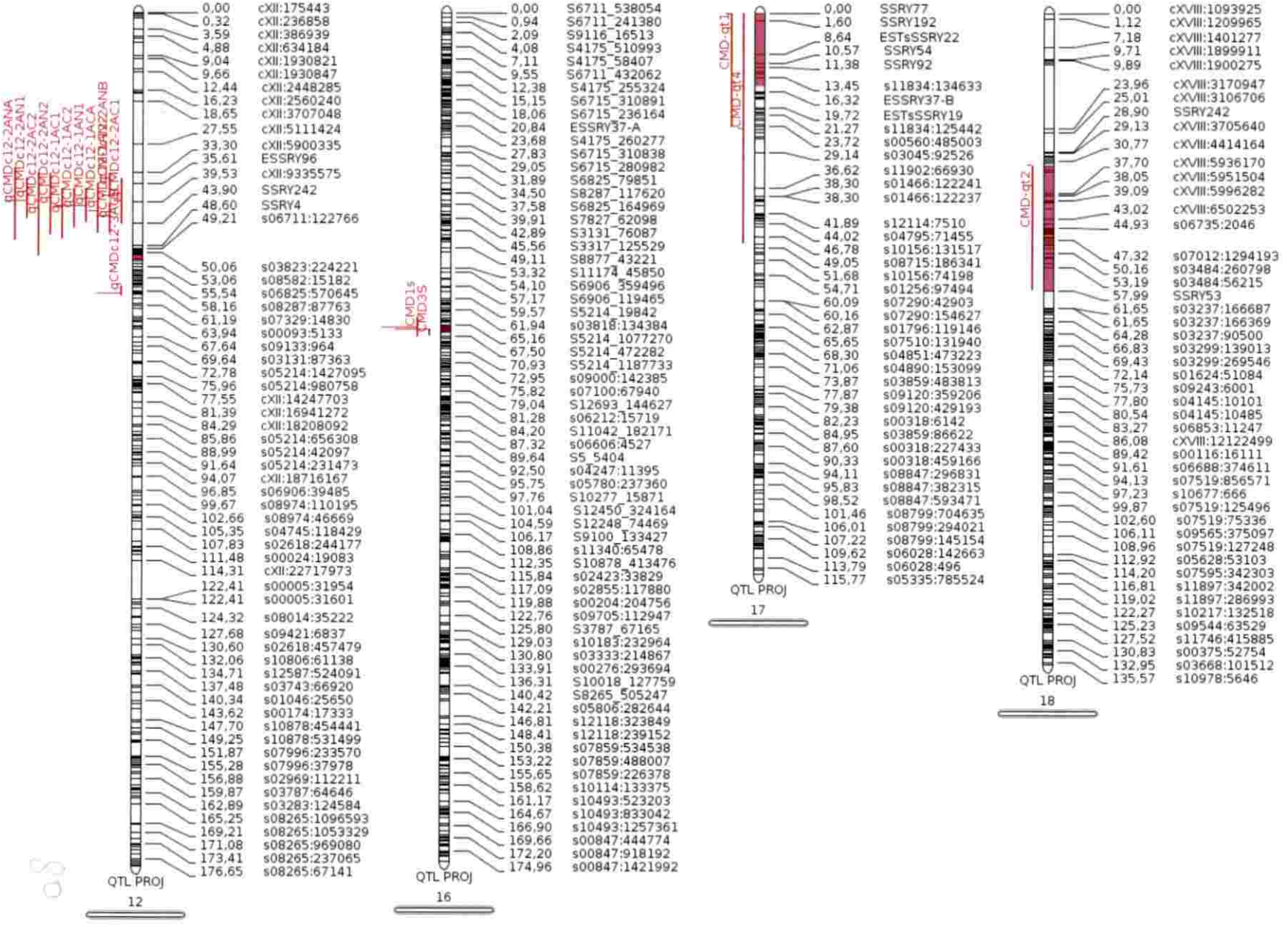

Figure 6. Meta-QTLs for CMD resistance QTLs in chromosome 12,16,17 and 18 . Vertical lines on the left of the chromosomes show the confidence interval of each QTL.

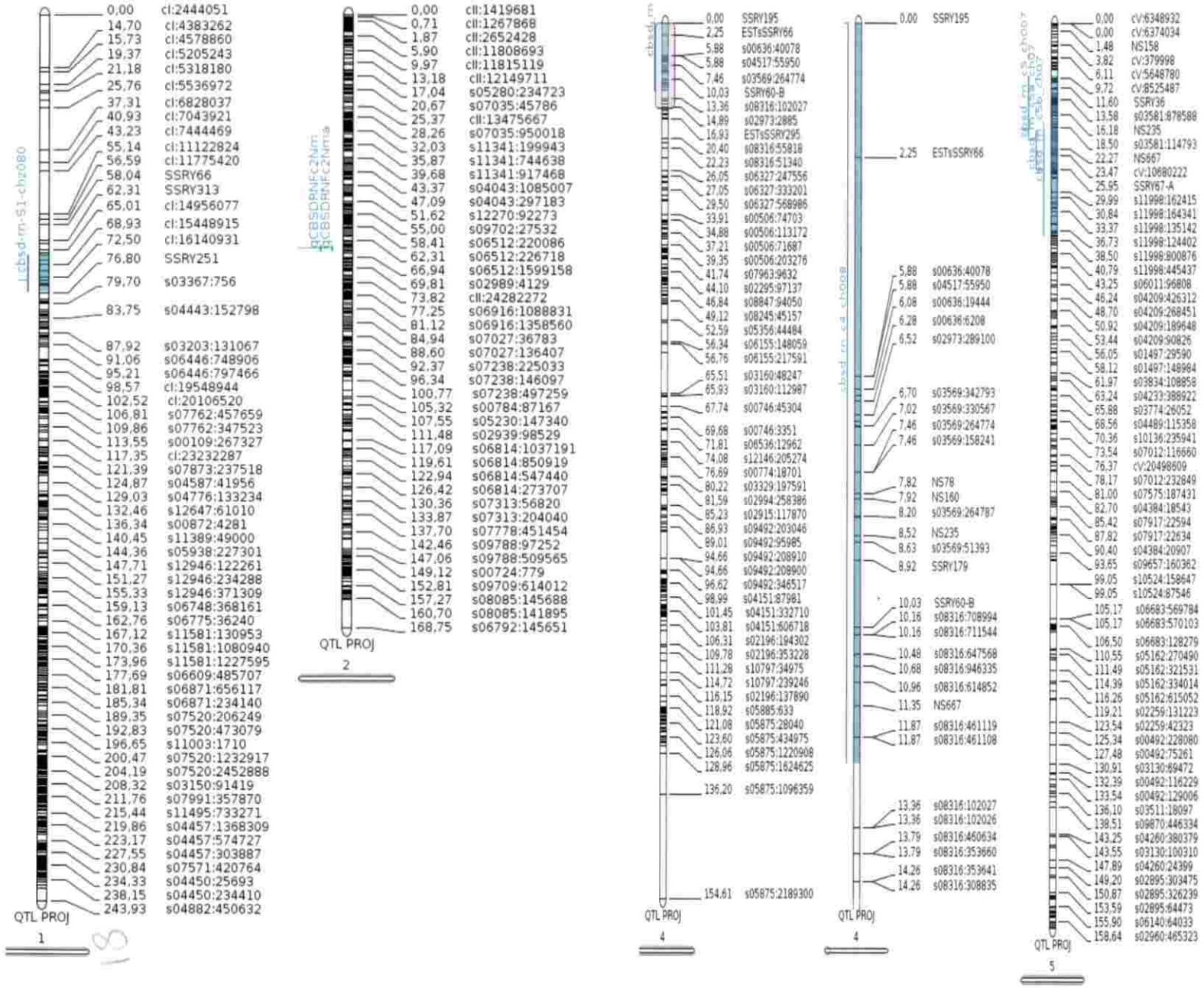

Figure 7. MetaQTLs for CBSD-RN resistance QTLs in chromosome 1,2,4 and 5. Vertical lines on the left of the chromosomes show the confidence interval of each QTL

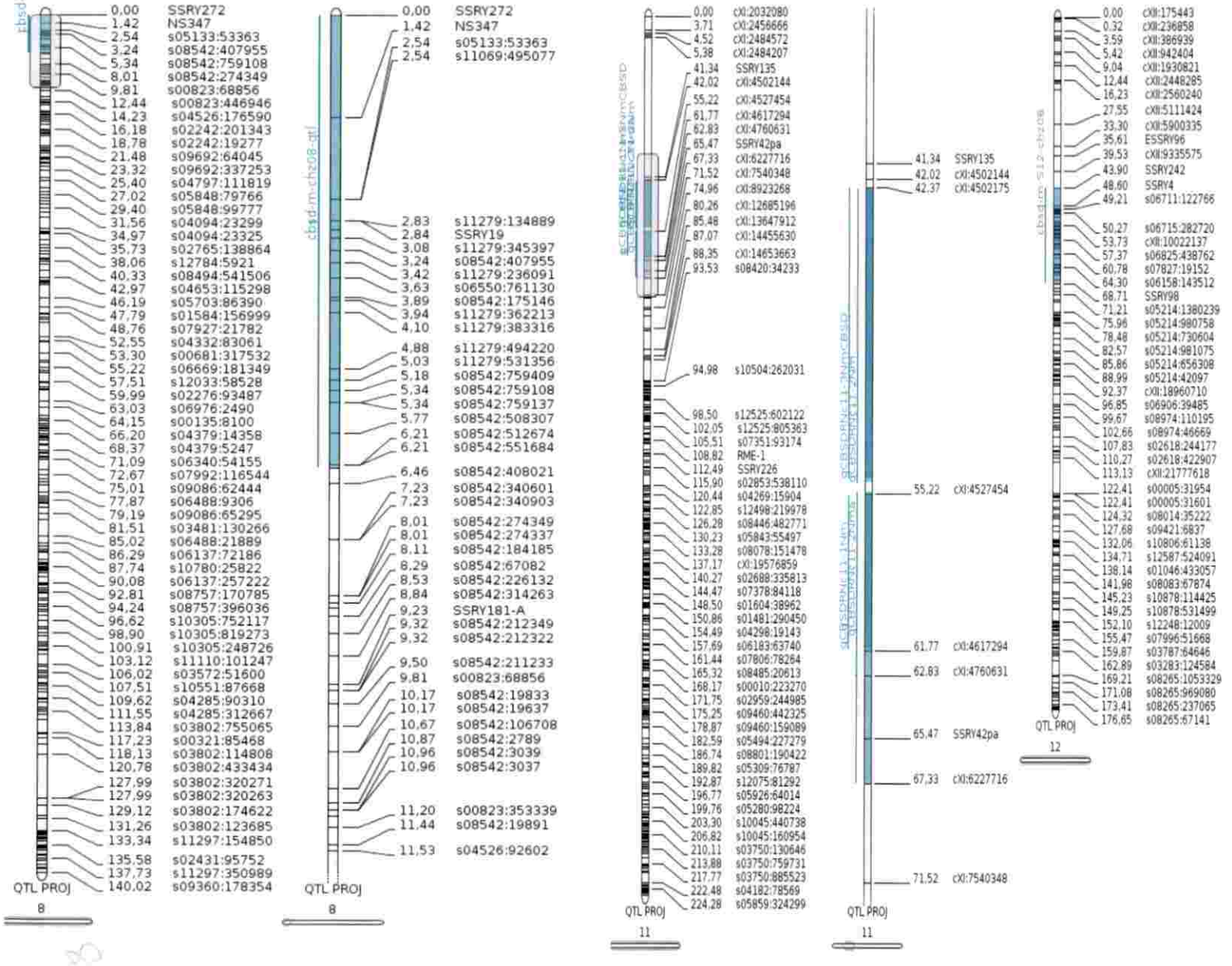

Figure 7. Meta-QTLs for CBSD-RN resistance QTLs in chromosome 8,11 and 12 . Vertical lines on the left of the chromosomes show the confidence interval of each QTL.

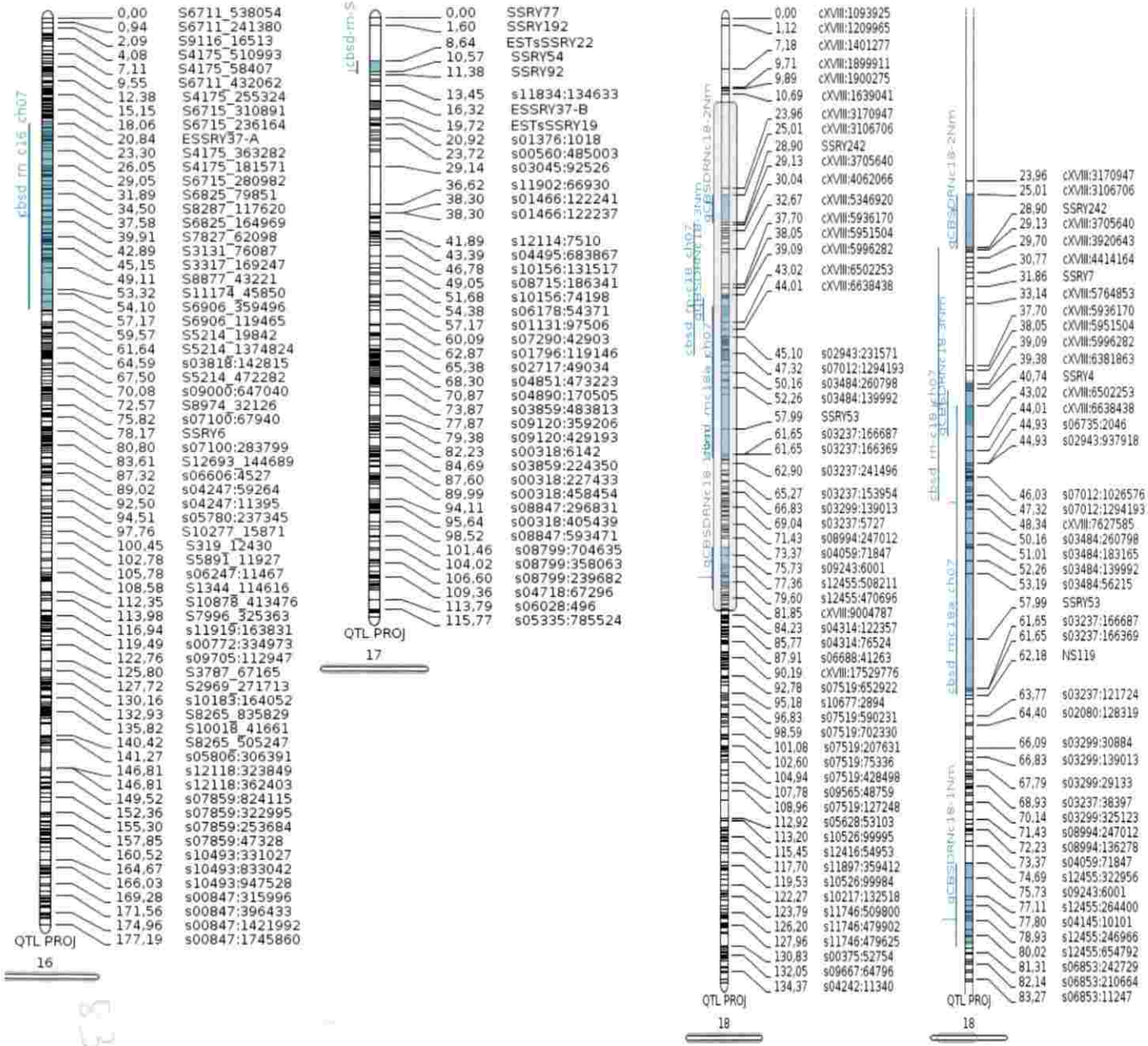

Figure 7. Meta-QTLs for CBSD-RN resistance QTLs in chromosome 16,17 and 18 . Vertical lines on the left of the chromosomes show the confidence interval of each QTLs.

Table 7. Details of MQTLs obtained for CMD resistance

Sl. No.	Chromosome no.	MQTL name	AIC criterion	QTLs in MQTL	Confidence Interval of MQTL (cM)	Flanking markers	
						Left marker	Right marker
1	1	MQTLCMD1.1	10.42	qCMDc1-1A-I	1.96	CI:12617571	CI: 14956077
				qCMDc1-2A-I	3.57	CI:15448915	CI:16140931
				qCMDc1-3A-I	4.46	CI:16467088	CI:17133348
2	2	MQTLCMD2.1	2.43	qCMDc2-1A	2.03	CII:11682886	CII:12149711
				qCMDc2-2A	1.01	CII:13956600	CII:14143547
3	3	MQTLCMD3.1	10.58	qCMDc3ACA	1.23	s07318:66188	s01782:129375
				qCMDc3ACB	0.65	s03975:73334	s07961:62978
4	5	MQTLCMD5.1	14.95	qCMDc5-1A-VN1	0.51	CV:6348932	CV:6374034
				qCMDc5-1A-VN2	0.51	CV:6348932	CV:6374034
				qCMDc5-2A-V	6.22	CV:5609145	CV:8525487
				CQTL1	2.41	CV: 5609145	CV:5609152
				qCMDc5-3A	5.12	s07359:9663	s08265:432615
5	7	MQTLCMD7.1	6.52	QTL2	3.45	s06914:1234619	s06914:1376671
6	10	MQTLCMD10.1	20.08	qCMDc10-1A	10.72	CX:5019983	CX:5667332
				qCMDc10-2A	2.99	CX:1986981	CX:2506662
				qCMDc10-3A	5.45	CX:8151902	CX:8859602
				qCMDc10-3AN2	8.74	CX:9048199	CX:10304862

				qCMDc10-4A	2.01	CX:8859608	CX:9048199
7	11	MQTLCMD11.1	19.62	CMD-qt5	4.45	CXI:4760631	CXI:6227716
				CMD-qt3	4.81	CXI:8832296	CXI:12685196
				CQTL2	2.41	s04324:82225	s04324:234003
8	12	MQTLCMD12.1	7.52	qCMDc12-2ANB	6.23	CXII:5900335	CXII:9335575
				qCMDc12-2AN2	6.23	CXII:5900335	CXII:9335575
				qCMDc12-2ANA	6.23	CXII:5900335	CXII:9335575
				qCMDc12-2AC1	6.23	CXII:5900335	CXII:9335575
				qCMDc12-2AC2	6.23	CXII:5900335	CXII:9335575
				qCMDc12-2AN1	6.23	CXII:5900335	CXII:9335575
				qCMDc-1AC2	3.25	s05214:371683	s05214:231473
				qCMDc12-1AN1	3.25	s05214:371683	s05214:231473
				qCMDc12-1ACA	2.65	s05214:42097	s05214:361175
				qCMDc12-1AN2	4.68	s05214:371707	s05214:283847
				qCMDc12-1AC1	3.25	s05214:371683	s05214:231473
				qCMDc12-3ACB	2.44	s11042:125671	s02618:324414
9	16	MQTLCMD16.1	8.3	CMD1S	0.68	s06419:73528	s09000:643658
				CMD3S	0.74	s09000:643658	s09000:363984
10	17	MQTLCMD17.1	3.2	CMD-qt1	5.35	s01493:61265	s01493:60896
				CMD-qt4	7.48	s0304:92526	s11902:66930
11	18	MQTLCMD18.1	30.86	CMD-qt2	1.48	CXVIII:4565306	CXVIII:5346920
					Average - 3.877		

Table 8. Details of MQTLs obtained for CBSD-RN resistance

Sl. No.	Chromosome no.	MQTL name	AIC criterion	QTLs in MQTL	Confidence Interval of MQTLs (cM)	Flanking markers	
						Left marker	Right marker
1	1	MQTLCBSD1.1	5.79	cbsd-rn-S1-chz080	10.5	CI:16140931	CI:16467088
2	2	MQTLCBSD2.1	2.92	qCBSDRNFc2Nm	1.61	s06512:959076	s06512:1204395
				qCBSDRNFc2Nma	1.61	s06512:959076	s06512:1204395
3	4	MQTLCBSD4.1	6.12	Cbsd-rn-c4-chz008	12.32	SSRY195	ESTsSSRY66
4	5	MQTLCBSD5.1	18.59	cbsd_rn_c5_ch007	10	S03581:114793	S03581:146480
				cbsd_rn_c5a_ch007	7.11	s03581:955446	s07359:9663
				cbsd_rn_c5b_ch007	19.1	s08265:398641	s08625:384532
5	8	MQTLCBSD8.1	4.76	Cbsd-rn-ch08qt1	6.2	s11279:134889	s11279:274068
6	11	MQTLCBSD11.1	26.7	qCBSDRNc11-2Nma	1.06	CXI:4617294	CXI: 4760631
				qCBSDRNc11-2Nm	12.35	CXI: 4502175	CXI:4527454
				qCBSDRNc11-1Nm	6.55	CXI:4527454	CXI:4617294
				qCBSDRNc11-3Nm	10.2	CXI:4502144	CXI:4527454
7	12	MQTLCBSD12.1	7.52	cbsd-rn-S12-chz08	7.1	s06906:361646	s06906:70963
8	16	MQTLCBSD16.1	8.3	cbsd_rn_c16_ch07	5.71	s05780:193032	s08127:46086
9	17	MQTLCBSD17.1	3.2	cbsd-rn-S17-nd08	1.32	s00560:99231	s00560:484985
10	18	MQTLCBSD18.1	30.86	cbsd_rn-c18_ch07	3.43	CXVIII:6320754	CXVIII:6502253
				qCBSDRNc18-2Nm	4.12	CXVIII:3106706	CXVIII:3705640
				qCBSDRNc18-3Nm	1.46	s04314:122357	s04314:105540
				cbsd_rnc18a_ch07	0.86	s06988:59802	s02080:46710
				qCBSDRNc18-1Nm	0.15	s12455:451573	s12455:580350
					Average - 6.43		

4.6. IDENTIFICATION OF CANDIDATE GENES

The MQTLs with narrowed genetic and physical interval were selected for mining candidate genes underlying the QTLs. MQTLs with CI less than 2 cM were selected for identifying candidate genes. Nine QTLs within six MQTLs associated with CMD resistance (Table 9) and seven QTLs within four MQTLs associated with CBSD-RN resistance (Table 10) were obtained within 2 cM CI. The physical positions of flanking/closer markers of QTLs were identified using ICGMC map. The QTLs with less than 1Mega base pair (Mb) length were selected. For CMD resistance seven QTLs were identified with physical length less than 1Mb and for CBSD-RN resistance all the seven QTLs were found within 1Mb length. These QTLs with narrowed physical intervals are then integrated with genomic information for identifying genes. The physical positions of flanking markers of selected QTLs were located in whole genome assembly of cassava (*Manihot esculenta* V.6 assembly) using Phytozome V.12.1.Database. The genes present in between the physical position of markers were selected and listed. A total of 73 genes were identified for CMD resistance QTLs (Table11) and 51 genes were obtained for CBSD-RN resistance QTLs (Table12).

4.7 FUNCTIONAL ANNOTATION OF CANDIDATE GENES USING PHYTOMINE

The selected candidate genes with less than 1Mb physical length were then subjected to functional annotation. The genes ID's of identified candidate genes were searched against Phytomine interface in Phytozome. The functional information such as PFAM (protein family) and Gene Ontology (GO) were obtained. From 73 genes identified for CMD resistance 38 genes were annotated (Table 13) other genes were uncharacterized or without any annotation. From 51 genes identified for CBSD-RN resistance 30 genes were annotated (Table 14).

Table 9. Details of CMD resistance MQTLs selected for mining Candidate genes

MQTL name	QTLs in MQTL	Position (cM)	CI (cM)	Flanking markers		Position in Cassava V.5.1 (bp)	Position in Cassava V.6.1 (bp)	Physical length (Mb)
				Left marker	Right marker			
MQTLCMD1.1	qCMDc1-1A-I	64.60	1.96	CI:12617571	CI: 14956077	12617571-14956077	20644083-22495692	1.7
MQTLCMD2.1	qCMDc2-2A	34.57	1.01	CII:13956600	CII:14143547	13956600-14143547	15197682-15391499	0.193
MQTLCMD3.1	qCMDc3ACA	57.24	1.23	s07318:66188	s01782:129375		3110072-3310263	0.2
	qCMDc3ACB	59.40	0.65	s03975:73334	s07961:62978		3740747-3846382	0.1
MQTLCMD5.1	qCMDc5-1A-VN1	43.39	0.51	CV:6348932	CV:6374034	6348932-6374034	6510374-6547440	0.037
	qCMDc5-1A-VN2	43.39	0.51	CV:6348932	CV:6374034	6348932-6374034	6510374-6547440	0.037
MQTLCMD16.1	CMD1S	65.64	0.68	s06419:73528	s09000:643658		1734878-1774334	0.039
	CMD3S	66.07	0.74	s09000:643658	s09000:363984		1774334-2054008	0.279
MQTLCMD18.1	CMD-qt2	31.86	1.48	CXVIII:4565306	CXVIII:5346920	4565306-5346920	4795191-5831555	5

Table10. Details of CBSD-RN resistance MQTLs selected for mining Candidate genes

MQTL name	QTLs in MQTL	Position (cM)	CI (cM)	Flanking markers		Position in Cassava V.5.1 (bp)	Position in Cassava V.6.1 (bp)	Physical length (Mb)
				Left marker	Right marker			
MQTLCBSD2.1	qCBSDRNFc2Nm	64.49	1.61	s06512:959076	s06512:1204395		7451724-7697043	0.245
	qCBSDRNFc2Nma	64.49	1.61	s06512:959076	s06512:1204395		7451724-7697043	0.245
MQTLCBSD11.1	qCBSDRNc11-2Nma	62.82	1.06	CXI:4617294	CXI: 4760631	4617294-4760631	5630044-5761172	0.131
MQTLCBSD17.1	cbsd-rn-S17-ndl08	10.60	1.32	s00560:99231	s00560:484985		1153442- 1539196	0.385
MQTLCBSD18.1	qCBSDRNc18-3Nm	40.56	1.46	s04314:122357	s04314:105540		5473689-5547914	0.206
	cbsd_rnc18a_ch07	47.89	0.86	s06988:59802	s02080:46710		6555344-6582765	0.027
	qCBSDRNc18-1Nm	78.91	0.15	s12455:451573	s12455:580350		10068084-10068441	0.0035

Table 11. Candidate genes identified for QTLs associated with CMD resistance

Sl. No.	QTL name	Candidate genes	Length (bp)
1	qCMDc2-2A	1. Manes.02G187400.1	10,579
		2. Manes.02G187500.1	7,705
		3. Manes.02G187600.1	4,084
		4. Manes.02G187700.1	748
		5. Manes.02G187800.1	276
		6. Manes.02G187900.1	614
		7. Manes.02G188000.1	7,254
		8. Manes.02G188100.1	1,151
		9. Manes.02G188200.1	345
		10. Manes.02G188300.1	1,542
		11. Manes.02G188400.1	10,039
		12. Manes.02G188500.1	636
		13. Manes.02G188600.1	3,304
		14. Manes.02G188700.1	225
		15. Manes.02G188800.1	934
		16. Manes.02G188900.1	3,464
2.	qCMDc3ACA	1. Manes.03G038400.1	2,448
		2. Manes.03G038500.1	682
		3. Manes.03G038600.1	3,120
		4. Manes.03G038700.1	1,437
		5. Manes.03G038800.1	2,729
		6. Manes.03G038900.1	1,938
		7. Manes.03G039000.1	648
		8. Manes.03G039100.1	3,421
		9. Manes.03G039200.1	3,835
		10. Manes.03G039300.1	397
		11. Manes.03G039400.1	921
		12. Manes.03G039500.1	1,648
		13. Manes.03G039600.1	4,585
		14. Manes.03G039700.1	2,169
		15. Manes.03G039800.1	921
		16. Manes.03G039900.1	3,052
		17. Manes.03G040000.1	4,752
		18. Manes.03G040100.1	2,816
		19. Manes.03G040200.1	2,139
		20. Manes.03G040300.1	6,688
		21. Manes.03G040400.1	5,707
		22. Manes.03G040500.1	797
		23. Manes.03G040600.1	3,287

3.	qCMDc3ACB	1. Manes.03G045900.1 2. Manes.03G046000.1 3. Manes.03G046100.1 4. Manes.03G046200.1 5. Manes.03G046300.1 6. Manes.03G046400.1	4,567 6,691 2,857 3,491 458 4,354
4	qCMDc5-1A-VN1 & qCMDc5-1A-VN2	1. Manes.05G084900.1 2. Manes.05G085000.1 3. Manes.05G085100.1 4. Manes.05G085200.1 5. Manes.05G085300.1	6,944 246 3,807 14,180 2,790
6	CMD1S	1. Manes.16G018100.1 2. Manes.16G018200.1 3. Manes.16G018300.1 4. Manes.16G018400.1 5. Manes.16G018500.1	1,962 3,123 3,131 2,690 3,392
7	CMD3S	1. Manes.16G018700.1 2. Manes.16G018800.1 3. Manes.16G018900.1 4. Manes.16G019000.1 5. Manes.16G019100.1 6. Manes.16G019200.1 7. Manes.16G019300.1 8. Manes.16G019400.1 9. Manes.16G019500.1 10. Manes.16G019600.1 11. Manes.16G019700.1 12. Manes.16G019800.1 13. Manes.16G019900.1 14. Manes.16G020000.1 15. Manes.16G020100.1 16. Manes.16G020200.1 17. Manes.16G020300.1 18. Manes.16G020400.1	3,399 3,618 4,504 3,344 4,968 1,937 3,207 8,260 5,772 1,127 1,565 526 678 3,250 3,368 1,725 117 7,140

Table 12. Candidate genes identified for QTLs associated with CBSD-RN resistance

Sl. No.	QTL name	Candidate genes	Length (bp)
1	qCBSDRNFc2Nm	1. Manes.02G100100.1	3,020
		2. Manes.02G100200.1	108
		3. Manes.02G101000.1	5,368
		4. Manes.02G101300.1	4,326
		5. Manes.02G100500.1	912
		6. Manes.02G101500.1	4,375
		7. Manes.02G101800.1	2,520
		8. Manes.02G100300.1	3,036
		9. Manes.02G100400.1	2,539
		10. Manes.02G100600.1	4,303
		11. Manes.02G100700.1	5,123
		12 Manes.02G100800.1	2,284
		13. Manes.02G100900.1	9,520
		14. Manes.02G101200.1	1,815
		15. Manes.02G101400.1	11,455
		16 .Manes.02G101700.1	8,758
		17. Manes.02G101900.1	4,668
		18. Manes.02G102100.1	2,378
2.	qCBSDRNFc2Nma	1. Manes.02G100100.1	3,020
		2. Manes.02G100200.1	108
		3. Manes.02G101000.1	5,368
		4. Manes.02G101300.1	4,326
		5. Manes.02G100500.1	912
		6. Manes.02G101500.1	4,375
		7. Manes.02G101800.1	2,520
		8. Manes.02G100300.1	3,036
		9. Manes.02G100400.1	2,539
		10. Manes.02G100600.1	4,303
		11. Manes.02G100700.1	5,123
		12 Manes.02G100800.1	2,284
		13. Manes.02G100900.1	9,520
		14. Manes.02G101200.1	1,815
		15. Manes.02G101400.1	11,455
		16 .Manes.02G101700.1	8,758
		17. Manes.02G101900.1	4,668
		18. Manes.02G102100.1	2,378
3.	CBSDRNc11-2Nma	1. Manes.11G058000.1	7,979
		2. Manes.11G058100.1	927
		3. Manes.11G058200.1	3,735

		4. Manes.11G058300.1 5. Manes.11G058400.1 6. Manes.11G058500.1 7. Manes.11G058600.1 8. Manes.11G058700.1 9. Manes.11G058800.1	4,056 3,057 5,068 2,327 5,224 12,766
4	cbsd-rn-S17-ndl08	1. Manes.17G005200.1 2. Manes.17G005300.1 3. Manes.17G005400.1 4. Manes.17G005500.1 5. Manes.17G005600.1 6. Manes.17G005700.1 7. Manes.17G005800.1 8. Manes.17G005900.1 9. Manes.17G006000.1 10. Manes.17G006100.1 11. Manes.17G006200.1 12. Manes.17G006300.1 13. Manes.17G006400.1	896 1,018 1,795 10,762 12,674 6,658 141 1,422 3,620 4,275 9,699 1,338 4,097
5	qCBSDRNc18-3Nm	1. Manes.18G063700.1 2. Manes.18G063800.1 3. Manes.18G063900.1 4. Manes.18G064000.1	3,597 1,596 1,764 468
6	cbsd_rnc18a_ch07	1. Manes.18G076400.1 2. Manes.18G076500.1 3. Manes.18G076600.1 4. Manes.18G076700.1	1,670 4,783 1,802 21,078
7	qCBSDRNc18-1Nm	1. Manes.18G109400.1 2. Manes.18G109500.1 3. Manes.18G109600.1	7,054 978 2,293

Table 13. Functional annotation of candidate genes identified for QTLs associated with CMD resistance

QTL name	Genes	PFAM	GO
qCMDc2-2A	Manes.02G187400.1	PF06046 : Exocyst complex component	GO:0006887 : Secretion by cells that results in the release of intracellular molecules
	Manes.02G187500.1	PF00069: Protein kinase domain PF00560: Leucine Rich Repeat	GO:0004672: catalysis of the phosphorylation of an amino acid residues
	Manes.02G187700.1	PF00188: Cysteine-rich secretory protein family	
	Manes.02G188000.1	PF00400: WD domain, G-beta repeat	GO:0005515: Interacting with any protein or protein complex
	Manes.02G188100.1	PF00213: ATP synthase delta	
	Manes.02G188300.1	PF02536:Mitochondrial transcription termination factor	GO:0003690: Interacting with double-stranded DNA
	Manes.02G188600.1	PF03016: Exostosin family	
	Manes.02G188800.1	PF00704: Glycosyl hydrolases family 18	GO:0004553: Catalysis of the hydrolysis of any O-glycosyl bond
qCMDc3ACA	Manes.03G038400.1	PF00924: Mechanosensitive ion channel	GO:0055085: Transportation of solute across a lipid bilayer
	Manes.03G038500.1	PF06596: Photosystem II reaction centre X protein	GO:0009523: A photosystem that contains a pheophytin-quinone reaction center with associated accessory pigments and electron carriers
	Manes.03G038800.1	PF00189: Ribosomal protein C-terminal domain	GO:0003723: Interacting selectively with an RNA molecule or a portion thereof.
	Manes.03G039000.1	PF07876: Stress responsive A/B Barrel Domain	
	Manes.03G039100.1	PF00067: Cytochrome P450	GO:0005506: Interacting selectively and non-covalently with iron (Fe) ions
	Manes.03G039500.1	PF00724: NADH oxidase family	GO:0006351: The cellular synthesis of RNA on a template of DNA
	Manes.03G039700.1	PF07883: Jasmonic acid biosynthesis	
	Manes.03G039800.1	PF00332: Glycosyl hydrolases family 17	GO:0004553:Catalysis of the hydrolysis of any O-glycosyl bond.
	Manes.03G040100.1	PF10250: fucosyl transferase	

	Manes.03G040200.1	PF00583: Acetyltransferase	
	Manes.03G040400.1	PF03133: Tubulin-tyrosine ligase family	
qCMDc3ACB	Manes.03G046000.1	PF01612: 3'-5' exonuclease	
	Manes.03G045900.1	PF01417: ENTH domain	
	Manes.03G046100.1	PF05907: Protein of unknown Function	
	Manes.03G046200.1	PF00069: Protein kinase domain	GO:0006468: The process of introducing a phosphate group on to a protein.
qCMDc5-1A-VN1& qCMDc5-1A-VN2	Manes.05G084900.1	PF01501: Glycosyl transferase family 8	GO:0016757 :Catalysis of the transfer of a glycosyl group
	Manes.05G085100.1	PF13426: RNA SPLICING PROTEIN	
	Manes.05G085200.1	PF00097: Zinc finger	GO:0046872: Interacting selectively and non-covalently with any metal ion
	Manes.05G085300.1	PF13414 : TPR repeat PF12796 : Ankyrin repeats	
CMD1S	Manes.16G018100.1	PF07714: Protein tyrosine kinase	GO:0004672: protein kinase activity
	Manes.16G018200.1	PF07714: Protein tyrosine kinase	GO:0006468: protein phosphorylation
	Manes.16G018300.1	PF08276: Protein tyrosine kinase	GO:0004672: protein kinase activity
	Manes.16G018400.1	PF01453: D-mannose binding lectin	
CMD3S	Manes.16G018700.1	PF07714: Protein tyrosine kinase	GO:0004674: protein serine/threonine kinase activity
	Manes.16G019100.1	PF12398: Receptor serine/threonine kinase	GO:0006468: The process of introducing a phosphate group on to a protein
	Manes.16G019300.1	PF01733: Nucleoside transporter	GO:0005337: nucleoside transmembrane transporter activity
	Manes.16G019800.1	PF12609: Wound-induced protein	
	Manes.16G020000.1		GO:0005789: endoplasmic reticulum membrane
	Manes.16G020100.1		GO:0055085: transmembrane transport
	Manes.16G020400.1	PF13516: Leucine Rich repeat	GO:0005515: protein binding, Interacting selectively with any protein or protein complexes

Table 14. Functional annotation of candidate genes identified for QTLs associated with CBSD-RN resistance

QTL name	Genes	PFAM	GO
qCBSDRNFc2Nm & qCBSDRNFc2Nma	Manes.02G100300.1	PF00171:Aldehyde dehydrogenase family	GO:0016491: Catalysis of an (redox) reaction
	Manes.02G100400.1	PF01486: K-box region	GO:0003677: Any molecular function by which a gene product interacts selectively and non-covalently with DNA
	Manes.02G100600.1	PF13325: N-terminal region of micro-spherule protein	GO:0005515: Interacting selectively and non-covalently with any protein or protein complex
	Manes.02G100700.1	PF13855: Leucine rich repeat	
	Manes.02G100800.1	PF00067: Cytochrome P450	GO:0005506:Interacting selectively with iron (Fe) ions
	Manes.02G101200.1	PF03126: Plus-3 domain	GO:0003677: Any molecular function by which a gene product interacts selectively and non-covalently with DNA
	Manes.02G101400.1	PF01095:Pectinesterase	GO:0042545:The series of events leading to chemical and structural alterations of an existing cell wall
	Manes.02G101700.1	PF05577:Serine carboxypeptidase	
	Manes.02G101900.1	PF00225:Kinesin motor domain	
	Manes.02G102100.1	PF00560:Leucine Rich Repeat PF07714:Protein tyrosine kinase	GO:0006468:The process of introducing a phosphate group on to a protein
qCBSDRNc11-2Nma	Manes.11G058000.1	PF00226: DnaJ domain	GO:0031072; heat shock protein
	Manes.11G058200.1	PF02978: Signal peptide binding domain	GO:0048500: signal recognition
	Manes.11G058300.1	PF04005: Hus1-like protein	GO:0000077: DNA damage
	Manes.11G058600.1	PF11976: Ubiquitin-3 like	GO:0005515; protein binding
	Manes.11G058800.1	PF13424: Tetratricopeptide repeat	GO:0005515: protein binding
	Manes.17G005200.1	PF00335: Tetraspanin family	GO:0016021: integral component of membrane

cbsd-rn-S17-ndl08	Manes.17G005600.1	PF00069: Protein kinase domain	GO:0004672: protein kinase activity
	Manes.17G005700.1	PF08263: Leucine rich repeat N-terminal domain	GO:0004674: Serine threonine protein kinase activity
	Manes.17G005900.1	PF07690: Major Facilitator Superfamily	GO:0055085: transmembrane transport,
	Manes.17G006000.1	PF07714: Protein tyrosine kinase	GO:0004672: protein kinase activity
	Manes.17G006200.1	PF13041: PPR repeat family	GO:0005515: protein binding, Interacting selectively and non-covalently with any protein or protein complex
	Manes.17G006300.1	PF02458: Transferase family	GO:0016747: Catalysis of the transfer of an acyl group,
	Manes.17G006400.1	PF00501: AMP-binding enzyme	GO:0003824: catalytic activity
qCBSDRNc18-3Nm	Manes.18G063800.1	PF02183:Homeobox associated leucine zipper	GO:0003677: Any molecular function by which a gene product interacts selectively and non-covalently with DNA
	Manes.18G063900.1	PF09809: Mitochondrial ribosomal protein L27	
Cbsd_rnc18a_ch07	Manes.18G076400.1	PF00582: Universal stress protein family	GO:0006950: Any process that results in a change in state or activity of a cell or an organism
	Manes.18G076500.1		GO:0055085:Transportation of solute across a lipid bilayer,
qCBSDRNc18-1Nm	Manes.18G109400.1	PF00862: Sucrose synthase	GO:0005985: The chemical reactions and pathways involving sucrose
	Manes.18G109500.1	PF00135: Carboxylesterase family	
	Manes.18G109600.1	PF00501 :AMP-binding enzyme	GO:0008152: The chemical reactions and pathways, including anabolism and catabolism

4.8 EXPERIMENTAL VALIDATION

The expression of selected candidate genes was detected using SYBR Green qPCR assay. Genes Manes.03G039700.1 which is associated with Jasmonic acid biosynthesis and Manes.16G020400.1 which is associated with LRR protein were selected for validation. Cassava resistant variety MNga and Susceptible variety H-165 used for validation.

4.8.2. Primer designing using Primer3plus

Primer sequences specific to genes Manes.03G039700.1 and Manes.16G020400.1 from MQTLCMD3.1 and MQTLCMD16.1 were designed using Primer 3plus software. The primer details are given in Table 1

Table 15. Details of primers used for validation

Primer Name	Primer Sequence (5'-3')	Product size (bp)
ME1-F	GCTGCGTGCAAAACTAAAAT	180
ME1-R	TTGACATGAGTGATATTTTCTTGAG	
ME2-F	GTGCGAAATGGAAATCAATG	319
ME2-R	GCCTTCTCAGCATATGGAGC	
Actin-F	CCCAAAAGCCAACAGAGAGA	150
Actin -R	CATCACCGAGTCCAACACAAT	

4.8.3. RNA isolation

RNA was isolated from both Resistant (MNga) and Susceptible (H-165) varieties available at ICAR-CTCRI. RNA was isolated using QIAGEN RNeasy plant mini kit; TRI reagent based method and Lithium chloride method. Out of the above methods better results were obtained in $LiCL_2$ method and the RNA was stored at -80°C. Two distinct bands were observed when resolved on agarose gel (1.5%), which indicated no apparent degradation (Plate1).

4.8.4. RNA quantification

The quantity and quality of the isolated RNA was analysed using nanodrop spectrophotometer (Table 16)

4.8.5. cDNA synthesis

The isolated RNA samples of resistant and susceptible varieties were converted into cDNA using RevertAid First Strand cDNA Synthesis Kit of Thermofisher scientific. The concentraion of cDNA samples were quantified and diluted to a final concentration of 100 ng μl^{-1} and was used for the expression study using real time PCR.

4.8.6. RT- qPCR

For studying gene expression the genes present in the resistant and susceptible variety were targeted using designed specific primers ME1-F, ME1-R and ME2-F, ME2-R.The SYBR green PCR assay was used for studying gene expression. The relative gene expression level of resistant and susceptible varieties is studied using $2^{-\Delta\Delta CT}$ method. Actin was used as the reference gene for the expression study.

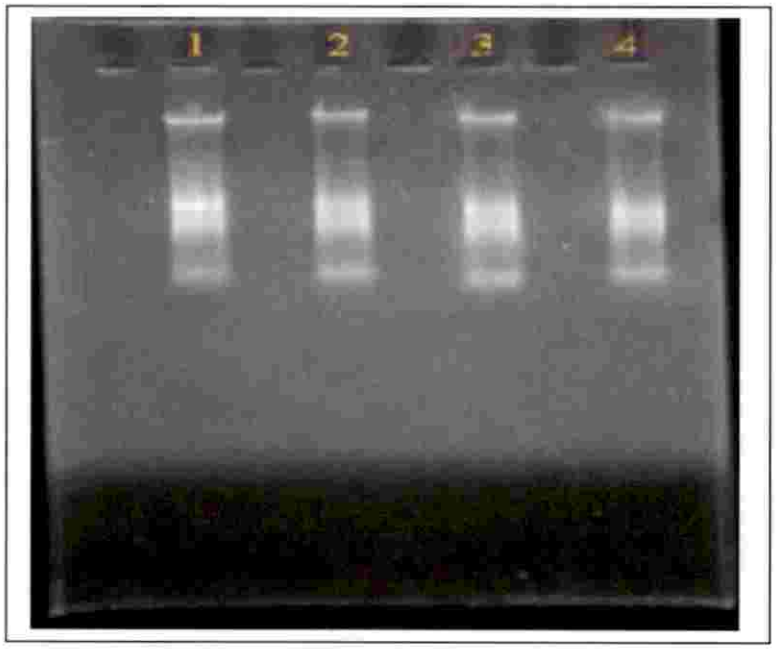

Plate1. - RNA isolated from Cassava leaf samples:

Lane 1: MNga healthy, Lane 2: MNga infected

Lane 3: H- 165 healthy, Lane 4: H- 165 infected.

Table 16. Concentration and absorbance ratio of isolated RNA estimated using nanodrop spectrophotometer

Samples	Concentration (ng/µl)	A260/A280
MNga-H	570.618	2.38
MNga-I	535.937	2.43
H-165-H	599.630	2.50
H-165-I	336.234	2.44

The standard fluorescent amplification representing exponential growth of PCR products was observed in each cycle, yielding threshold cycle (Ct) values. The Ct values is given in the logarithmic scale and inversely proportional to the quantity of cDNA. Thus highly expressed genes have low Ct value and low expressed genes have high Ct value. The fold change (-ΔΔCT) can be calculated by comparing the normalized expression (ΔCt) of the two conditions. The fold change, *viz.* the expression ratio, indicated the upregulation and downregulation of the gene.

The resistant variety (MNga) had a fold change of about 3.5 by amplifying with the primer ME-1 and had 6 fold change by amplifying with the primer ME-2. The Susceptible variety (H-165) had 1.2 fold change with primer ME-1 and 2.2 fold change with primer ME-2 (Figure 8)

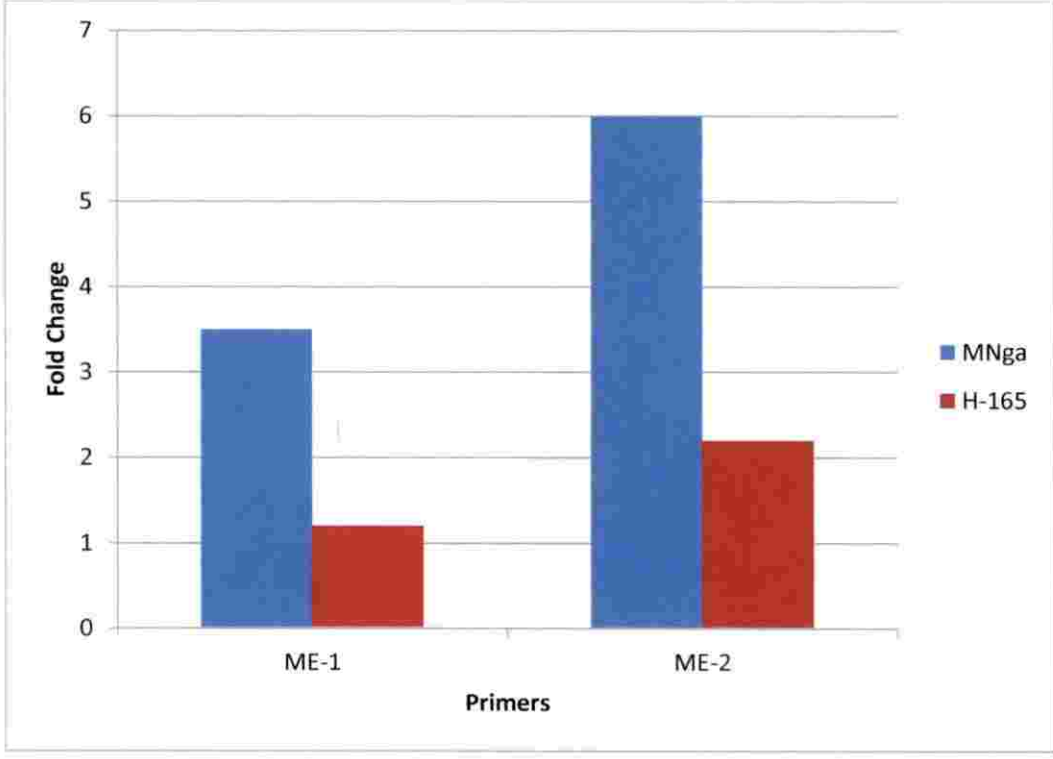

Figure 8. Relative gene expression of resistant (MNga) and susceptible (H-165) variety

DISCUSSION

5. DISCUSSION

Cassava is a starchy root crop that is grown almost entirely within tropics. It is the fourth most important source of calories for humans; it is particularly noteworthy as a basic energy source for food, animal feed and industrial uses that can be produced on marginal agricultural land (Cock. 2019).

Cassava yields are hampered by susceptibility to biotic and abiotic stresses. In particular, bacterial and viral diseases can cause severe yield losses. Of note are cassava bacterial blight (CBB), cassava mosaic disease (CMD), and cassava brown streak disease (CBSD), all of which can cause catastrophic losses for growers (Lin *et al.*, 2019).

Most complex resistance traits are controlled by multiple loci and they are called as Quantitative Trait Loci (QTL) (Young. 1996).With QTL mapping, the roles of specific resistance loci can be described, race-specificity of partial resistance genes can be assessed, and interactions between resistance genes, plant development, and the environment can be analyzed (Castro *et al.*, 2015).

A number of QTL for disease resistance in Cassava have been identified and reported (Kayondo *et al.*, 2017; Soto *et al.*, 2017; Rabbi *et al.*, 2014; Masumba *et al.*, 2017 ; Kulembeka *et al.*, 2010; Lokko *et al.*, 2005; Mohan *et al.*, 2013) . However, these QTLs have been obtained from different populations, grown under different environmental conditions, and tested with different analytical methods; this has made the QTL positioning interval too large. Therefore, integration of original disease resistance QTLs from different studies needs to be performed to identify consistent and accurate loci.

In the present study, 61 QTLs associated with CMD and CBSD resistance genes in cassava were collected from 6 published studies and integrated using Meta-analysis method. Meta-QTL analysis which uses QTL from previous studies is useful in the identification of regions rich in QTLs (Said *et al.*,2015).

Currently QTLs associated with some crop related traits have been integrated with meta-analysis method and candidate genes related to them have been identified for root traits in maize (Guo *et al.*, 2018), virus resistance in maize (Wang *et al.*, 2016) and fatty acid content in soyabean (Qin *et al.*, 2018), Fusarium Head Blight Resistance in Bread Wheat (Venske *et al.*, 2019). In cassava, studies related to meta-analysis of QTLs is not done yet.

In the present study, a consensus map was constructed by integrating all the individual linkage map with a reference map. Cassava genetic linkage map developed by ICGMC was used as reference map which consist of 22,403 markers on 18 linkage groups with a map length of 2,412cM. The resulted consensus map consists of 22,621 markers on 18 linkage groups with a map length of 2,412.35 cM. 58 QTLs out of 61 initial QTLs were projected to this consensus map. The QTLs which were not projected consist of larger confidence interval length (more than 60cM) than the projected QTLs.

Biomercator V.4.2.software is used for QTL meta-analysis. This software has been used for QTL meta-analysis to integrate information from multiple QTL studies, detect MQTLs, and estimate their genetic effects precisely (Arcade *et al.*, 2004). A number of studies have reported on QTL meta-analysis using this software (Zhang *et al.*, 2014; Lan *et al.*, 2011). In the present study meta -analysis of 58 projected QTLs associated with CMD and CBSD resistance were done separately using Goffinet and Gerber method (Goffinet and Gerber. 2000). Meta-analysis has reduced the average confidence interval (CI) of QTLs to half as compared to the average confidence interval of projected QTLs. Only the QTLs with CI less than 2cM are selected for identifying candidate genes because only from narrowed genetic interval the potential candidate genes can be identified. Six MQTLs for CMD resistance and four MQTLs for CBSD-RN resistance were obtained within the confidence interval of 2cM. The MQTLs with narrowed confidence interval were distributed in Chromosomes 1,2,3,5,11,16,17 and 18. The MQTLs obtained for CMD resistance were from the QTL studies conducted in Tanzania (Masumba *et al.*, 2017) and from Nigeria (Rabbi *et al.*, 2014). The

MQTLs for CBSD-RN resistance were from studies conducted in Tanzania and Nigeria (Masumba *et al.*, 2017; Kulembeka *et al.*, 2010).

In this study MQTLs with physical length less than 1Mb were selected for mining candidate genes. Functional annotation of candidate genes of QTLs within MQTLCMD2.1, MQTLCMD3.1, MQTLCMD5.1 and MQTLCMD16.1 for CMD resistance and QTLs within MQTLCBSD2.1, MQTLCMD11.1, MQTLCMD17.1 and MQTLCMD18.1 were carried out.

The functional annotation of candidate genes of these major QTLs reveals that genes underlying them are involved in defence mechanism against pathogens. In this study the genes Manes.02G187500.1, Manes.03G046200.1, Manes.16G018100.1, Manes.16G018200.1, Manes.16G018300.1and Manes.16G018700.1from MQTLCMD2.1, MQTLCMD3.1 and MQTLCMD16.1 respectively and Manes.02G102100.1, Manes.17G005600.1, Manes.17G005600.1 from MQTLCBSD-RN2.1 and MQTLCBSD-RN17.1 respectively found to be associated with the protein kinases family. These protein kinases are involved in mediating defence related cell signalling in plants (Ding *et al.*, 2009).

Genes Manes.02G187700.1, Manes.16G020400.1 from MQTLCMD2.1 and MQTLCMD16.1 respectively and genes Manes.02G100700.1, Manes.02G102100.1 and Manes.17G005700.1 from MQTLCBSD-RN2.1 and MQTLCBSD-RN17.1 respectively were associated with Leucine rich repeat protein families. The LRR proteins are involved in specific protein–protein interactions. They are involved in the activation of defence genes after contact with gene products of pathogens (Andersen *et al.*, 2018).

The genes Manes.03G039100.1, and Manes.02G100800.1 from MQTLCMD3.1 and MQTLCBSD-RN2.1 respectively were found to be associated with protein families of cytochromeP450. These cytochromes P450 are involved in the biosynthetic pathway of major phytoalexins (chemicals

synthesized by plants to deter hostile organisms) which mediate defense response. (Jun *et al.*, 2015). Gene Manes.03G039700.1 is associated with Jasmonic acid biosynthesis. Jasmonic acids are members of jasmonates, studies explains that jasmonate pathway plays an essential role in wound-induced defense responses (Schilmiller *et al.*, 2005).

Genes Manes.05G085200.1 and Manes.05G085300.1 from MQTLCMD5.1 and Manes.18G063800.1 from MQTLCBSD-RN18.1 were associated with protein families of Zinc finger and TPR repeats and leucine zipper respectively. Zinc finger and leucine zipper are transcription factors. Recent studies shows that plant transcription factors have role in resistance of plants towards pathogens (Noman *et al.*,2019, Perotti *et al.*, 2017). Tetratrico Peptide Repeats (TPR) containing proteins are involved in plant hormones mediated signal transduction pathway which leads to defense responses of plants (Schapire *et al.*, 2006).

Gene Manes.16G019100.1 from MQTLCMD16.1 were associated with receptor serine/threonine kinases protein family. These Receptor like Kinases appear to play a central role in signalling during pathogen recognition, the subsequent activation of plant defense mechanisms, and developmental control (Afzal *et al.*, 2008).

Gene Manes.02G101400.1 from MQTLCBSD-RN 2.1 is associated with pectin esterase protein family. Pectin esterase is esterified form of pectin, which is a plant cell wall component and they are directly involved in plant defense (Fan *et al.*, 2017).

Genes Manes.11G058000.1 andManes.11G058200.1.1from MQTLCBSD-RN11.1 was found to be associated with Heat shock protein (HSP) and signal recognition protein respectively. HSPs play an indispensable role in plant

immunity as molecular chaperones in the quality control of intracellular resistance (R) proteins against potential invaders (Park et al., 2015).

In the present study the validation of the identified QTLs were done by using real time PCR. Primers specific to annotated candidate genes of QTLs qCMD3ACA and CMD3S were used for amplification. Results of real time PCR shows that the expression of genes in resistant variety was more than that of susceptible variety, thereby confirming the reliability and accuracy of identified QTLs from this study.

Marker Assisted selection is an important strategy for crop improvement and the success of MAS depends on the development of reliable markers. The available markers in Cassava have been predominantly generated from individual QTL investigations. The Meta-analysis approach, by integrating several QTLs should allow the development of markers with improved precision. The candidate genes, especially those co-located with the ranked MQTL could be strong candidates for developing perfect markers. The meta-QTL analysis conducted in this study accurately compared genomic positions of individual QTL identified from different studies and refined the confidence intervals of the main genomic regions associated with resistance.

SUMMARY

6. SUMMARY

The study entitled "META-ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA" was conducted at the ICAR-CTCRI, Sreekaryam, Thiruvanathapuram during October 2018 to August 2019. The objectives of this study were to combine the results of QTL detection studies conducted on Cassava Mosaic Disease resistance genes in Cassava, to integrate this data with genomic information of Cassava and to validate the results obtained using resistant and susceptible varieties.

A total of 61 QTLs associated with CMD and CBSD-RN resistance genes were collected from six published QTL mapping studies from 2003-2017. A consensus map was then constructed by integrating individual linkage maps of these selected QTLs with a genetic linkage map of cassava (ICGMC map). The resulted consensus map consists of 18 linkage groups comprising 22,621 markers and with map length of 2412.35cM.

A total of 58 QTLs from 61 initial QTLs were then projected to this consensus map. The QTLs which were not projected consist of larger confidence interval than the projected QTLs.

Meta-analysis of these 58 projected QTLs were done by using Biomercator V.4 software. Meta-analysis of QTLs associated with CMD and CBSD-RN resistance were done separately. A total of 11 Meta-QTLs (MQTLs) were obtained for CMD resistance and a total of 10 MQTLs obtained for CBSD-RN resistance. Then the MQTLs with confidence interval less than 2 cM were selected for identifying candidate genes. Six MQTLs associated with CMD resistance and Four MQTLs associated with for CBSD-RN resistance were found within the 2 cM confidence interval length.

From the selected MQTLs only the QTLs with less than 1Mb physical length were used for mining candidate genes. The candidate genes were detected using Phytozome database. A total of 73 candidate genes were identified for

QTLs associated with CMD resistance and a total of 51 candidate genes were identified for QTLs associated with CBSD-RN resistance.

Functional annotation of these candidate genes were carried out by using Phytomine. For QTLs related to CMD resistance 38 candidate genes were annotated and for QTLs related to CBSD-RN resistance 30 candidate genes were annotated. Defence related proteins like Leucine Rich Repeats (LRR), Jasmonic acid, Receptor like kinases, protein kinases, Heat shock protein and trascription factors like Zinc finger and Leucine Zipper were found to be associated with these annotated genes.

Validation of the identified QTLs was done by using real time PCR. Genes Manes.03G09700.1 and 1Manes.16G020400.1 were used for validation. In real time PCR, the expression of genes in resistant variety was found to be more than that of susceptible variety.

REFERENCES

7. REFERENCES

Afzal, A.J., Wood, A.J. and Lightfoot, D.A. 2008. Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Mol. Plant-Microbe Interactions*, 21(5): 507-517.

Akano, A.O., Dixon, A.G.O., Mba, C., Barrera, E. and Fregene, M. 2002. Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor. and Appl. Genet.* , 105(4): 521-525.

Alabi, O.J., Kumar, P.L. and Naidu, R.A. 2011. Cassava mosaic disease: a curse to food security in sub Saharan Africa.

Andersen, E., Ali, S., Byamukama, E., Yen, Y. and Nepal, M. 2018. Disease resistance mechanisms in plants. *Genes*, 9(7): 339.

Arcade, A., Labourdette, A., Falque, M., Mangin, B., Chardon, F., Charcosset, A. and Joets, J. 2004. BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinforma.* , *20*(14): 2324-2326.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, *3*(10): 3376.

Balyejusa Kizito, E., Rönnberg Wästljung, A.C., Egwang, T., Gullberg, U., Fregene, M. and Westerbergh, A. 2007. Quantitative trait loci controlling cyanogenic glucoside and dry matter content in cassava (Manihot esculenta Crantz) roots. *Hereditas*, *144*(4): 129-136.

Bellotti, A. and van Schoonhoven, A. 1978. Mite and insect pests of cassava. *Annual Review of Entomology*, *23*(1): 39-67.

Bellotti, A.C., 2008. Cassava pests and their management. *Encyclopedia of Entemol.*: 764-794.

Blair M, Fregene M, Beebe S, Ceballos H . 2007 Marker-assisted breeding in common beans and cassava. Marker-assisted selection: Current status and future perspectives in crops, livestock, forestry and fish. Food and Agriculture Organisation of the United Nations (FAO), Rome: 81–115

Boonchanawiwat, A., Sraphet, S., Whankaew, S., Boonseng, O., Smith, D.R. and Triwitayakorn, K. 2016. Mapping of quantitative trait loci underlying

resistance to cassava anthracnose disease. *The J. of Agric. Sci.*, *154*(7): 1209-1217.

Brown, J.K, Frohlich, D. and Rosell, R. 1995. The sweet potato or silver leaf whiteflies

Bull, Simon E., Rob W. Briddon, William S. Sserubombwe, Kahiu Ngugi, Peter G. Markham, and John Stanley. 2006. "Genetic diversity and phylogeography of cassava mosaic viruses in Kenya." *J. of General Virol.* 87, no. 10: 3053-3065.

Castro-Álvarez, F.F., William, M., Bergvinson, D.J. and García-Lara, S., 2015. Genetic mapping of QTL for maize weevil resistance in a RIL population of tropical maize. *Theor. and Appl. Genet.*, *128*(3): 411-419.

Cock, J.H., 2019. *Cassava: new potential for a neglected crop*. CRC Press.

Collard, B.C., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, *142*(1-2): 169-196.

Collins, N.C., Tardieu, F. and Tuberosa, R. 2008. Quantitative trait loci and crop performance under abiotic stress: where do we stand?. *Plant Physiol.* *147*(2): 469-486.

Conesa, A. and Götz, S. 2008. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int. J. of plant genomics*, *2008*.

Delourme, R., Falentin, C., Fomeju, B.F., Boillot, M., Lassalle, G., André, I., Duarte, J., Gauthier, V., Lucante, N., Marty, A. and Pauchon, M. 2013. High-density SNP-based genetic map development and linkage disequilibrium assessment in Brassica napus L. *BMC genomics*, *14*(1) : 120.

Ding, X., Richter, T., Chen, M., Fujii, H., Seo, Y.S., Xie, M., Zheng, X., Kanrar, S., Stevenson, R.A., Dardick, C. and Li, Y. 2009. A rice kinase-protein interaction map. *Plant Physiol.*, 149(3): 1478-1492.

Fan, H., Dong, H., Xu, C., Liu, J., Hu, B., Ye, J., Mai, G. and Li, H. 2017. Pectin methylesterases contribute the pathogenic differences between races 1 and 4 of Fusarium oxysporum f. sp. cubense. *Scientific reports*, 7(1): 13140.

Fanou, A.A., Amégnikin Zinsou, V. and Wydra, K. 2018. Cassava bacterial blight: A devastating disease of cassava. *Cassava. V. Waisundara, ed. IntechOpen, Erfurt, Germany. https://doi. org/10.5772/intechopen*, *71527*: 13-36.

Fanou, A.A., Zinsou, V.A. and Wydra, K. 2017. Cassava bacterial blight: A devastating disease of cassava. In *Cassava*. IntechOpen.

FAO [Food and Agriculture Organization].2016.

FAO [Food and Agriculture Organization].2017.

FAO [Food and Agriculture Organization].2018.

FAO [Food and Agriculture Organization].2019.

Ferguson, M., Rabbi, I., Kim, D.J., Gedil, M., Lopez-Lavalle, L.A.B. and Okogbenin, E. 2012. Molecular markers and their application to cassava breeding: past, present and future. *Tropical plant Biol.*, *5*(1): 95-109.

Fregene, M., Bernal, A., Duque, M., Dixon, A. and Tohme, J. 2000. AFLP analysis of African cassava (Manihot esculenta Crantz) germplasm resistant to the cassava mosaic disease (CMD). *Theor. and Appl. Genet. 100*(5): 678-685.

Fregene, M.A., Suarez, M., Mkumbira, J., Kulembeka, H., Ndedya, E., Kulaya, A., Mitchel, S., Gullberg, U., Rosling, H., Dixon, A.G. and Dean, R. 2003. Simple sequence repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. *Theor. and Appl. Genet.*, *107*(6): 1083-1093.

Fregene, M.A., Vargas, J., Ikea, J., Angel, F., Tohme, J., Asiedu, R.A., Akoroda, M.O. and Roca, W.M. 1994. Variability of chloroplast DNA and nuclear ribosomal DNA in cassava (Manihot esculenta Crantz) and its wild relatives. *Theor. and Appl. Genet.*, *89*(6): 719-727.

Fulton, T.M., Beck-Bunn, T., Emmatty, D., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D. and Tanksley, S.D. 1997. QTL analysis of an advanced backcross of Lycopersicon peruvianum to the cultivated tomato and comparisons with QTLs found in other wild species. *Theor. and Appl. Genet.*, *95*(5-6): 881-894.

Geiger HH, Heun M. 1989. Genetics of quantitative resistance to fungal diseases. *Annu. Rev. Phytopathol.* 27: 317–41

Giovannoni, J.J., Wing, R.A., Ganal, M.W. and Tanksley, S.D. 1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research*, *19*(23) : 6553-6568.

Goffinet, B. and Gerber, S. 2000. Quantitative trait loci: a meta-analysis. *Genetics*, *155*(1) : 463-473.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S. 2011. Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, *40*(D1): 1178-D1186.

Guo, J., Chen, L., Li, Y., Shi, Y., Song, Y., Zhang, D., Li, Y., Wang, T., Yang, D. and Li, C. 2018. Meta-QTL analysis and identification of candidate genes related to root traits in maize. *Euphytica*, *214*(12): 223.

Hahn, S.K. and Janet, K. 1985. Cassava: a basic food of Africa. *Outlook on Agriculture*, *14*(2): 95-99.

Islam, M., Ontoy, J. and Subudhi, P.K. 2019. A meta-analysis of Quantitative Trait Loci Associated with Seedling-Stage Salt Tolerance in Rice (Oryza sativa L.). *Plants*, *8*(2) : 33.

Izquierdo, P., Astudillo, C., Blair, M.W., Iqbal, A.M., Raatz, B. and Cichy, K.A. 2018. Meta-QTL analysis of seed iron and zinc concentration and content in common bean (Phaseolus vulgaris L.). *Theor. and Appl. Genet.*,*131*(8): 1645-1658.

Jansen, R.C. 1993. Interval mapping of multiple quantitative trait loci. *Genetics*, *135*(1): 205-211.

Jun, X.U., WANG, X.Y. and GUO, W.Z. 2015. The cytochrome P450 superfamily: key players in plant development and defense. *J.of Integrative Agric.*, 14(9): 1673-1686.

Kayondo, S.I., Del Carpio, D.P., Lozano, R., Ozimati, A., Wolfe, M., Baguma, Y., Gracen, V., Offei, S., Ferguson, M., Kawuki, R. and Jannink, J.L. 2018. Genome-wide association mapping and genomic prediction for CBSD resistance in Manihot esculenta. *Scientific reports*, *8*(1): 1549.

Kizito, E.B., Chiwona-Karltun, L., Egwang, T., Fregene, M. and Westerbergh, A. 2007. Genetic diversity and variety composition of cassava on small-scale farms in Uganda: an interdisciplinary study using genetic markers and farmer interviews. *Genetica*, *130*(3): 301-318.

Kowalski, S.P., Lan, T.H., Feldmann, K.A. and Paterson, A.H. 1994. Comparative mapping of Arabidopsis thaliana and Brassica oleracea chromosomes reveals islands of conserved organization. *Genet.*, *138* (2): 499-510.

Kulembeka, H.P.K., 2010. Genetic linkage mapping of field resistance to cassava brown streak disease in cassava *(Manihot esculenta* Crantz*)* landraces from Tanzania (Doctoral dissertation, University of the Free State).

Lanaud, C., Fouet, O., Clément, D., Boccara, M., Risterucci, A.M., Surujdeo-Maharaj, S., Legavre, T. and Argout, X. 2009. A meta–QTL analysis of disease resistance traits of Theobroma cacao L. *Molecular Breeding*, *24*(4): 361-374.

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

Lan-Lan, H., Kai-Zhen, Z., Qi-Bin, M.A., Hai, N. and Cun-Yi, Y. 2011. Integrated QTLs map of phosphorus efficiency in soybean by Meta-analysis. *Chinese J. of oil crop Sci.*, *33*(1).

Li, W.T., Liu, C., Liu, Y.X., Pu, Z.E., Dai, S.F., Wang, J.R., Lan, X.J., Zheng, Y.L. and Wei, Y.M. 2013. Meta-analysis of QTL associated with tolerance to abiotic stresses in barley. *Euphytica*, *189*(1): 31-49.

Lin, Y.R., Schertz, K.F. and Paterson, A.H. 1995. Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. *Genet.*, *141*(1): 391-411

Lin, Z.D., Taylor, N.J. and Bart, R. 2019. Engineering Disease-Resistant Cassava. *Cold Spring Harbor perspectives in biology.*

Liu, S., Hall, M.D., Griffey, C.A. and McKendry, A.L. 2009. Meta-analysis of QTL associated with Fusarium head blight resistance in wheat. *Crop Sci.*, *49*(6): 1955-1968.

Lokko, Y., Gedil, M. and Dixon, A., 2004. QTLs associated with resistance to the cassava mosaic disease. *Proc 4th Intl Crop Sci Congr.*

Lokko, Y., Danquah, E.Y., Offei, S.K., Dixon, A.G. and Gedil, M.A. 2005. Molecular markers associated with a new source of resistance to the cassava mosaic disease. *African J. of Biotechnol.*, *4*(9).

López, C. E., and Bernal, A. J. 2012. Cassava bacterial blight: using genomics for the elucidation and management of an old problem. *Trop. Plant Biol.* 5: 117–126

López, C.E., Quesada-Ocampo, L.M., Bohórquez, A., Duque, M.C., Vargas, J., Tohme, J. and Verdier, V., 2007. Mapping EST-derived SSRs and ESTs

involved in resistance to bacterial blight in Manihot esculenta. *Genome*, *50*(12): 1078-1088.

Lozano, J.C. and Booth, R.H. 1974. Diseases of cassava (Manihot esculenta Crantz). *PANS Pest Articles & News Summaries*, *20*(1): 30-54.

Maruthi, M.N., Hillocks, R.J., Mtunda, K., Raya, M.D., Muhanna, M., Kiozia, H., Rekha, A.R., Colvin, J. and Thresh, J.M. 2005. Transmission of Cassava brown streak virus by Bemisia tabaci (Gennadius). *J. of Phytopathol.*, *153*(5): 307-312.

Masumba, E.A., Kapinga, F., Mkamilo, G., Salum, K., Kulembeka, H., Rounsley, S., Bredeson, J.V., Lyons, J.B., Rokhsar, D.S., Kanju, E. and Katari, M.S. 2017. QTL associated with resistance to cassava brown streak and cassava mosaic diseases in a bi-parental cross of two Tanzanian farmer varieties, Namikonga and Albert. *Theor. and Appl. Genet.*, *130*(10): 2069-2090.

Miah, G., Rafii, M., Ismail, M., Puteh, A., Rahim, H., Islam, K. and Latif, M. 2013. A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. *Int. J. of Mol. Sci.*, *14*(11): 22499-22528.

Miles, C.M. and Wayne, M. 2008. Quantitative trait locus (QTL) analysis.

Mkumbira, J., Chiwona-Karltun, L., Lagercrantz, U., Mahungu, N.M., Saka, J., Mhone, A., Bokanga, M., Brimer, L., Gullberg, U. and Rosling, H. 2003. Classification of cassava into 'bitter'and 'cool'in Malawi: From farmers' perception to characterisation by molecular markers. *Euphytica*, *132*(1) : 7-22.

Mohan, C., Shanmugasundaram, P., Maheswaran, M., Senthil, N., Raghu, D. and Unnikrishnan, M. 2013. Mapping new genetic markers associated with CMD resistance in cassava (Manihot esculenta Crantz) using simple sequence repeat markers. *J.of Agric. Sci.*, *5*(5): 57.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G., Yano, M., Bhatia, C.R. and Sasaki, T. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breed.*, *3*(2): 87-103.

Monger, W.A., Seal, S., Cotton, S. and Foster, G.D. 2001. Identification of different isolates of Cassava brown streak virus and development of a diagnostic test. *Plant Pathol.*, *50*(6): 768-775.

Noman, A., Aqeel, M., Khalid, N., Islam, W., Sanaullah, T., Anwar, M., Khan, S., Ye, W. and Lou, Y. 2019. Zinc finger protein transcription factors:

Integrated line of action for plant antimicrobial activity. Microbial pathogenesis.

Nzuki, I., Katari, M.S., Bredeson, J.V., Masumba, E., Kapinga, F., Salum, K., Mkamilo, G.S., Shah, T., Lyons, J.B., Rokhsar, D.S. and Rounsley, S. 2017. QTL mapping for pest and disease resistance in cassava and coincidence of some QTL with introgression regions derived from Manihot glaziovii. *Frontiers in plant Sci.*, *8*: 1168.

Okogbenin E, PortoM, Egesi C, Mba C, Espinosa E, Santos L, Ospina C, Marin J, Barrera E, Gutierrez J, Ekanayake C, Iglesias C, Fregene MA. 2007 Marker-assisted introgression of resistance to cassava mosaic disease into Latin American germplasm for the genetic improvement of cassava in Africa. *Crop Sci.* 47:1895–1904

Okogbenin, E. and Fregene, M. 2002. Genetic analysis and QTL mapping of early root bulking in an F 1 population of non-inbred parents in cassava (Manihot esculenta Crantz). *Theor. and Appl. Genet.*, *106*(1): 58-66.

Okogbenin, E., Egesi, C.N., Olasanmi, B., Ogundapo, O., Kahya, S., Hurtado, P., Marin, J., Akinbo, O., Mba, C., Gomez, H. and de Vicente, C., 2012. Molecular marker analysis and validation of resistance to cassava mosaic disease in elite cassava genotypes in Nigeria. *Crop Sci.*, *52*(6): 2576-2586.

Pandey, A.K., Sharma, R., Singh, Y., Prakash, B. and Ahlawat, S.P.S. 2006. Evaluation of genetic variability in Kenkatha cattle by microsatellite markers. *Asian-australasian J. of Anim. Sci.*, *19*(12): 1685-1690.

Park, C.J. and Seo, Y.S. 2015. Heat shock proteins: a review of the molecular chaperones for plant immunity. *The plant Pathol. J.*, 31(4): 323.

Perotti, M.F., Ribone, P.A. and Chan, R.L. 2017. Plant transcription factors from the homeodomain-leucine zipper family I. Role in development and stress responses. *IUBMB life*, 69(5): 280-289.

Pita, J.S., Fondong, V.N., Sangare, A., Otim-Nape, G.W., Ogwal, S. and Fauquet, C.M. 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J. of Gen. Virol.*, *82*(3): 655-665.

Qi, Z., Sun, Y., Chen, L., Guo, Q., Liu, C., Hu, G. and Chen, Q. 2009. Meta-analysis of 100-seed weight QTLs in soybean. *Scientia Agricultura Sinica*, *42*(11): 3795-3803.

Qi, Z.M., Sun, Y.N., Wang, J.L., Zhang, D.W., Liu, C.Y., Hu, G.H. and Chen, Q.S. 2011. Meta-Analysis of 100-seed weight QTLs in soybean. *Agric. Sci. in China*, *10*(3): 327-334.

Qin, H., Liu, Z., Wang, Y., Xu, M., Mao, X., Qi, H., Yin, Z., Li, Y., Jiang, H., Hu, Z. and Wu, X. 2018. Meta- analysis and overview analysis of quantitative trait locis associated with fatty acid content in soybean for candidate gene mining. *Plant breeding*, *137*(2): 181-193.

Rabbi, I.Y., Hamblin, M.T., Kumar, P.L., Gedil, M.A., Ikpan, A.S., Jannink, J.L. and Kulakow, P.A. 2014. High-resolution mapping of resistance to cassava mosaic geminiviruses in cassava using genotyping-by-sequencing and its implications for breeding. *Virus research*, *186*: 87-96.

Said, J.I., Song, M., Wang, H., Lin, Z., Zhang, X., Fang, D.D. and Zhang, J. 2015. A comparative meta-analysis of QTL between intraspecific Gossypium hirsutum and interspecific G. hirsutum× G. barbadense populations. *Mol. Genet. and Genomics*, *290*(3): 1003-1025.

Saunders, K., Salim, N., Mali, V.R., Malathi, V.G., Briddon, R., Markham, P.G. and Stanley, J. 2002. Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: evidence for acquisition of a DNA B component by a monopartite begomovirus. *Virology*, *293*(1) .63-74.

Save, F.A.O. 2013. grow: Cassava. A guide to sustainable production intensification. Rome: Food and Agriculture Organization of the United Nation.

Schapire, A.L., Valpuesta, V. and Botella, M.A. 2006. TPR proteins in plant hormone signaling. Plant signaling & behavior, 1(5): 229-230.

Schilmiller, A.L. and Howe, G.A. 2005. Systemic signaling in the wound response. Current opinion in plant biology, 8(4): 369-377.

Sosnowski, O., Charcosset, A. and Joets, J. 2012. BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinforma.*, *28*(15): 2082-2083.

Soto Sedano, J.C., Mora Moreno, R.E., Mathew, B., Léon, J., Gómez Cano, F.A., Ballvora, A. and López Carrascal, C.E. 2017. Major Novel QTL for Resistance to Cassava bacterial blight identified through a multi-environmental analysis. *Frontiers in Plant Sci.*, *8*: 1169.

Storey, H.H. and Nichols, R.F.W. 1938. Studies of the mosaic diseases of cassava. *Annals of Appl. Biol.*, *25*(4): 790-806.

Storey, H.H. 1936. Virus diseases of East African plants. VI. A progress report on studies of disease of cassava. *East African Agric. J.*, *2*: 34-9.

Storey, H.H. 1939. Transmission of plant viruses by insects. *The Botanical Review*, *5*(4): 240.

SU, E., Rabbi, I.Y., Nwachukwu, E.C. and Adamu, C.L.Y. 2016. MAPPING OF QUANTITATIVE TRAIT LOCI ASSOCIATED WITH CASSAVA MOSAIC DISEASE (CMD) RESISTANCE. *Int. J. of Sci. Env. and Technol.*, 5(4) : 2637 – 2645

Swamy BM, Sarla N. 2011 Meta-analysis of yield QTLs derived from inter-specific crosses of rice reveals consensus regions and candidate genes. *Plant Mol Biol.* Rep 29: 663–680

Swamy, B.M., Vikram, P., Dixit, S., Ahmed, H.U. and Kumar, A. 2011. Meta-analysis of grain yield QTL identified during agricultural drought in grasses showed consensus. *BMC genomics*, *12*(1): 319.

Tappiban, P., Sraphet, S., Srisawad, N., Smith, D.R. and Triwitayakorn, K. 2018. Identification and expression of genes in response to cassava bacterial blight infection. *J. of Appl. Genet.*, *59*(4): 391-403.

Truntzler, M., Barrière, Y., Sawkins, M.C., Lespinasse, D., Betran, J., Charcosset, A. and Moreau, L. 2010. Meta-analysis of QTL involved in silage quality of maize and comparison with the position of candidate genes. *Theor.and Appl. Genet.*, *121*(8):1465-1482.

Van, K. and McHale, L. 2017. Meta-analyses of QTLs associated with protein and oil contents and compositions in soybean [Glycine max (L.) Merr.] seed. *Int. J. of Mol. Sci.*, *18*(6): 1180.

Vasconcellos, R.C., Oraguzie, O.B., Soler, A., Arkwazee, H., Myers, J.R., Ferreira, J.J., Song, Q., McClean, P. and Miklas, P.N. 2017. Meta-QTL for resistance to white mold in common bean. *PloS one*, *12*(2): 0171685.

Venske, E., Dos Santos, R.S., Farias, D.D.R., Rother, V., Maia, L.C.D., Pegoraro, C. and Costa De Oliveira, A. 2019. Meta-analysis of the QTLome of fusarium head blight resistance in bread wheat: refining the current puzzle. *Frontiers in plant Sci.10*: 727.

Veyrieras, J.B., Goffinet, B. and Charcosset, A. 2007. MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *BMC Bioinforma.*, *8*(1): 49.

Vidya, P., Nair, A.G., Ambu, V., Sreekumar, J. and Mohan, C., 2017. Identification of candidate genes responsible for CMD resistance in cassava and mapping the CMD resistance using SSR markers. *Skin* 3(9).

Wang, Y., Xu, J., Deng, D., Ding, H., Bian, Y., Yin, Z., Wu, Y., Zhou, B. and Zhao, Y. 2016. A comprehensive meta-analysis of plant morphology, yield, stay-green, and virus disease resistance QTL in maize (Zea mays L.). *Planta*, *243*(2): 459-471.

Wang, Y., Yao, J., Zhang, Z. and Zheng, Y. 2006. The comparative analysis based on maize integrated QTL map and meta-analysis of plant height QTLs. *Chinese Science Bulletin*, *51*(18): 2219-2230.

Weising, K., Nybom, H., Pfenninger, M., Wolff, K. and Kahl, G. 2005. *DNA fingerprinting in plants: principles, methods, and applications*. CRC press.

Winter, S., Koerbler, M., Stein, B., Pietruszka, A., Paape, M. and Butgereitt, A. 2010. Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa. *J. of Gen. Virol.*, *91*(5):1365-1372.

Wu, X.L. and Hu, Z.L. 2012. Meta-analysis of QTL mapping experiments. In *Quantitative Trait Loci (QTL)*: 145-171

Wu, Y., Huang, M., Tao, X., Guo, T., Chen, Z. and Xiao, W. 2016. Quantitative trait loci identification and meta-analysis for rice panicle-related traits. *Mol. Genet. and Genomics*, *291*(5): 1927-1940.

Young, N.D. 1996. QTL mapping and quantitative disease resistance in plants. *Annual review of Phytopathol.*, *34*(1): 479-501.

Zeng, Z.B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proceedings of the National Academy of Sciences*, *90*(23): 10972-10976.

Zeng, Z.B., 1994. Precision mapping of quantitative trait loci. *Genetics*, *136*(4): 1457-1468.

Zhang, H., Uddin, M.S., Zou, C., Xie, C., Xu, Y. and Li, W.X. 2014. Meta- analysis and candidate gene mining of low- phosphorus tolerance in maize. *J. of integrative plant Biol.*, *56*(3): 262-270.

Zhang, L.Y., Liu, D.C., Guo, X.L., Yang, W.L., Sun, J.Z., Wang, D.W. and Zhang, A. 2010. Genomic distribution of quantitative trait loci for yield and yield- related traits in common wheat. *J. of integrative plant Biol.*, *52*(11): 996-1007.

APPENDICES

APPENDIX I

CTAB RNA extraction buffer

Tris HCL (pH=8.0)	100mM	
EDTA	25mM	
NaCl	2M	
CTAB	2%	
Mercaptoethanol	2% (v/v)	Freshly prepared
PVP	2% (w/v)	

Prepared in DEPC treated water

APPENDIX II

TBE Buffer (10 X)

Tris base	107 g
Boric acid	55 g
0.5 M EDTA (pH 8.0)	40 mL

Final volume made up to 1000 mL with distilled water and autoclave before use.

APPENDIX III

70% Ethanol

100% Ethanol -70 mL

Distilled water- 30 mL

META-ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA

By

RESHMA PARVEEN J.

(2014-09-104)

Abstract of Thesis

Submitted in partial fulfilment of

the requirement for the degree of

B. Sc. - M. Sc. (INTEGRATED) BIOTECHNOLOGY

Faculty of Agriculture

Kerala Agricultural University, Thrissur

DEPARTMENT OF PLANT BIOTECHNOLOGY COLLEGE OF AGRICULTUREVELLAYANI, THIRUVANANTHAPURAM 695 522 KERALA, INDIA

2019

ABSTRACT

Plant disease resistance is a complex trait which is controlled by quantitative trait loci (QTLs). Several QTLs have been found for cassava disease resistance. Meta-analysis provides a simple, reliable, and economical method for integrating information from multiple QTL studies across various environmental and genetic backgrounds, detecting consistent QTLs powerfully and estimating their genetic positions precisely.

The study entitled "META-ANALYSIS OF QTLs ASSOCIATED WITH PEST AND DISEASE RESISTANCE GENES IN CASSAVA" was conducted at the ICAR-CTCRI, Sreekaryam, Thiruvanathapuram during October 2018 to August 2019. The objectives of this study was to combine the results of QTL detection studies conducted on Cassava Mosaic Disease (CMD) resistance genes in cassava, to integrate this data with genomic information of cassava and to validate the results obtained using resistant and susceptible varieties.

61 QTLs from 6 different published QTL mapping studies related to Cassava Mosaic Disease (CMD) and Cassava Brown Streak Disease (CBSD) were selected. A consensus map was constructed by integrating individual linkage maps of these QTLs with a reference map (Cassava genetic linkage map). 58 QTLs from the initial 61 QTLs were then projected to this consensus map.

Using Biomercator V.4 software meta-analysis of these projected QTLs were carried out. Meta-analysis of CMD and CBSD-RN resistance QTLs were done separately. 11 MQTLs for CMD resistance and 10 MQTLs for CBSD-RN resistance were obtained. For mining candidate genes and for identifying their functional information the QTLs within the MQTLs with confidence interval less than 2cM and physical length less than1Mb were selected.

A total of 38 candidate genes associated with CMD resistance and 30 genes associated with CBSD-RN resistance were annotated. Defence related protein and transcription factors were found to be associated with these genes. Validation of the identified QTLs was done by using real time PCR and the genes were found to be expressed in resistant variety of cassava.

The MQTLs found in this study that have small physical and genetic intervals are useful for marker-assisted selection for developing disease resistance varieties of cassava.