Biniam Tesfagaber Hizbai

Comparative mapping of QTLs affecting oil content, oil composition, and other agronomically important traits in Oat (*Avena sativa* L.)

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
University of Ottawa in partial fulfillment of the requirement for the M.Sc. degree in the Ottawa-Carleton Institute of Biology

Thèse soumise à la Faculté des études supérieures et postdoctorales
Université d'Ottawa en vue de l'obtention de la maîtrise ès science
L'Institut de biologie d'Ottawa-Carleton

August, 2012

Abstract:

Groat oil content and composition are important quality traits in oats (*Avena sativa* L). These traits are controlled by many genes with additive effects. The chromosomal regions containing these genes, known as quantitative trait loci (QTL), can be discovered through their close association with markers. This study investigated total oil content and fatty acid components in an oat breeding population derived from a cross between high oil ('Dal') and low oil ('Exeter') parents. A genetic map consisting of 475 DArT (Diversity Array Technology) markers spanning 1271.8 cM across 40 linkage groups was constructed. QTL analysis for groat oil content and composition was conducted using grain samples grown at Aberdeen, ID in 1997. QTL analysis for multiple agronomic traits was also conducted using data collected from hill plots and field plots in Ottawa, ON in 2010. QTLs for oil content, palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2) and linolenic acid (18:3) were identified. Two of the QTLs associated with oil content were also associated with all of the fatty acids examined in this study, and most oil-related QTL showed similar patterns of effect on the fatty acid profile. These results suggest the presence of pleiotropic effects on oil-related traits through influences at specific nodes of the oil synthesis pathway. In addition, 12 QTL-associated markers (likely representing nine unique regions) were associated with plant height, heading date, lodging, and protein content. The results of this study will provide information for molecular breeding as well as insight into the genetic mechanisms controlling oil biosynthesis in oat.

Résumé :

La teneur en huile et la composition du grain sont des traits importants dans l'avoine (Avena sativa L). Ces traits sont contrôlés par plusieurs gènes aux effets additifs. Les régions chromosomiques contenant ces gènes, aussi connues comme locus de caractères quantitatifs (LCQ), peuvent être découvertes par leur association de proximité avec des marqueurs. Cette étude s'est concentrée sur la teneur totale en huile et les composantes des acides gras retrouvés dans une population d'avoine cultivée dérivée d'un croisement entre un parent riche en huile (« Dal ») et un antre pauvre en huile (« Exeter »). Une carte génétique composée de 475 marqueurs DArT (Diversity Array Technology) couvrant 1271.8 cM comprenant 40 groupes de liaison fut construite. L'analyse de LCQ sur la teneur en huile et la composition du grain fut menée en utilisant des échantillons cultivés à Aberdeen, ID en 1997. Une analyse LCQ sur de multiples traits agronomiques fut aussi menée en utilisant des données provenant de parcelles de champ et de parcelles buttes à Ottawa, ON en 2010. Des LCQs furent identifiés pour la teneur en huile, acide palmitique (16:0), acide stéarique (18:0), acide oléique (18:1), acide linoléique (18:2) ainsi que pour l'acide linolénique (18:3). Deux des LCQs liés à la teneur en huile furent aussi liés à tous les acides gras étudiés dans cette étude et la majeure partie des LCQs associés à l'huile ont démontré des modèles similaires d'effets sur le profil des acides gras. Ces résultats suggèrent la présence d'effets pléiotropiques sur des traits reliés à l'huile en influençant à certaines jonctions de la voie de synthèse de l'huile. De plus, 12 marqueurs liés aux LCQs (représentant probablement neuf régions uniques) furent associés avec la hauteur des plants, la date de floraison, la verse et la teneur en protéines. Les résultats de cette étude fournira de l'information pour la sélection moléculaire ainsi qu'une vue sur les mécanismes génétiques contrôlant la biosynthèse de l'huile chez l'avoine.

Acknowledgements:

First and foremost I would like to thank my supervisors Dr. Nicholas Tinker and Dr. Douglas Johnson for accepting me as their student. They offered me invaluable assistance, support and guidance throughout the course of this project. Dr. Tinker constantly challenged me to go further and do better. Dr. Johnson provided me with both moral and material support and career advice.

I am indebted to Charlene Wight and Kyle Gardner for their help with day to day lab activities and with preparing and editing my manuscript. They were great sounding boards for my ideas, but above all great friends.

Thanks to Dr. Judith Frégeau-Reid for granting me access to her labs and for providing protocols for oat quality analysis. I gratefully acknowledge the technical assistance provided by Eglantine Imbeault and Dorothy Sibbitt. I would also like to thank to Dr. Stephen Molnar for the time spent editing my manuscript and for useful discussions.

I would like to express my undying gratitude to Sandra McSheffrey and Tim Wehrle who helped me overcome impossible odds to come to Canada and study. Without their dedication and help the realization of my dreams would have been a much tougher proposition. They are the kind of friends that one can only hope for.

I would like to express my gratitude to Philippe Couroux for his help in solving IT related issues and for translating my abstract. Many people have helped me during these demanding years. I thank them all for their kindness and encouragement.

Finally, I would like to express my everlasting gratitude and love to my daughter, Ariam, and my wife, Saron. Ariam helped keep things in perspective for me. Saron was patient, caring and never wavered in her support throughout the course of my studies. I dedicate this thesis to them.

Scope and outline of the thesis

This thesis has been prepared in a manuscript based format. The manuscript has been accepted in the Plant Genome journal and it has been presented here in its original format. The thesis is organised in five chapters.

Chapter one: This is a general introductory chapter that explains the importance of the oat crop and provides the rationale for conducting the study.

Chapter two: This is an overview of genetic mapping and QTL analysis. This chapter provides the basic principles and methods of the genetic approach used in this study as well as the salient requirements for conducting QTL analysis. It is provided as general background for readers who are less familiar with these principles and procedures.

Chapter three: This is a review of literature relevant to oil and fatty acid composition in oat and other plant species, as well as a general assessment of QTL analysis in oat. These topics are highly relevant to the objectives of this thesis, and this review provides a more detailed background than what is presented in the submitted manuscript.

Chapter four: In this chapter the manuscript is presented as submitted to The Plant Genome journal. The tables and figures have been placed in the main text. The figures and tables in this chapter have been labeled slightly differently from the journal manuscript in order to take into account the figures used in the literature review and to maintain continuity. Supplemental tables and figures that are relevant to the study but were not included in the main body of the manuscript have been provided at the end of the thesis. A large amount of data which includes raw marker data and supplemental tables is provided digitally on a CD.

Chapter five - This section presents additional detailed discussion and conclusions, including discussion of results that were omitted from the submitted manuscript.

Table of Contents

List of Figures

List of tables

List of abbreviations

ACCase	Acetyl coenzyme-A
ACP	Acyl Carrier protein
AFLP	Amplified fragment length polymorphism
BC	Backcross
CAPs	Cleaved amplified polymorphic sequence
CoA	Coenzyme-A
DAG	Diacylglycerol
DArT	Diversity array technology
DGAT	Diacylglycerol acyltransferase
DH	Double-haploid
ER	Endoplasmic reticulum
ESTs	Expressed sequence tag
FA	Fatty acid
FAD	Fatty acid desaturase
FAE	Fatty acid elongase
FAS	Fatty acid synthase
G3P	Glycerol-3-phosphate
GPAT	G3P-acyltransferase
INDEL	Insertion or deletion
KAS	B-ketoacyl-ACP synthase
LEC2	Leafy cotyledon 2
LPA	Lysophosphatidic acid
LPAAT	Lysophosphatidic acid acyltransferase
MAS	Marker assisted selection
PA	Phosphatidic acid
PAP	Phosphatidic acid phosphatase
QTL	Quantitative trait locus/loci
RAPD	Random amplified polymorphic DNA
RF	Recombination frequency
RFLP	Restriction fragment length polymorphism
RIL	Recombinant inbred lines
SAD	Stearoyl-ACP Δ9-desaturase
SCAR	Sequence characterized amplified region
sCIM	Simplified composite interval mapping
SIM	Simple interval mapping
SNPs	Single nucleotide polymorphism
SSR	Simple sequence repeats
TAG	Triacylglycerol

Chapter 1- Introduction

1.1 Oat and its importance

Oat (*Avena sativa* L.) is a cereal grain that belongs to the Gramineae (Poaceae) family. Oat and its wild relatives occur at three ploidy levels: diploid, tetraploid, and hexaploid, with a base chromosome number of 7. The center of origin of oat is unknown, although the Mediterranean basin and the Middle East are thought to be the most likely possibilities. The domestication of oats is a relatively recent phenomenon, occurring much later than the domestication of other important cereals such as wheat and barley. The primary cultivated oat (*A. sativa*) is a hexaploid of 2n=6x = 42 originating as an aggregation of three diploid genomes: AA, CC and DD. (Rines *et al.,* 2006).

Oats are adapted to a wide range of soil types, thus temperature and moisture conditions are the usual factors that limit where oats are grown (Gibson and Benson, 2002). Oats are chiefly a European and North American crop used for both human and animal feed. These areas have the cool, moist climate to which oats are best adapted. Oat production ranks sixth in world grain production following corn, wheat, rice, barley, and sorghum. Russia, Canada, Australia, Poland, USA and Spain are the leading oat producing countries (Figure1-1). The production of oats worldwide has been very stable, and long term forecasts show it to remain at these levels with little growth. Since 2005-2006, the world production of oats has averaged around 24.6 million tonnes (Mt). Canada accounts for 15% of the annual world oat production, making it the second largest producer after the Russian federation. Canada is, however, the largest exporter of oats, accounting for roughly 60% of global exports (Strychar, 2011).

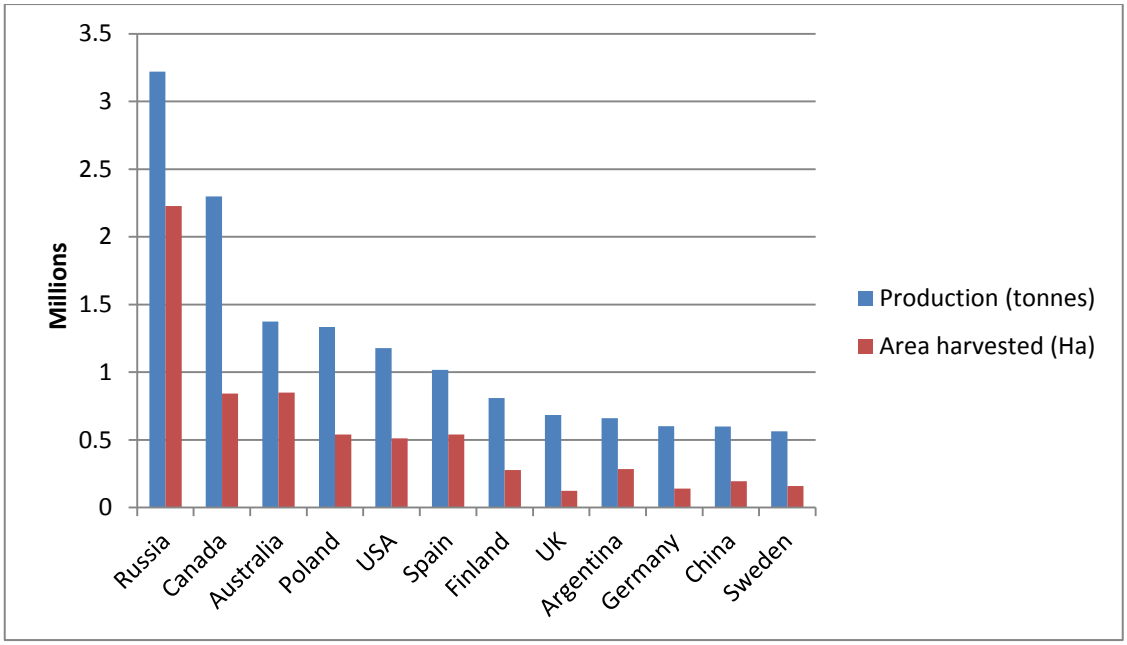


Figure 1- 1 *The leading producers of oat ranked by total production in million tonnes. (Figure constructed using data obtained from FAOSTAT, 2012).*

A combination of good quality traits has given oats a reputation as a human health-food. The high nutritional value of oat is mainly attributed to its composition of protein, fibre and lipid (Gibson and Benson, 2002). The oat grain has the highest protein level among cereals with 12-20% in the de-hulled kernel (groat) and 9-15% in the whole grain (Peterson, 1992). Furthermore, it has a superior amino acid profile compared with wheat, barley or maize with higher levels of all essential amino acids (Rines *et al.*, 2006). *β*-glucan, a high-molecular-weight water-soluble fibre, lowers total plasma cholesterol and LDL cholesterol levels of blood and balances the glucose and insulin contents of serum after meals, thus reducing the risks of cardiovascular diseases in humans (Meydani, 2009).

There are three different perspectives on what makes a good oat variety. Farmers are interested mainly in the agronomic attributes that determine yield per unit area with a given set of inputs. The milling industry is concerned primarily with seed traits that influence the weight of good-quality de-hulled seed per tonne of purchased crop. The milling industry also seeks to purchase oat varieties

with high *β*-glucan, since they must meet a minimum level of this trait to market oat products with a health claim (FDA, 1997). Consumers are interested in flavour, quality, and nutritional composition, but they have little direct influence on varieties that are produced. Because of consumer awareness and improved product labelling, consumer requirements and nutritional value may become increasingly important in setting objectives in a breeding program.

Oil content in oats is of significant importance from the nutritional standpoint. The high oil content is preferred for animal feed as the caloric content of the oil provides high energy. In contrast, oats for human food use should have low oil to ensure low calorific contents and reduce the potential for rancidity (Youngs, 1978). As a result, a breeding program may need to select among two different oil content profiles depending on the end use of its intended varieties. Similarly there is a dichotomy in selecting for fatty acid (FA) composition. The composition of oat oil with its high proportion of unsaturated FAs is favorable for nutrition. However, high proportions of unsaturated FAs reduce the shelf life of oat products. Therefore, identifying the genetic factors and mechanisms controlling these traits is important for breeding programs.

The oil content, primarily in the form of triacylglycerols (TAGs), in oat groats is a quantitative trait controlled by multiple genetic factors (Thro and Frey, 1985). The method used for inferring the presence and position of such genetic factors in the genome is based upon analysis as quantitative trait loci (QTLs). The basic principles and methods of QTL analysis are discussed in greater detail in Chapter Two. Rapid developments in genome sequencing, genetic marker technology and computing power have led to a proliferation of genetic maps and the identification of QTLs for many important traits of most crops. Several studies have identified QTLs for oil content in oats (Kianian *et al.*, 1999; Tanhuanpää *et al.*, 2010; Zhu *et al.*, 2004). QTLs for FAs in oats, however, have not been

reported. Many of the reported oil content QTLs in oat show consistency among at least two different studies, but additional studies in other germplasm are required to provide a more complete inventory of these QTL regions and to understand each QTL's sensitivity to environment and genetic background. Furthermore, there is a need to identify and map more genetic markers that are common among these populations in order to provide more accurate and systematic comparisons of chromosomal location of QTLs. To date, most of the maps in oat have been generated using a variety of markers based on sequence characterized amplified region (SCAR), simple sequence repeats (SSR), amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), and restriction fragment length polymorphism (RFLP) technologies. The lack of common genetic markers among different experiments has made it difficult to compare the positions of QTLs among oat maps. A recently developed DArT marker platform for oat and a revised DArT-based reference map of the 'Kanota'(*Avena byzantine* C. Koch) x 'Ogle' (*Avena sativa* L.) (KxO) mapping population (Tinker *et al.,* 2009) provide new opportunities for rapid map development and comparative mapping in oat.

The objectives of this thesis were

(i) to review the literature about the current state of knowledge in oat oil content and composition, and the molecular and genetic factors regulating these traits in oat;

(ii) to develop a DArT-based linkage map based on progenies from a cross between oat parents with high *vs.* low groat oil content;

(iii) to measure groat oil contents, FA profiles, and additional agronomic traits in the progenies;

(iv) to identify major QTLs affecting these traits;

(v) to conduct comparative mapping with locations of other QTLs and candidate genes in oat**.**

Chapter 2 - An overview of genetic mapping and QTL analysis

2.1 Genetic Mapping and QTL Analysis

A large number of traits of economic importance are quantitative in nature. Unlike qualitative traits, which are governed by single major genes, quantitative traits show continuous variation owing to the many genes that govern them. Quantitative traits also show a greater degree of influence from the environment. The regions containing these genes, known as quantitative trait loci (QTL), can be discovered through their close association with genetic markers. In QTL analysis, the association between observed trait values and presence or absence of marker alleles that have been mapped onto a linkage map (discussed in section 2.1.1.4) is analysed. When it is significantly clear that the observed correlation is not random, the presence of a QTL is inferred. Once the presence of a QTL is inferred, its position and effect are estimated based on the nearest one or two markers with the strongest relationship to the trait.

In summary, the objectives of QTL mapping are

i) To infer the presence of chromosomal regions that affect a trait of interest;

ii) To estimate the effects of alternate alleles in each specific region.

2.1.1 Requirements for QTL mapping

There are certain requirements for QTL mapping; these include:

1. A suitable mapping population of segregating progeny;
2. Measurement of polymorphic genetic markers in the progeny;
3. A saturated linkage map;
4. Phenotypes of mapping progeny.

2.1.1.1 Mapping populations

With natural populations, consistent marker-QTL association exists only in situations where the marker is very closely linked to a QTL. A new body of research called association mapping is devoted to identifying QTL in natural populations (reviewed by Stich and Melchinger, 2010). However, most QTL analysis to date has been carried out in structured bi-parental populations segregating for traits of interest. The parents used to generate mapping populations should be genetically divergent enough to exhibit sufficient polymorphism. In addition they should have a wide contrast in terms of the specific trait or traits under investigation. The larger the phenotypic difference between the parents the greater the chances of detecting the QTLs affecting the trait.

The choice of a mapping population depends on the objectives of the study as well as availability of time and resources. In self-pollinating species, mapping populations originate from parents that are highly homozygous (inbred). In cross pollinating (outcrossing) species, the situation is more complicated since most of these species do not tolerate inbreeding (Collard *et al.,* 2005). F_2 populations, developed by selfing of F_1 hybrids between two parents, and backcross (BC) populations, derived by crossing the F_1 hybrid to one of the parents, are the simplest types of mapping populations. Their main advantages are that they are easy to construct and require only a short time to produce (Semagn *et al.,* 2006). However, each genotype is represented by a single progeny plant, so replications over time or space cannot be carried out. Recombinant inbred lines (RILs) are developed by single-seed descent from individual plants of an F_2 population. Only one seed is harvested from each F_2 plant to grow the F3 generation. The remaining seeds are discarded. Each seed constitutes a line. The same process is repeated for six to eight generations. At this point, all of the seed from an individual plant is bulked. However, advancing the population for six or more generations is expensive in time and labor. An alternative is the use of a double-haploid (DH)

population. A DH population is produced by regenerating plants using tissue culture techniques from pollen grains or microspores whose chromosomes have been artificially doubled. This technique can only be utilized on species that are amenable to tissue culture manipulations. The major advantages of RILs and DH populations are that they produce homozygous or 'true-breeding' lines that can be multiplied and reproduced without genetic change occurring.

The size of a mapping population is an important factor that determines the power of QTL detection, defined as the probability of detecting a QTL at a given level of statistical significance (Manly and Olson, 1999). Small population size often results in the detection of few QTLs with large phenotypic effects (Beavis, 1998; Melchinger *et al.,* 1998; Schon *et al.,* 2004). Melchinger *et al.* (1998) evaluated the power of QTL detection using two independent samples of different size (107 vs 344) from a maize population. The total number of QTLs detected for all traits in the larger population (N=344) was almost triple to that of the numbers detected in the smaller population (N=107). While a large population size permits the detection of QTLs with smaller effects, for practical purposes identification of QTLs with large effects would serve the purpose of marker assisted selection (MAS).

2.1.1.2 Selection of genetic markers

A genetic marker can be described as a variation in a gene or DNA sequence that can be observed directly (such as a morphological marker) or indirectly through any assay. These differences originate as mutations or other genetic events that cause single nucleotide polymorphisms (SNPs), insertions or deletions (INDELs) or rearrangements (translocations or inversions).

Different types of marker systems such as phenotypic, cytological, and protein assays have been used for genetic research before the development of DNA based markers. However, they have limitations which hindered their wide applicability in genetic mapping studies (Langridge and Chalmers, 2005). DNA-based markers gained popularity because they have advantages including:

1. they are phenotypically neutral and seldom influenced by environments;
2. they are widely distributed through the whole genome, enabling construction of genome wide high density maps;
3. they have a high level of allelic polymorphism;
4. they are highly reproducible;
5. only small amounts of plant tissue are needed;
6. they are amenable to automation, providing opportunities for undertaking complex tasks with reduced time and cost.

Selecting the type of marker to use is important since there are a wide variety of DNA markers in use today. RFLPs, SSRs, RAPDs, AFLPs, SSRs, DArTs, ESTs (expressed sequence tag), CAPs (cleaved amplified polymorphic sequence), and SNPs (single nucleotide polymorphism) have all been used for map construction (Semagn *et al.*, 2006). Each system has advantages and disadvantages. Selection of which marker type to use depends on the species under investigation, the objective of the study and on the resources available. Here, the most widely used molecular marker technologies are described.

i) Restriction Fragment Length Polymorphism (RFLP)

RFLPs are differences in restriction fragment lengths caused by SNPs that change restriction endonuclease recognition sites, or INDELs between restriction sites that change their size. In RFLP, DNA polymorphism is detected by hybridizing labelled DNA probes to digested DNA that has been

separated by size and immobilized on a membrane (Southern 1975). Marker alleles are identified by size differences of the restriction fragments to which these probes hybridize. Developments in RFLP technology enabled the first large scale efforts to produce genetic maps in plants (Helentjaris *et al.*, 1986; Burr *et al.*, 1988). The RFLP marker type is considered superior due to its co-dominant inheritance and high reproducibility (Agrawal *et al.*, 2008). There are, however, several limitations to RFLP such as

a) It requires high quantity and quality of DNA;

b) It depends on the development of specific probe libraries for the species;

c) The technique is not suitable for automation;

d) The level of polymorphism is low;

e) It is time consuming and laborious;

f) It may require radioactively labeled probes;

g) RFLP probes must be physically maintained and thus are difficult to share between laboratories.

Efforts to overcome these limitations have led to the development of some of the methods that follow, all of which are based on PCR amplification.

ii) Random Amplified Polymorphic DNAs (RAPDs).

RAPD markers are DNA polymorphisms produced by rearrangements or deletions of genomic sites that are at or between the binding sites of oligonucleotide PCR primers (Welsh and McClelland, 1990; Williams *et al.,* 1990). The RAPD technique utilizes a randomly chosen set of PCR primers containing 8-10 nucleotides which are used to amplify DNA fragments in a low-stringency PCR (35-45°C). These amplified fragments are separated on conventional agarose gels, and RAPDs are identified by the presence or absence of a particular fragment (i.e. band).

The RAPD technique has several advantages: it is easy to use, economical, doesn't require sequence information, and only a small amount of DNA is required to generate the markers. However, many

RAPD markers show poor reproducibility, which has made this technique less popular for current genetic mapping studies (Nguyen and Wu, 2005).

iii) Amplified Fragment Length Polymorphism (AFLP)

AFLPs are molecular markers obtained by selective PCR amplification of restriction fragments from genomic DNA. The AFLP marker system was developed by Vos *et al.* (1995) to overcome the limitations of RAPD markers. Generating AFLP markers involves three steps: (i) restriction of the DNA and ligation of oligonucleotide adapters, (ii) selective amplification of sets of restriction fragments with specific restriction enzymes, and (iii) gel analysis of the amplified fragments. AFLPs are dominant markers and polymorphisms are scored as presence/absence of data rather than determination of sizes at various loci. AFLP markers are highly polymorphic, reproducible, and do not require prior sequence information, making them a good choice for genetic mapping. They generally provide good genome coverage and have been used for the construction of whole-genome and high resolution maps in many species (Chalmers *et al.*, 2001; Groh *et al.*, 2001; Lionneton *et al.*, 2002). The main criticism of the AFLP technique is that scoring of markers relies on gel-based systems which limits its throughput and it is a labour intensive procedure (Brugmans *et al.* 2003; Polanco *et al.* 2005).

iv) Microsatelites (SSR)

Microsatellites, also known as short tandem repeats or simple sequence repeats, are repetitions of very short (usually one to six) nucleotide motifs. These repeats are interspersed in many eukaryotic genomes (Tautz and Renz 1984). Microsatelites are highly polymorphic due to the variation in the number of tandemly repeated units (Farook and Azam, 2002). SSR loci are amplified by PCR using pairs of oligonucleotide primers specific to unique DNA sequences flanking the SSR sequence. Their codominant inheritance, high abundance, high allelic diversity, locus specificity and their reproducibility have made microsatellite highly popular markers (Gupta *et al.,* 1999; Agrawal *et al.*, 2008). However, SSRs are expensive to discover, and the technique has low throughput. As a result

of these limitations, microsatellites have been used extensively in only a few agriculturally important crops (Squirrell *et al.,* 2003).

v) Single Nucleotide Polymorphism (SNP)

SNPs are single nucleotide positions in a given stretch of genomic DNA at which variations between different individuals occur within a species. The change in nucleotides can be a result of a mutational transition (C ↔ T, G ↔ A) or a transversion (C ↔ A, C ↔ G, G ↔ T, A ↔ T), the former being more common. Insertion or deletion (INDELs) are generally not classified as SNPs (Brooks, 1999), although they may often be detected by technologies used to detect SNP variants. SNPs are the most abundant form of DNA sequence polymorphisms (Collins *et al.,* 1997; Van *et al.,* 2005). They are widely distributed throughout genomes, although their occurrence and distribution varies among species (Agrawaal *et al.,* 2008). The binary (bi-allelic) character and stability from generation to generation make SNPs amenable to automated, high throughput genotyping. This makes them attractive tools for QTL mapping studies. The development of DNA chips to analyse genotypes for single nucleotide polymorphisms (SNPs) has increased their utility. However, the technology depends on intensive genomic sequencing which is expensive and time consuming. Hence, SNPs have not yet been widely used in crops where large-scale public or commercial investment has not occurred.

vi) Diversity Arrays Technology (DArT)

DArT markers were developed to overcome the shortcomings of other marker systems. DArT assays are able to detect thousands of markers in one reaction on a microarray platform. This allows the potential for increasing marker density within a short time and at low cost (Jaccoud *et al.,* 2001; Mace *et al.,* 2008). DArT markers are based on a complexity-reduced reproducible subset of a genome, called a representation, which is generated using restriction enzymes and size selection (Akbari *et al.,* 2006). The genomic representation obtained from pooled samples of many different accessions is cloned and individual inserts are printed on to a microarray resulting in a “discovery

array." Genomic representations from genomes of individual accessions are prepared using the same complexity reduction methods and these are hybridized to the discovery array. Each individual fragment of a genomic representation will only hybridize to matching fragments on the genotyping array, thereby displaying a unique hybridisation pattern. Polymorphic clones (DArT markers) are identified on the basis of the intensities of their hybridization signals. After an initial marker discovery process, clones that do not reveal polymorphisms may be eliminated from the array to provide a more cost-effective "typing array" for further work.

Sequencing is not necessary for identification and scoring of polymorphic DArT markers, which makes the method applicable to all species regardless of how much DNA sequence information is available for that species. This is particularly useful for crop species such as oats in which genetic research is lagging due to lack of resources. Although sequencing is not a pre-requisite, DArT markers provide sequence-ready clones that can be useful for further genetic analysis, including the discovery of SNPs.

2.1.1.3 Genotyping

Genotyping involves the identification of polymorphic molecular markers that reveal differences between the two parents. The polymorphic markers are screened across either the entire mapping population, including the parents, or a subset of the population selected on the basis of phenotype (Lander and Botstein, 1989) to identify segregation patterns of each marker. Genotyping an entire mapping population with markers distributed across the whole genome is more reliable but requires more resources (Xu and Crouch, 2008).

2.1.1.4 Linkage Mapping

Linkage mapping is the process of arranging markers in groups, then placing them in order and estimating inter-marker distances. This is done based on the principle that genes (markers) segregate following chromosome recombination during meiosis. Most often, when two loci are linked on the same chromosome, the alleles originating from one parent remain associated in the resulting gametes and progeny. However, during meiosis, a homologous chromosome pair may cross over and exchange sections, such that parental alleles become dissociated (recombine). The distance between two loci is estimated by counting the number of recombinant gametes, and expressing this as a proportion of the total gametes. If the recombination frequency (RF) is statistically less than 50% (observed when two loci are on different chromosomes) then the loci are said to be linked. Recombination frequencies are not additive and mathematical correction factors, known as mapping functions, are used to translate recombination fractions to additive genetic distances which are expressed as centimorgans (cM). There are two widely used mapping functions. The first one proposed by Haldane (1919) assumes no cross-over interference. However, this is not reliable for distances of over 15cM. The second mapping function known as the Kosambi mapping function (Kosambi, 1943) takes into account some crossover interference.

Linkage groups then are formed by clustering together loci that are linked. Following this, the order of loci within linkage groups is estimated. This is the most computationally challenging step because it involves testing all (or many) possible orders to find the order that is most probable. Orders can be optimized based on the cumulative distances that they produce, and/or by minimizing the number of crossovers that must have occurred to produce the observed data. The reader is directed to Kearsy and Pooni (1996) for a more detailed explanation of the principles and practice of linkage mapping.

2.1.1.5 Phenotyping of the mapping population

The basic phenotypic data required for QTL mapping are the estimates of phenotypic performance of individuals across environments. The accuracy and precision of phenotyping influences the quality of the QTL mapping results. A reliable QTL map can only be produced from reliable phenotypic data. The phenotype of quantitative traits can be affected by environment. This can pose a problem for QTL analysis, especially when this influence is large. Trials across multiple environments are often analysed in order to account for the genotype-environment interaction. This ensures that results will have relevance beyond just the environment where a QTL is discovered.

The power and precision of a QTL analysis is determined by many factors including population size, genome coverage of polymorphic markers, and phenotype accuracy. Due to the availability of relatively cheap, high throughput techniques, the cost of genotyping is now less of a limiting factor (Semagn, 2010). However, the cost and logistics involved in phenotyping often impose limits on population size.

2.1.2 Methods of detecting QTLs

Three widely-used methods for detecting QTLs are single-marker analysis, simple interval mapping (SIM) and composite interval mapping (CIM) (Liu, 1998; Tanksley, 1993). Single-marker analysis (also single-point analysis) is the simplest method for detecting QTLs. In this method, an independent test for QTL/trait association is performed for each marker locus using analysis of variance (ANOVA), t-test, or linear regression. This method does not require a complete linkage map and can be performed with basic statistical software programs. Although it is computationally simple, this method has some drawbacks: (i) the likelihood of detecting a QTL is significantly reduced as the distance between the QTL and the marker increases, (ii) the magnitude of the QTL's effect can be

underestimated because recombination may occur between the marker and the QTL, (iii) association of QTLs with more than one marker cannot be determined through this method (Tanksley, 1993).

SIM was first proposed by Lander and Botstein (1989) to overcome the shortcomings of single marker analysis. This method makes use of linkage maps and evaluates the association between trait values and the genotype of a hypothetical QTL at multiple points between adjacent pairs of linked markers along chromosomes. The use of linked markers accounts for the possibility of recombination between the markers and the QTL, thereby increasing the probability of detecting the QTL and providing a better estimate of the QTL effect (Lander and Botstein, 1989; Liu, 1998). However SIM also has a drawback in that it considers one QTL at a time in the model ignoring the effects of other segregating QTLs. This can reduce the power of detecting a given QTL and it can lead to biased estimation of the effect and position of a QTL when other QTLs are located in the same linkage group (Martinez and Curnow, 1992). CIM provides a better estimate of the position and effect of a QTL by using a model that incorporates the potential presence of other QTLs. Like SIM, CIM evaluates the possibility of a target QTL at multiple points within the interval between two adjacent markers. However, at each point, it also includes the effect of one or more background markers that are commonly referred to as cofactors (Jansen and Stam, 1994; Zeng, 1993). The cofactors are used to minimize the effects of QTLs in the remainder of the genome when attempting to identify a QTL in a particular region.

2.2 Marker assisted breeding

The recent developments in DNA markers and bioinformatics have had a big impact on plant genetics and plant breeding. A large number of QTL mapping studies have been conducted for

diverse crop species resulting in an abundance of marker–trait associations. The most important contribution of these fields is perhaps marker-assisted selection (MAS). MAS refers to the use of molecular markers that are linked to genes or QTLs of interest as an aid to phenotypic screening in a breeding program. Plants that possess particular genes or QTLs may be identified based on their genotype rather than their phenotype. Using molecular markers has advantages in that early identification of cultivars with the desired genes can save breeders a lot of time, effort and money. Furthermore, environmental conditions which influence phenotypes do not affect genotype, which makes the use of molecular markers more reliable. Another benefit from using MAS is that the total number of lines that need to be tested may be reduced. Since many lines can be discarded after MAS at an early generation, this permits a more effective breeding design (Ribaut and Betran, 1999).

However, in recent years it has become widely accepted that QTL validation may be required prior to use in MAS. QTL validation refers to the verification that a QTL is effective in different genetic backgrounds as well as in a range of representative environments (Langridge *et al.,* 2001). There are several factors that make QTL validation necessary. The most important factor is presence of the QTL in the target population. A QTL detected in one population may not be present in another population if the alleles are identical. The QTL may also be influenced by the presence of other non-target QTLs that interact with it (epistasis). The chromosomal region associated with the trait of interest may hold several QTLs and recombination between those QTLs or epistatic interactions may influence the expression of the target QTLs. Errors in QTL mapping can also result in spurious QTLs or overestimation of the QTL effect (Semagn *et al.,* 2006). These factors are particularly important for more complex quantitative traits with many QTLs of small effects (e.g. drought tolerance, yield). Therefore, QTL validation and/or additional testing is often recommended prior to MAS (Foolad *et al.*, 2001).

Chapter 3 Literature Review

3.1 Oat oil content and composition

Oil content is one of the important quality traits in oats. The concentration of oil in the groats of most oat cultivars ranges from 2% to 11% (Youngs, 1986). The upper values are the highest among cereals. However, oat groats are not normally utilized as a source of purified edible oil because this amount is too low to make extraction economically feasible given the value of the commodity and the fact that there is little value to the by-product of the process (White, 2007). It has been estimated that a groat oil content of at least 17% would be required to make the crop a commercial source of edible oil (Frey and Hammond, 1975). Frey and Holland (1999) developed oat lines with oil contents up to 18%, however, the increased oil content was achieved at the cost of other agronomic traits such as yield, grain weight and disease resistance.

The FA composition of oat oil is also of considerable interest to oat breeders, both from the nutritional and technological standpoints. Oat oil contains saturated and unsaturated FAs. For example, linoleic and linolenic acids are essential unsaturated fatty acids in mammalian nutrition while the saturated palmitic acid increases oil stability by protecting against peroxidation. Polyunsaturated FAs may adversely affect the flavour and storage quality of oats. For example linolenic acid causes oil instability (Thro *et al.,* 1983). The three most abundant FA components of oat oil: palmitic acid (16:0), oleic acid (18:1) and linoleic acid (18:2), account for 90-95% of total FAs in oats (Zhou *et al.*, 1999). In addition, oat oil is composed of stearic (18:0), linolenic (18:3), myristic (14:0) and eicosenoic (20:1) acids. Many cultivars also contain trace amounts of palmitoleic (16:1), arachidic acid (20:0) and other minor FA (Frey and Hammond, 1975).

The FA composition of oats varies with the increase and decrease of oil content. Several studies have reported a significant increase in the oleic acid proportion as the overall oil content increased (Frey and Hammond, 1975; Schipper and Frey 1991; Holland *et al.*, 2001). It has been suggested that this positive correlation between oil content and oleic acid is a result of increase in the sink capacity in the form of TAGs, where oleic acid is preferred, relative to structural lipids where it is less dominant (White, 2007). However it has not been established whether increased sink capacity drives higher oil levels or whether this is related to a different causal factor.

3.2 Overview of fatty acid synthesis and TAG accumulation in seeds

In order to understand the possible reasons for the variation in plant lipid composition, an overview of the salient features of lipid biosynthesis in plants is helpful. The TAGs of most seeds contain the common fatty acids: palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2) and linolenic acid (18:3). These FAs are also found in membrane lipids, but whereas proportions of FAs in membrane lipids are conserved, composition of fatty acids in seed storage lipids varies widely. The biosynthesis of FAs for TAG storage is essentially a scaled up version of the membrane lipid synthesis pathway (Voelker and Kinney, 2001).

TAG biosynthesis starts with the *de novo* synthesis of long chain fatty acids from precursors derived from photosynthates in the plastid of cells (Fig 3-1). Two enzyme systems are involved in this process; acetyl-CoA carboxylase (ACCase) and fatty acid synthase (FAS). The first step is conversion of acetyl coenzyme A (acetyl-CoA) to malonyl-CoA in a process catalyzed by ACCase. The second enzyme system of FA synthesis, FAS, transfers the malonyl-CoA to an acyl carrier protein (ACP) and catalyses the extension of the hydrocarbon chain by sequential additions of malonyl moieties. The

final products of FAS are usually 16 carbon (palmitic) and 18 carbon (stearic) fatty acids. The synthesis of a 16 carbon fatty acid requires seven cycles of the FAS complex and the condensation, reduction, and dehydration of 7 malonyl groups. FAS is not a single enzyme but rather a multi-enzyme complex used for acyl-transfer in the sequential reactions involved in 2-carbon addition, and in termination of the overall reaction. The first condensation reaction is catalysed by β-ketoacyl-ACP synthase III (KAS III) which uses acetyl-CoA and malonyl-ACP substrates. The next six condensations are catalysed by KAS I, while the final reaction between palmitoyl-ACP and malonyl-ACP utilises KAS II (Harwood, 1996).

The 16 and 18 carbon fatty acids are then exported to the endoplasmic reticulum (ER) for further elongation or desaturation depending on the genotype of the plant and environmental conditions. However, some species accumulate shorter fatty acids (e.g. 8:0, 10:0, 12:0, 14:0) as a result of the action of thioesterases that cleave the growing fatty acid from the acyl carrier protein (ACP) and thereby prevent further elongation (Broun *et al.*, 1999). Many genes for thioesterases have been cloned. These have been used to create experimental transgenic canola plants that accumulate caprylic acid (8:0), capric acid (10:0), and lauric acid (12:0) rather than 16- and 18-carbon fatty acids (Voelker *et al.*, 1996). The saturated fat, stearic acid (18:0) is desaturated to the monounsaturated FA oleic acid (18:1) in a process catalyzed by stearoyl-ACP Δ9-desaturase (SAD) while it is still in the plastid. Desaturation can take place in the endoplasmic reticulum (FAD 2) or within the plastid (FAD 6). For seed oils, FAD 2 is the main pathway while in leaves the eukaryotic pathway utilising FAD 2 will operate to varying extents (Harwood, 1996). In the ER, 18:1 is desaturated to 18:2 and 18:3 (Figure 3-1) by two specialized microsomal membrane-associated desaturases, FAD2 (ω6) and FAD3 (ω3), respectively (Voelkerand Kinney, 2001).

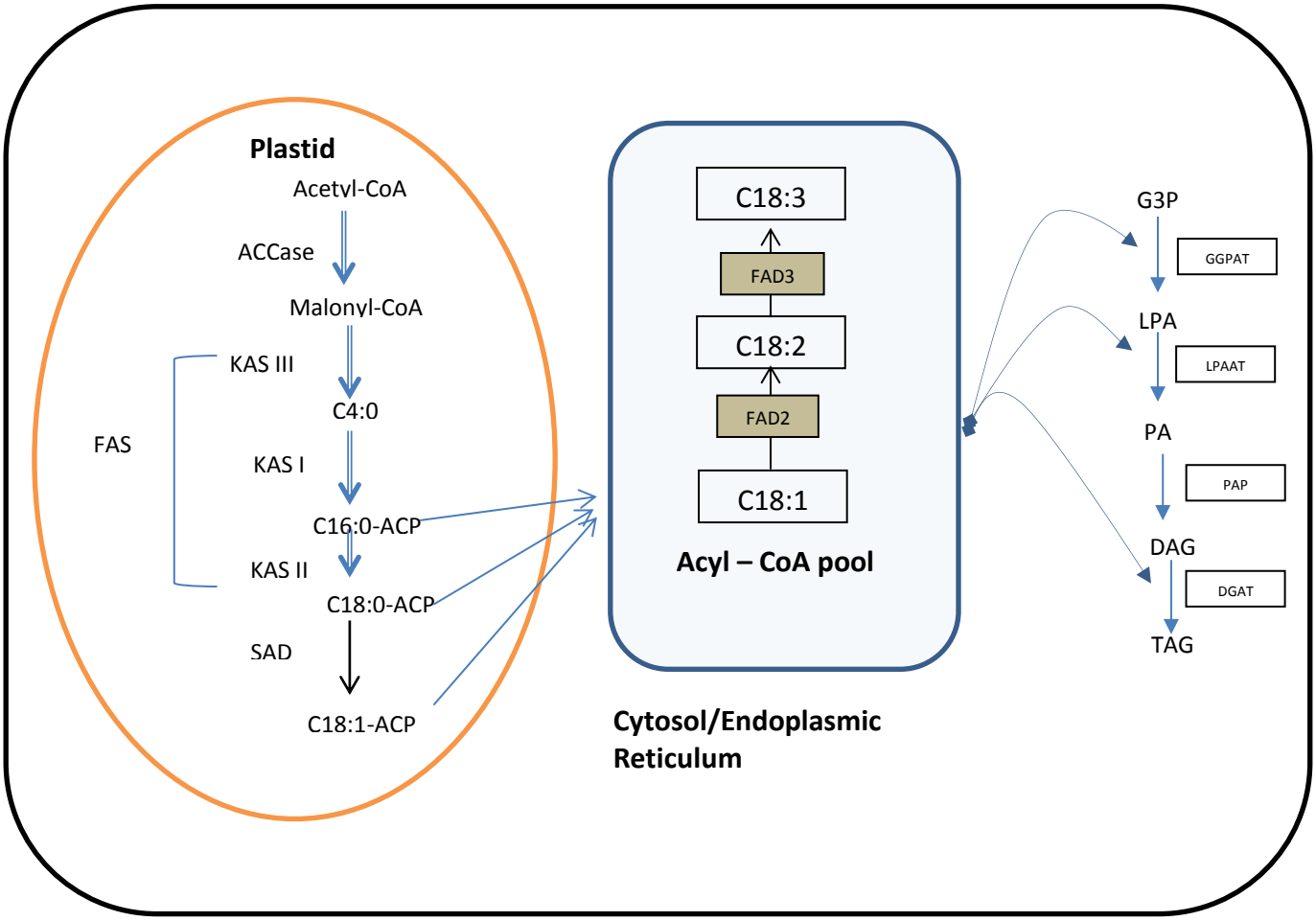


Figure3- 1*A Simplified diagram of fatty acid synthesis and TAG assembly in seeds. (Figure modified from Voelker and Kinney, 2001)*

The free fatty acids exported to the ER become esterified to CoA and serve as substrates for the ER-based eukaryotic lipid formation through the glycerol-3-phosphate (G3P) or Kennedy pathway (Frentzen, 1998). The process is initiated by G3P-acyltransferase (GPAT) which transfers the acyl chain from CoA to the *sn*-1 position of glycerol-3-phosphate (G3P), forming lysophosphatidic acid (LPA). The next step is catalyzed by the enzyme lysophosphatidic acid acyltransferase (LPAAT) which transfers an acyl-chain from the CoA ester to *sn*-2, creating phosphatidic acid (PA). The phosphate group is then removed by phosphatidic acid phosphatase (PAP), forming diacylglycerol (DAG). The final step is catalyzed by diacylglycerol acyltransferase (DGAT), which transfers an acyl group from

acyl-CoA to *sn*-3 of DAG and forms TAG. TAG is thought to accumulate within the ER membranes between the two monolayers to form inclusions bounded by a single leaflet of the ER membrane. Subsequent budding of the inclusions results in maturing lipid bodies that are stabilized by a surface layer of phospholipids and oleosins (Harwood, 1996).

3.3 Regulation of TAG synthesis in plants

Currently there is a good understanding of the biochemistry and the metabolic pathways of FA and lipid synthesis in plants. Many of the genes involved in the process have been identified through expression analysis studies (Beisson *et al.*, 2003). The regulatory elements that direct gene expression in the pathway, however, remain largely unknown. This lack of understanding of the regulatory mechanism has been cited as the main reason for the disappointing results obtained with many transgenic oil crops (Ohlrogge and Jaworski, 1997). Identifying the factors that direct the expression of genes in these pathways is important in order to identify targets for manipulation of the oil content and composition of oil crops. This is even more important in the context of oats where the breeding objectives are complex. Different approaches have been used to study the regulation of FA and oil biosynthesis. Some of the genetic and molecular approaches have been explored here.

3.4 Molecular mechanisms regulating seed oil content

Although the molecular mechanisms of oil biosynthesis may vary from one species to another, the pathway consists of two primary systems: fatty acid synthesis and TAG assembly. The regulatory mechanisms that determine differences in seed oil content and composition have not been completely elucidated. ACCase has been implicated in regulating the extent of TAG production. Studies using developing rapeseed showed that ACCase and malonyl-CoA levels dropped

dramatically when maximum TAG storage levels were achieved even though other FAS enzymes remained at normal levels (Tumham and Northcote, 1983). This suggests a central regulatory role for ACCase. A feedback inhibition mechanism has been suggested (Shintani and Ohlrogge, 1995) in which synthesis of ACCase is reduced by the increased concentration of FAs, particularly 18:1, in the plastid (Andre *et al.,* 2012).

Gene expression analyses on *A. thaliana* seeds have shown a distinct temporal difference in the expression activities of genes involved in the two systems. The transcripts for many core fatty acid synthesizing enzymes such as ACCase subunits, KAS I and FAD2 (oleate desaturase) exhibit a bell-shaped expression pattern. They begin to be down-regulated during the active oil synthesis which indicates global regulation of the entire fatty acid biosynthetic mechanism. In contrast, transcripts for some fatty acid modifying enzymes, such as linoleate desaturase (FAD3), elongase (FAE1), and oleosins are induced later and remain high during the maturation phase (Ruuska *et al.,* 2002). The distinct dissimilarity in the expression patterns suggests that the FA synthesis in the plastid and TAG assembly in the ER are controlled by two different transcriptional programs (Ohlerogge *et al.,* 2004; Baud and Lepiniec, 2009). These findings were supported by additional evidence from similar studies in *Zea mays* (Lee *et al.,* 2002).

In contrast to the above, other studies have suggested that both FA synthesis and TAG assembly may be controlled by a common regulatory mechanism. Studies of transgenic and mutant plants have demonstrated that although the FA composition of seeds may be manipulated to a great extent, the change in composition also results in altered levels of TAG accumulation. This indicates the pleiotropic effect of genetic factors regulating fatty acid synthesis and oil assembly. For example Katavic *et al.* (1995) generated a seed lipid mutant line in *A. thaliana*, *AS11***,** which accumulated low

amounts of TAG in comparison to wild type (WT). The mutation in the *AS11* line is in a gene designated as *TAG1* that encodes DGAT (Zou *et al.*, 1999). Characterization of the mutant revealed a dramatic change to the FA composition of the oil. In comparison to WT, *AS11* had reduced levels of eicosenoic acid (20:1) and oleic acid, while linolenic acid (18:3) was doubled and constituted the major portion of the oil. Moreover, several enzymes of lipid metabolism had altered activity, suggesting that changes to TAG accumulation as a result of a single gene mutation also results in dramatic changes in fatty acid composition of *A. thaliana* seeds

These observations suggest a global regulation of the TAG biosynthesis pathway. Transcription analysis in *A. thaliana* has implicated the transcription factor WRINKLED1 (WRI1) as a potential global regulator of *de novo* fatty acid biosynthesis that specifies the regulatory action of the transcription factor Leafy Cotyledon 2 (LEC2), one of the master regulators of seed maturation which is associated with oil accumulation in maturing seeds (Baud *et al.,* 2007).

3.5 Genetic control of oil synthesis in seeds

Oil content and fatty acid composition are typically quantitative traits under polygenetic control and influenced by environmental conditions. Oil content is a highly heritable trait with several studies reporting heritability ranging between 50 and 90% (Thro and Frey, 1985). High heritability indicates that a trait has greater potential to respond rapidly to selection (Ackerly, 2000).

QTL analysis for oil content has been performed in a variety of plant species including oats (Kianian *et al.*, 1999; Tanhuanpää *et al.*, 2010; Zhu *et al.*, 2004), oilseed rape (Ecke *et al.*, 1995; Burns *et al.*, 2003), sunflower (Mokrani *et al.*, 2002; Leon *et al.*, 2003) and many others. QTL analysis studies in oats have identified several genomic regions affecting variation in oil content. Kianian *et al.* (1999)

found four QTLs in each of two segregating populations. Similar analyses have identified six QTLs for groat oil content in 'Terra' X 'Marion' (TxM) (De Koeyer *et al.*, 2004) and six QTLs in 'Ogle' X ' MAM17-5' (OxM)(Zhu *et al.*, 2004). Additional regions are probably involved, but their individual effects are below the threshold for detection by the QTL approach. Meanwhile, ACCase, which catalyses the first committed step in fatty acid synthesis, has been implicated as a candidate gene-product that affects major oil content differences in oats (Kianian *et al.*, 1999).

The fatty acid composition of seed oil varies considerably both intra and inter-specifically, with fatty acids varying in both chain length and degrees of desaturation. Determining the genetic basis for the variation in fatty acid composition of oilseeds has great potential not only for modification of the oil profiles of commercially important oilseeds designed to meet the requirements of the end users through classical breeding (Keurentjes *et al.*, 2006), but also for unravelling metabolic, regulatory, and developmental pathways (Jansen and Nap, 2001; Barker *et al.*, 2007). Genes involved in the desaturation and elongation processes were identified through mutagenesis experiments on the model species *Arabidopsis thaliana* (Lemieux *et al.,* 1990; Okuley *et al.,* 1994; Beisson *et al.,* 2003). QTL analysis has also been used to identify genomic regions that affect the enzymatic systems involved in the production of specific fatty acids in the lipid biosynthetic pathway (Barker *et al.*, 2007). Several studies have identified loci that are associated with both oil content and the most common fatty acids. Qiu *et al.* (2006) and Ecke *et al.*(1995) reported QTLs associated with oil content in *B. napus* that showed close linkage to the 22:1n9 genes fatty acid elongase1.1 (FAE1.1) and FAE1.2.

3.6 QTL Analysis in Oats

Compared to other crops such as wheat, rice and barley, the mapping of the oat genome has lagged. This is partly due to the large genome size of oat and its polyploid nature. The first large-scale mapping effort in oat was conducted in diploid species. The resulting map was based on F2 families from the cross '*A. atlantica' X* '*A. hirtula'* (O'Donoughue *et al.,* 1992). This map was based primarily on RFLP loci detected using cloned gene sequences from oats and barley. The first molecular linkage map of hexaploid oat was developed by O'Donoughue *et al.* (1995), using a RIL population from a cross between Kanota and Ogle. The map consisted of 561 markers covering a distance of 1482 cM, with 38 linkage groups. Since then, several maps of hexaploid oat have been reported; the most recent published by Tinker *et al.* (2009) which improved the previous KxO maps with the addition of DArT markers. This map contained 1166 markers spanning 1890 cM of the oat genome. The 1166 markers produced 29 linkage groups, with 43 unlinked loci. This map is currently the most complete hexaploid oat molecular linkage map. The KxO maps together with maps developed from other crosses can be used as useful references for oat genomic research.

The first QTL analysis for oil was conducted using KxO population and another mapping population derived from 'Kanota' and 'Marion' (KxM) (Kianian *et al.*, 1999). The authors found four QTLs in each population. Similar analyses using RIL mapping populations have yielded six QTLs for groat oil content in TxM (De Koeyer *et al.*, 2004) and six QTLs in OxM (Zhu *et al.*, 2004). Additional genes are probably involved, but their individual effects are below the threshold of detection by the QTL approach. Meanwhile ACCase was implicated in oil content differences in oats. Hybridization experiments with the candidate gene for ACCase gene detected the presence of three loci in hexaploid oats (Kianian *et al*., 1999). One locus (*Accase A* or *Accase 1*) was mapped in KxO (Kianian *et al.,* 1999, Wight *et al.,* 2003) and two loci (*Accase A, Accase B*) were mapped in KxM (Kianian *et*

al., 1999, Groh *et al.,* 2001). These results indicate consistent relationships between the oil QTLs identified so far. Additional QTL studies in other germplasm are needed to more completely inventory the QTL regions that affect oat oil content and to understand each QTL's environmental sensitivity (Molnar *et al.*, 2011). The aim of the present study therefore, is to identify additional QTLs through detailed QTL analysis of oil in an oat breeding population derived from a cross between high oil ('Dal') and low oil ('Exeter') parents (DxE) or confirm previously discover QTLs and their homologues.

The following chapter is a reproduction of a paper accepted by the journal The Plant Genome by "Biniam T. Hizbai', K. M. Gardner', C. P. Wight, R. K. Dhanda, S. J. Molnar, D. Johnson, J. Frégeau-Reid, W. Yan, B. Rossnagel, J. Holland, N. A. Tinker" (TPG-2012-07-0015). My role was to collect the data for total oil content, protein content, heading date, plant height and lodging; analyse these and additional datasets for oil profile obtained from the University of Saskatchewan for construction of the linkage map and identification of QTLs. I also did the DNA extractions in preparation for DArT analysis. In addition to that I did sample preparation for oil extraction which involved cleaning, de-hulling and grinding. I wrote the first draft of the paper and edited comments received from the co-authors into the final draft, for comments and submission.

The Figures and Tables in Chapter 4 are used with permission of the publisher The Plant Genome Journal in a letter signed by its Managing Editor, Elizabeth Gebhardt.

Biniam Tesfagaber Hizbai

Chapter 4 QTLs affecting oil content, oil composition, and other agronomically important traits in Oat (*Avena sativa* L.)

Biniam T. Hizbai[1,2], K. M. Gardner[1,3], C. P. Wight[1], R. K. Dhanda[4], S. J. Molnar[1], D. Johnson[2], J. Frégeau-Reid[1], W. Yan[1], B. Rossnagel[4], J. Holland[5], N. A. Tinker[1]*

Author affiliations

[1.] *Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada (AAFC), 960 Carling Avenue, Central Experimental Farm, Ottawa, ON, K1A 0C6, Canada*

[2.] *Department of Biology, University of Ottawa, 30 Marie Curie St, Ottawa ON, K1N 6N5, Canada*

[3.] *Nova Scotia Agricultural College, Truro, NS B2N 5E3*

[4.] *Crop Development Centre, 51 Campus Drive, Saskatoon, SK, S7N 5A8, Canada*

[5.] *Department of Crop Science, North Carolina State University, 1238 WIlliams Hall, Raleigh, NC 27695-7620 USA*

Received _____.

** Corresponding author (e-mail: nick.tinker@agr.gc.ca, Telephone: 613-759-1398, Fax: 613-759-1677)*

Abbreviations

AACC, American Association of Cereal Chemists; ACCase, acetyl-coA carboxylase; DAG, diacylglycerol; DGAT, Diacylglyserol acyltransferase; DArT, Diversity Arrays Technology; DH, double haploid; LG, linkage group; NIT, near infrared transmittance ; QTL, quantitative trait loci; RIL, recombinant inbred line; sCIM, simplified composite interval mapping; SIM, simple interval mapping; TAG, triacylglycerol

Abstract

Groat oil content and composition are important determinants of oat quality. We investigated these traits in a population of 146 recombinant inbred lines from a cross between 'Dal' (high oil) and 'Exeter' (low oil). A linkage map consisting of 475 DArT markers spanning 1271.8 cM across 40 linkage groups was constructed. QTL analysis for groat oil content and composition was conducted using grain samples grown at Aberdeen, ID in 1997. QTL analysis for multiple agronomic traits was also conducted using data collected from hill plots and field plots in Ottawa, ON in 2010. Using simple and composite interval mapping methods, QTLs for oil content, palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2) and linolenic acid (18:3) were identified. Two of the loci associated with oil content were associated with all of the fatty acids examined in this study, and most oil-related QTL showed similar patterns of effect on the fatty acid profile. These results suggest the presence of pleiotropic effects on oil-related traits through influences at specific nodes of the oil synthesis pathway. In addition, 12 QTL-associated markers (likely representing nine unique regions) were associated with plant height, heading date, lodging, and protein content.

Key Words: Oat (*Avena sativa*), DArT markers, quantitative trait loci (QTL), oil content, oil composition, ACCase

4.1 Introduction

Cultivated hexaploid oat (*Avena sativa* L.) is an important food and feed crop in many countries. Because oat is consumed directly as food, many aspects of oat quality are important objectives for breeding and molecular research (reviewed by Molnar *et al.,* 2011). The high nutritional value of the oat grain is attributed to its composition of protein, fibre and lipid (Ranhotra and Gelroth, 1995). The oat grain has the highest protein level among cereals, with levels of 12-20% in the groat (dehulled kernel) and 9-15% in the whole grain (Peterson 1992). Furthermore, it has an amino acid profile superior to that of wheat, barley or maize, with higher levels of all essential amino acids (Rines *et al.,* 2006). β-glucan, a high-molecular weight water-soluble fibre, contributes to the health benefits of oat by lowering the serum cholesterol levels of blood and by balancing postprandial glucose and insulin levels (Van Horn *et al.,* 1988; Shinnick *et al.,* 1991). The content and quality of oil in oat groats is of particular importance. For human food, oat varieties with low levels of oil and proportionately higher levels of unsaturated fat are desired. For animal feed, varieties with high levels of oil are desired because they have higher energy content.

Oil content in oat, primarily in the form of triacylglycerols (TAGs), is a highly heritable trait. Thro and Frey (1985) observed a polygenic pattern of inheritance with primarily additive gene action and Schipper and Frey (1991) estimated that heritability of the trait ranged from 63% to 93%. The first quantitative trait locus (QTL) analysis for oil content-associated loci in hexaploid oat was conducted by Kianian *et al.* (1999), using the 'Kanota' X 'Ogle' (KxO) and Kanota X 'Marion' (KxM) mapping populations. They found there were three or four QTLs in each population, depending on the specific QTL analysis method used. Similar analyses using recombinant inbred line (RIL) mapping populations identified six QTLs for groat oil content in 'Terra' X 'Marion' (De Koeyer *et al.,* 2004) and six QTLs in Ogle X 'MAM17-5' (Zhu *et al.,* 2004). Additional genes are probably involved, but their

individual effects would have been below the statistical threshold for detection. Kianian *et al.* (1999) further proposed that acetyl-coA carboxylase (ACCase), which catalyzes the first committed step in fatty acid (FA) synthesis, was a candidate gene for a QTL on linkage group (LG) 11 of the KxO map. Hybridization experiments using an oat ACCase clone detected the presence of three loci in hexaploid oat (Kianian *et al.*, 1999). One locus (*AccaseA* or *Accase1*) was mapped in KxO (Kianian *et al*., 1999, Wight *et al*., 2003) and two loci (*AccaseA*, *AccaseB*) were mapped in KxM (Kianian *et al.,* 1999, Groh *et al.,* 2001). In a more recent study, Tanhuanpää *et al.* (2010) identified eight QTLs associated with oil content. These accounted for 50% of the phenotypic variation in a homozygous doubled haploid (DH) population from the cross 'Aslak' x 'Matilda' (AxM).

In addition to overall oil content, the FA composition of oat oil is also of considerable interest to oat breeders because of the nutritional significance of unsaturated FAs for both humans and animals. Although higher proportions of unsaturated FAs are desired for human consumption, polyunsaturated FAs may adversely affect the flavour and storage quality of oats. Several studies have investigated the FA composition of seeds from different oat varieties grown in various environments (Saastamoinen *et al.,* 1989; Holland *et al.,* 2001) but to our knowledge, genetic studies designed to identify the QTLs affecting these traits have not been conducted. QTL mapping of the genomic regions involved in the variation of oat grain FA content could provide information for the selective breeding of varieties with specific FA profiles.

Many of the reported oil content QTLs in oat show consistency among at least two different studies, but additional studies in other germplasm are required to provide a more complete inventory of these QTL regions and to understand each QTL's sensitivity to environment and genetic background. Furthermore, there is a need to identify and map more genetic markers that are common among

these populations in order to provide more accurate and systematic comparisons of chromosomal location of QTLs. To date, most of the maps in oat have been generated using a variety of markers based on SCAR, SSR, AFLP, RAPD, and RFLP technologies. The lack of common genetic markers among different experiments has made it difficult to compare the positions of QTLs among oat maps. The DArT (Diversity Array Technologies) marker system is a cost effective hybridisation-based marker technology that offers a high level of multiplexing (Mace *et al.,* 2008). A recently developed DArT marker platform for oat and a revised DArT-based reference map in KxO (Tinker *et al.,* 2009) provide new opportunities for rapid map development and comparative mapping in oat.

The objectives of this study were (i) to develop a DArT-based linkage map based on progenies from a cross between oat parents with high *vs.* low groat oil content; (ii) to measure groat oil contents, FA profiles, and additional agronomic traits in the progenies; (iii) to identify major QTLs affecting these traits; (iv) to conduct comparative mapping with locations of other QTLs and candidate genes in oat**.**

4.2 Materials and Methods

An F_5 oat population consisting of 146 RILs was developed by single seed descent from a cross between the two hexaploid oat varieties 'Dal' (high oil) and 'Exeter' (low oil). The Dal x Exeter (DxE) RILs and the parents were grown in Aberdeen, ID, USA in 1997 in four-row irrigated field plots seeded with bulked $F_{5:8}$ grain. Seed from these tests was used to plant two independent non replicated field tests in Ottawa in 2010. The first Ottawa trial was planted in hill plots on May 19, 2010. The second trial was planted in four row plots with a spacing of 25 cm between rows on June 02, 2010. The RILs were randomized separately for the two tests.

Samples for genetic analysis were based on 8-10 seeds that were selected randomly from the $F_{5:8}$ Aberdeen-grown bulks. Seeds were planted in Cyg™ seed germination pouches (Mega Intl., West St. Paul, MN) with one lower corner of each pouch cut so that it could be irrigated from below by capillary action. When the seedlings were approximately 5cm tall, the leaves were harvested and placed in paper envelopes, to which had been added a mixed silica gel desiccant (VWR International, Mississauga, ON, Canada). These were left to dry at room temperature, then ground using a TissueLyser (Qiagen, Mississauga, Canada). DNA was extracted from the parents and the mapping population using Qiagen DNeasy plant mini kits (Qiagen, Maryland, USA) and manufacturer protocols.

DArT marker assays were conducted at Diversity Arrays Technology P/L (Yarralumla, Australia) based on methods described by Tinker *et al.* (2009). The standard hexaploid oat marker array available in 2010 was used, supplemented by an expanded array of markers derived from tetraploid oat (Oliver *et al.,* 2011). Molecular mapping was performed using the maximum likelihood algorithm of the mapping program "Joinmap 4" (Van Ooijen, 2006). Primary LG assignment was performed at

a minimum log of odds ratio (LOD) of 5.0. The grouping threshold was relaxed to LOD 4.0 to identify potential LGs among unassigned loci. Markers that showed high segregation distortion were excluded.

The DxE map was compared to the latest version of the KxO map with linkage groups denoted KOD 1 though KOD 42 to indicate that these are from the DArT based reference map published by Tinker *et al.* (2009). Putative homologies between LGs from the two populations were established based on three types of evidence. Firstly, LGs were declared to be homologous if they shared four or more markers. Secondly, smaller LGs where the proportion of shared markers was high in one or both groups were declared to be homologous. Thirdly, when other supporting evidence was available, either from previous work or the current study, the two groups were declared to be homologous.

Groat oil was measured in 2010 from grain samples grown at Aberdeen in 1997. These samples had been kept at minus 20°C since harvest. Following extraction using an Accelerated Solvent Extractor (ASE-200), total lipid content was determined. Samples from each line and the parents were cleaned thoroughly and mixed in a Seedboro quality mixer (Seedboro Equipment, Chicago, Illinois). Two sets of sub-samples were taken: 100 g for oil extraction and 500 g for protein measurement. The 100 g sub-samples were de-hulled using a CODEMA laboratory huller (CODEMA LLC, Minneapolis, MN, USA) and ground with a cutter mill (Arthur H. Thomas Scientific Apparatus, Philadelphia, PA, USA) fitted with a 20 µm mesh screen. Ground samples (flour) weighing between 400 and 500 mg for each line and the parents were then placed in extraction cells. A ratio of 3:2 of hexane:isopropanol was used as solvent. The conditions for extraction were: pressure of 6.90 x 10^3kPa, temperature of 125°C, heating period of 6 min, extraction (static) time of 20 min, with three static cycles per sample. The total recovered solvent volume was then evaporated at controlled temperature (80°C).

Recovered lipid mass was then weighed on an analytical balance to determine the percentage of oil in the original sample. Oil extraction was performed in two repetitions and the average of the two results was used as the final oil content. Moisture content of each repeated sample was measured at the same time by using the air-oven method (AACC Method 44-15-02). The oil content of samples was calculated on a dry weight basis.

The lipid oil profile was assayed by analysing lipid content extracted from oat flour heated at 90^{o}C for 80 minutes with 2.5 mL of methanolic sulphuric acid (2% v/v), then cooled down and supplemented with a solution of 1 mL hexane spiked with 0.2 mg methyl heptadecanoate per mL of hexane. The mixture was shaken and left overnight for separation of the hexane. The methylated FA samples were analysed by gas chromatography using a Hewlett-Packard 6890 Series GC system (Hewlett-Packard Technologies) equipped with an Agilent 6890 Series injector (Agilent Technologies, Wilmington, DE, USA) and fitted with a DB-Wax capillary column (Agilent 127-7013, 10 m x 100 μm x 0.20 μm).

Protein levels were determined by Near Infrared Transmittance (NIT) on whole groats using an Infratec 1241 system (FOSS, Hillerød, Denmark). The machine is routinely calibrated using a calibration equation developed in house at the Eastern Cereal and Oilseed Research Centre, Ottawa. The calibration is updated and validated every year with wet chemistry.

Phenotypic data for other agronomic traits were collected on all 146 lines of the mapping population planted in the two Ottawa, 2010 field trials. Heading date was defined as the time when 50% of the panicles in a plot were completely emerged. Plant height at maturity was defined as the

mean height of five randomly selected plants per plot or three stems per hill. Visual scoring on a scale of 1-9 was used to determine lodging.

The program “MQTL” (Tinker and Mather 1995) was used to detect and estimate QTL effects. Both simple interval mapping (SIM) and simplified composite interval mapping (sCIM) were used for QTL analysis. The linkage map was scanned at all marker positions and at 5 cM intervals between markers. The QTL with the largest effect on each LG, identified by SIM was used as background marker for sCIM in order to test for additional smaller QTL, and also to refine QTL positions. QTL peaks separated by more than 20 cM were generally regarded as distinct QTL. Significant QTL main effects were used as anchors in testing for pairwise epistatic effects. Each anchor ‘A’ was held constant while the genome was scanned at a walking speed of 5 cM for a second QTL 'B' that interacted with ‘A’. Epistasis was declared when the test statistic for the multiplicative interaction term was greater than the average epistatic test statistic in permuted data. Statistical significance thresholds to achieve 5% genome-wide Type-I error were determined by 10,000 random permutations. Final estimates of QTL positions were based on sCIM or on locations of an epistatic model, and final effects were estimated in a multi-locus regression model incorporating all reported QTL positions for a given trait.

4.3 Results

4.3.1 Molecular mapping

The DArT assay detected 523 polymorphic markers (Table S-1a), 143 of which showed segregation distortion at the 0.05 level of significance, with Dal alleles overrepresented at 86 loci and Exeter at the remaining 57 (Table S-1b). Of these, 48 were excluded because of extreme segregation distortion ($\chi^2 \geq 15$). The remaining 475 markers formed a linkage map with 40 LGs (denoted DE 1 through DE 40) covering a distance of 1271.8 cM (Figure S-2), which is 43% of the 2932 cM genome size estimated by O'Donoughue *et al*. (1995). Seven markers remained unassigned to any LG. Markers were not evenly distributed along the chromosomes. Some LGs had a high concentration of markers, while others were sparsely populated. Thirty nine percent (185) of the markers mapped to only four LGs (DE 4, DE 5, DE 11, and DE 1). Fifteen LGs were made up of 12 to 23 markers each, while the smallest groups (25 of 40 LGs) consisted of two to eight markers each. Linkage group DE4 had the highest number of markers, with 75 markers covering a distance of 86.3 cM. In addition, two LGs (DE 7, DE 17) showed broad regions of skewed segregation ratios favouring the Dal allele (Table S-1c).

The DxE map is considerably shorter than the updated 2,028 cM KO map to which DArT markers had been added (Tinker *et al.,* 2009). This may be a result of inadequate polymorphism in some regions of the parental genomes that correspond to regions of similar ancestry. Dal and Exeter are genetically closer than Kanota and Ogle, with a kinship co-efficient (r) of 0.062% ($p < 0.05$) vs 0.047% for Kanota and Ogle, as calculated using the program "KIN" (Tinker and Mather, 1993). Alternately, it can also be due to the fact that the DArT markers were discovered in KxO and therefore biased to being polymorphic in that population (Tinker *et al.*, 2009).

4.3.2 Map comparison

After comparative mapping, 27 LGs in DxE were found to share markers with 20 LGs in the KxO map (Tinker *et al*., 2009) (Table S-1d). Of these, four LGs from DxE each shared markers with two different KxO LGs, while seven KxO LGs each shared markers with two or three DxE LGs. The number of shared markers ranged from one to 32. Twelve of the associations were based on four or more shared markers. These homologues are: DE 1 with KOD 2, DE 2 and DE 12 with KOD 4_12_13, DE 3 and DE 26 with KOD 16_23, DE 4 with KOD 17, DE 6 with KOD 24_26_34, DE 11 with KOD 19+25+27, DE 13 with KOD 11_41_20_45, DE 14 with KOD 33, DE 17 with KOD 22_44_18, and DE 27 with KOD 1_3_38_X1. Of these homologue sets, the strongest evidence for homology was observed between DE 4 and KOD 17, which shared 32 loci. Although the remaining associations were based on small numbers of shared markers, many of these involve very short LGs with a high proportion of shared markers.

A number of the LGs in the initial KxO map were subsequently joined together using evidence from aneuploidy and from other comparative mapping (Kianian *et al.,* 1997; Wight *et al.,* 2003). Some of the LGs in the DxE map (e.g., DE 19-23-38, DE 6-34, DE 7-13) could potentially be joined, based on their relationships with the KOD LGs. However, doing so could introduce errors in orientation. Since we are aware that comprehensive efforts in developing an oat consensus map are underway, we feel it is preferable to present the current map based solely on evidence from this study.

Several DxE LGs shared markers with more than one KxO LG. While this may indicate groups that could be joined in KxO, it may also indicate homoeologous associations (Figure S-3d). For example, KOD 4_12_13 and KOD 5_30, which both share markers on DE 2, are likely to be homoeologues,

based on other map comparisons (Wight *et al.,* 2003). Similarly, LGs KOD 11_41_20_45 and KOD 37, which both share markers on DE 7, are also likely to be homoeologues (Kianian *et al.,* 1999; Tinker *et al.,* 2009).

4.3.3 Quantitative traits and QTL

Phenotypic data was collected for total oil content, the major FA components, protein, as well as agronomic traits such as plant height, heading date and lodging (Table S-1e). The frequency distributions for all the traits are presented in Figure 4-1. Transgressive segregation in both extremes was observed for stearic acid, oleic acid, linoleic acid, and linolenic acid, while oil content and palmitic acid exhibited transgressive segregation only in one direction. Transgressive segregation was also observed in both extremes for plant height and heading date, but only in one direction for lodging. The presence of QTLs was initially inferred based on SIM. Further scanning using sCIM confirmed the positions of most of the QTLs discovered by SIM and identified some new QTLs. A few of the QTLs discovered using SIM were not confirmed by sCIM, but are being reported as they are statistically highly significant. A summary of QTLs affecting all oil-related traits is presented in Table 4-1, and those QTLs affecting other traits are summarized in Table 4-2.

4.3.4 QTLs affecting oil content

A total of six QTLs associated with oil content were identified in genomic regions belonging to DE 1, DE 4, DE 7, and DE 13 (Table 4-1). LGs DE 4 and DE 13 had two QTLs each and one QTL was observed on each of DE 1 and DE 7. All alleles for higher oil content were contributed by Dal, and their effects ranged from 0.22 (oPt-13269) to 1.87% (oPt-6135) oil content.

The QTL with the largest effect, which accounts for 32% of the phenotypic variation in the trait, was located on DE 13 between oPt-17088_A and oPt-6135. The position of this QTL collocates with a QTL near cdo665b found on KO 11 and KM 11 (Kianian *et al.,* 1999) and a QTL near rz69 on group OM3 (Zhu *et al.*, 2004). Both of these were tightly linked to the *Accase 1* locus. Similarly, a QTL for oil

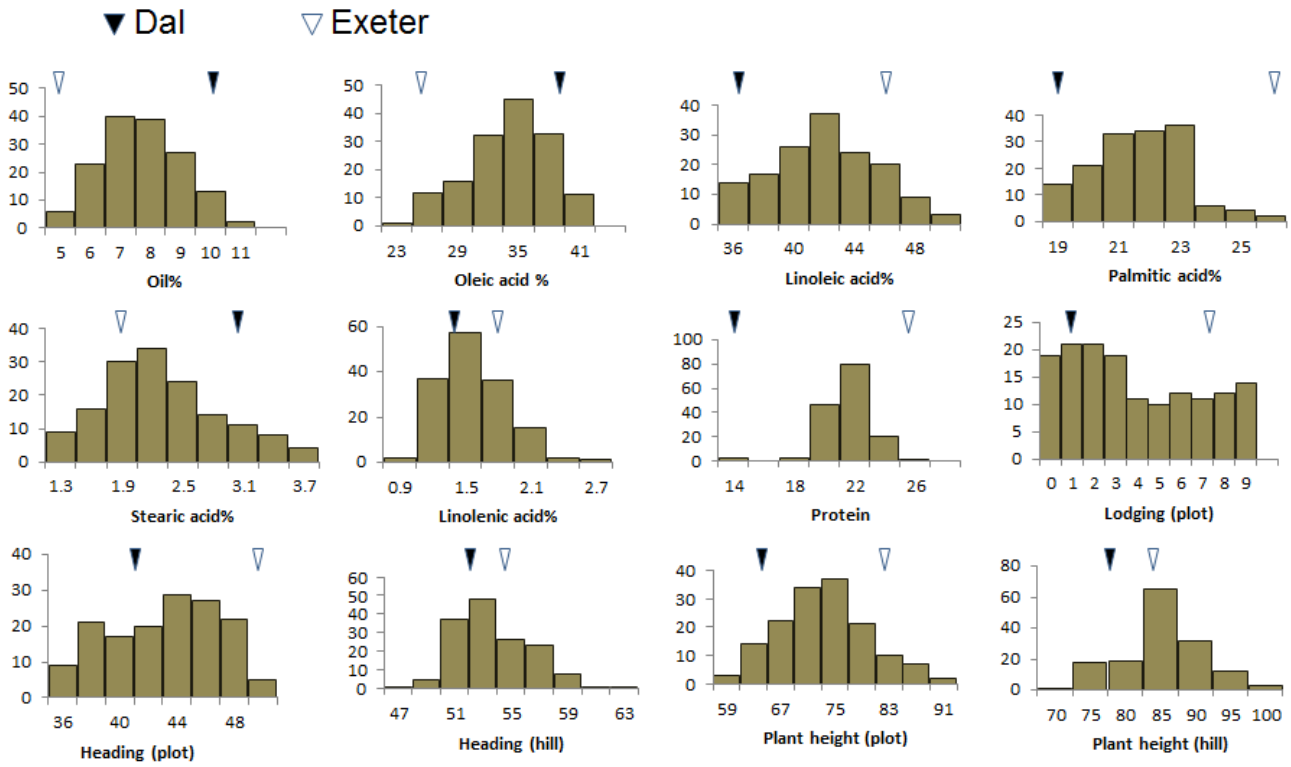


Figure 4- 1 *Frequency distributions for the traits studied in the 146 RILs of the DxE hexaploid oat mapping population. Phenotypes of the parents are marked by arrows.*

content was identified by Tanhuanpää *et al.* (2012) closely linked to *Accase 1* locus on AM 12 which is homologous with KO 11. This QTL has been consistently identified in four different populations. Moreover, in three out of the four studies it was linked to *Accase 1.* This can be considered a validation of this QTL.

The second major QTL, accounting for 19% of the phenotypic variation, was located near oPt-17489 on DE 7, which shared markers with two KxO LGs (KOD 11_41_20_45 and KOD 37) on which QTLs for oil have been mapped.

Homology of DE 7 to either KOD 11_41_20_45 or KOD 37 is ambiguous. DE 7 and KOD 11_41_20_45 share two markers and both LGs contain a major QTL for oil content. In addition, both LGs have regions of segregation distortion (O'Donoughue *et al.,* 1995). The shared markers, the co-location of QTLs for oil content, and the presence of regions of segregation distortion on both DE 7 and KOD 11_41_20_45 would suggest that these two LGs are homologous. However, DE 7 has three markers in common with KOD 37, which also contains a major QTL for oil (Kianian *et al.,* 1999). In addition, DE 7 shares markers with AM 11, which is homologous to KOD 37, and on which the candidate gene *Accase*2 was mapped (Tanhuanpää *et al.,* 2012). More evidence is required to resolve the issue of homology of DE 7 with either KOD 11_41_20_45 or KOD 37; however, the relationships established here suggest that KO 37 and KOD 11_41_20_45 are homoeologous.

Two other QTLs with minor effects were identified on DE 4. The position of one of these loci, oPt-16384, corresponds with a QTL identified on LG 1 of the AxM population by Tanhuanpää *et al.* (2012). Our study is the second to detect QTLs for oil content in this region.

The position of the QTL on DE 1 corresponds to that of the QTL identified on OM 6 (Zhu *et al.,* 2004). The combined effect of the QTLs in this study explains 76% of the phenotypic variation in oil content among the progenies.

4.3.5 QTLs affecting oil composition

A total of four to eight QTLs associated with individual FA components were identified on six LGs (Table 4-1). Four QTLs for palmitic acid, two with major and two with minor effects, were detected

on DE 7, DE 13, and DE 30. These QTLs explain 54% of the total phenotypic variation of the trait among the progenies. All alleles for higher palmitic acid content were contributed by Exeter. Variation in stearic acid content was also associated with four QTLs located on DE 7, DE 13, and DE 30. These QTLs explain 56% of the total phenotypic variation for this trait. Oleic acid content was associated with eight QTLs explaining 76% of the total phenotypic variation. The loci with the highest effects were located on DE 7 and DE 13. All but one of the alleles for increased oleic acid content were contributed by Dal. Linoleic acid content was associated with seven QTLs on DE 4, DE 7, and DE 13, with effects ranging from 0.59 to 2.81%. The alleles for higher levels of this trait were contributed by Exeter, except for one locus, which was derived from Dal. These seven QTLs explain 75% of the phenotypic variation for this trait. The variation in linolenic acid was associated with four QTLs located on DE 3, DE 7, DE 8, and DE 13. The combined effect of these loci explains 42% of the total phenotypic variation for the trait.

Table 4 - 1 Estimated effects of QTLs for total oil and FA content

DxE linkage group	Marker or marker interval	Linkage group distance (cM)	Distance from marker (cM) †	Oil content (%)	Palmitic acid (%)	Stearic acid (%)	Oleic acid (%)	Linoleic acid (%)	Linolenic acid (%)
1	oPt-794644	49.5	0	0.40*					
3	oPt-0373_A	48.8	0						0.16
4	oPt-11790	14.5	0	**0.59***			**0.96***		
4	oPt-17576	18.8	0					**-0.99***	
4	oPt-11705	45.3	0					**-1.10***	
4	oPt-793441	45.8	0				**1.59***		
4	oPt-16384	71	0	**0.30***					
4	oPt-15097	71.5	0				**1.47***	**-2.04***	
7	oPt-15501	12.7	0				**3.44***		
7	oPt-17489	12.7	0	**1.15***	-0.80*	0.35*		**-2.81***	-0.14
8	oPt-17391	8.2	0						-0.15
13	oPt-17088 - oPt-6135	45.5	25/37	**1.87***					
13	oPt-17088 - oPt-6135	55.5	35/37				**3.39***		
13	oPt-6135	57.1	0		**-0.91***	**0.28***		**-2.45***	**-0.27***
13	oPt-17524	85.5	5			**0.21***	**-0.35***	**0.59***	
13	oPt-17524	90.5	10		-0.18				
13	oPt-13269	110.2	5	**0.22***			0.60*		
13	oPt-13269	115.2	10					-1.06*	
30	oPt-11992	0	0		**-1.61***	**0.53***	1.36*		
	Total effect ††			4.53	3.5	1.37	13.16	11.04	0.72

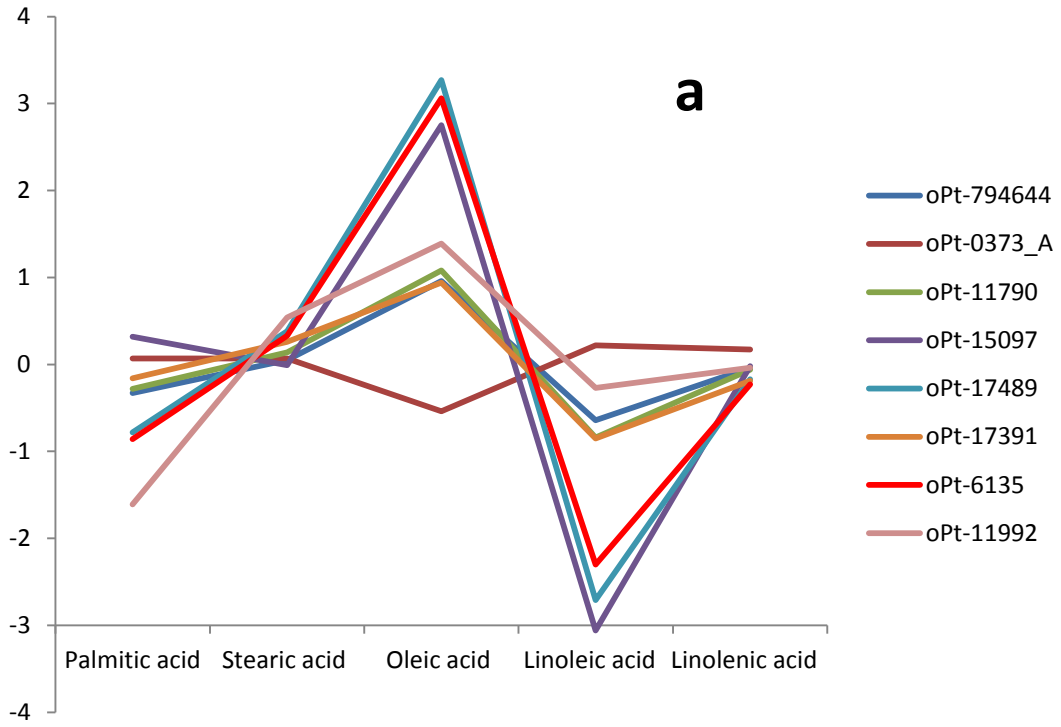
Effect is the measured change in genotypic mean resulting from substituting two 'Exeter' alleles by two 'Dal' alleles.

**All reported effects are significant with an experiment-wide type I error rate of p<0.05 determined by permutation. Those effects with an asterisk are significant with an experiment wide type I error rate of p<0.01.*

† *For marker intervals, this is the distance from the first marker, followed by the total interval distance between the markers.*

†† *Total effect is the sum of absolute values of effects. It predicts the range of progeny values. Shaded cells indicate QTL detected with SIM only while cells with effects in bold fonts were detected with SIM and confirmed with sCIM. The rest were detected with sCIM only. The p value for the sCIM is used as the level of significance when QTL positions are detected with both SIM and sCIM.*

A graphical summary of all QTL regions affecting oil components is presented in Figure 4-2a. This figure shows multi-locus estimates (presented in Table S-1f) for each measured FA regardless of whether the QTL was significant for that specific FA. In order to visualize possible mechanisms for changes in oil components, we re-estimated QTL effects for each oil component measured as a proportion of total seed weight (Figure 4-2b, Table S-1g). When expressed this way, the same QTL regions were significant for at least one of the FA components, though not all QTLs remained significant. When expressed relative to total oil, the general trend for all QTLs was that the Dal alleles increased the proportion of oleic acid and decreased the proportion of linoleic acid. However, when expressed as a proportion of seed weight, the Dal alleles increased or maintained all FA components, with the greatest increase being in oleic acid.



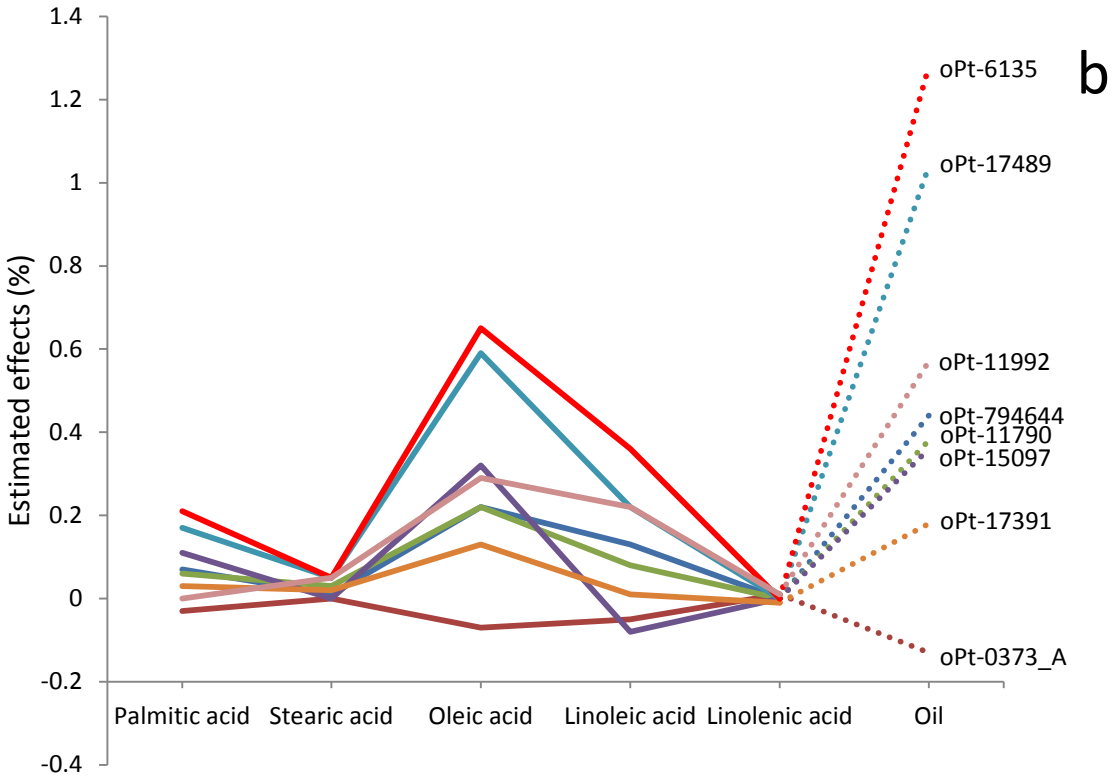


Figure 4- 2 *A graphical summary of multi-locus effects of all QTL regions affecting total oil and FA components when* ***(a)*** *the FA components are expressed as a percentage of oil* ***(b)*** *the total oil and FA contents are expressed as a percentage of seed dry weight. Effects are presented in the order in which FA biosynthesis would occur and different colored lines are used to connect QTL effects at common loci (identified on the right).*

4.3.6 QTL affecting agronomic traits

A total of 12 QTL-associated markers in nine localized regions were associated with plant height, heading date, lodging, and protein content (Table 4-2). Found between oPt-6441 and oPt-15763 on DE 17, only one minor QTL was detected for protein content. This locus accounted for 12% of the total phenotypic variation for the trait. The allele for higher protein content was derived from Exeter.

The variation in plant height for the hill plots was associated with four QTLs. One of these, oPt-10654, was detected only with SIM. When tested with sCIM, the statistic for this locus fell just under threshold. The four loci explain a total of 39% of the total phenotypic variation among the progenies. Both parents contributed alleles for increased height but the alleles from Exeter had the largest effects. Three QTLs were detected with SIM for plant height in the population grown in row plots; however, the test statistics for all three loci fell under the significance threshold upon further scanning with sCIM. Of these, two QTLs were in the same position as those identified in hill plots, while one, oPt-795082, was located on a different LG.

Similar to the other agronomic traits, observations for lodging were carried out on the DxE population planted in hill plots and in row plots. The hill plots were damaged by extremely strong winds. Therefore, QTL analysis was carried out only on measurements taken from the row plots. Three QTLs, found on DE 14, DE 15, and DE 30, explained 68% of the phenotypic variation in lodging (Table 4-2). Both parents contributed alleles that decrease lodging. Two of the alleles were derived from Dal and one was contributed by Exeter. The effects of these loci ranged from 17% to 30%. Overall, reduced lodging was contributed by Dal, which is the shorter parent. All three QTLs were detected with SIM and their positions confirmed with sCIM. The map positions of the loci for lodging in the current study did not correspond to QTLs reported in earlier studies.

Heading date was associated with two QTLs in the hill plots. Both loci (on DE 14 and DE 30) were identified with SIM and confirmed with sCIM. These same loci were also identified in the row plots, as were two more QTLs detected on different LGs (DE 5 and DE 17). One of these, oPt-6446, was detected only with sCIM.

Table 4 - 2 Estimated effects of QTLs for agronomic traits

DxE linkage group	Marker or marker interval	Linkage group distance (cM)	Distance from marker (cM) †	Protein content (%)	Plant height (cm)		Lodging	Heading date (days)	
					hill	plot		hill	plot
5	oPt-6446	9	0						-2.26*
11	oPt-795082	63.8	0			3.59			
12	oPt-18014	16.4	5		**3.24**				
14	oPt-10256	17.8	0		**-5.4***	-2.15			
14	oPt-14507	20	0				**-2.77***		
14	oPt-13141	25.5	0						**-2.81***
14	oPt-16892 - oPt-1427	36.1	5/6					**-2.47***	
14	oPt-14611	39.6	0		**0.56***				
14	oPt-14611 - oPt-13661	44.6	5/8			-2.02			
15	oPt-0594	10	0				**1.8***		
17	oPt-6441 - oPt-15763	11.9	5/5.3	**-0.91***					
17	oPt-6441	6.9	0						**-2.09***
23	oPt-10654	0	0		-2.78				
30	oPt-11992	0	0				**-1.56**	**-1.48**	**-2.44***
	Total effect ††			-0.91	11.98	7.76	6.13	3.95	9.6

Effect is the measured change in genotypic mean resulting from substituting two 'Exeter' alleles by two 'Dal' alleles.

**All reported effects are significant with an experiment-wide type I error rate of p<0.05 determined by permutation. Those effects with an asterisk are significant with an experiment wide type I error rate of p<0.01.*

† *For marker intervals, this is the distance from the first marker, followed by the total interval distance between the markers.*

†† *Total effect is the sum of absolute values of effects. It predicts the range of progeny values. Shaded cells indicate QTL detected with SIM only while cells with effects in bold fonts were detected with SIM and confirmed with sCIM The rest were detected with sCIM only. The p value for the sCIM is used as the level of significance when QTL positions are detected with both SIM and sCIM*

For most QTLs, parental alleles had the expected effects on traits. For example, alleles from the high oil content parent Dal increased oil content, while alleles from the low oil content parent Exeter had the opposite effect. However, a few QTLs exhibited effects in the opposite direction from that expected. For example, oPt-17524 from Dal decreased oleic acid content while the rest of the QTLs from the same parent increased the trait.

4.3.7 QTL Epistasis

Based on 10000 permutations for each trait, a consensus threshold statistic of 15 was established for controlling global type I error rate below 5% per full-genome scan. Scans for epistatic interactions conducted for all traits revealed six potential two-way interactions. Of these, only three were accepted as possible interactions (Table 4-3). The remaining three interactions showed exaggerated genotypic estimates that were likely due to very small sample numbers in one genotypic class. Furthermore, it is recognized that every epistatic scan, of which there were at least several per trait, introduced a new error rate of 5%. Three epistatic loci were associated with plant height. Two of these (oPt-10256 and oPt-14611) interacted with the locus oPt-795758, which is an unlinked marker.

Table 4 - 3 Epistatic QTL interactions for plant height (cm).

Locus A[1]	Locus B[1]	Test Statistic	aabb[2]	AAbb[2]	aaBB[2]	AABB[2]
oPt-795082 (LG 11)	oPt-795525 (LG 38)	17.7	72.29	72.19	66.73	75.3
oPt-10256 (LG 14)	oPt-795758 (LG 41)	16.5	76.74	67.8	71.58	71.29
oPt-14611 (LG 14)	oPt-795758 (LG 41)	17	76.53	66.81	71.48	71.39

[1]Locus A was selected based on the presence of a significant QTL main effect. This locus was held constant while scanning the genome for the presence of Locus B for which there was a significant deviation from the predicted sum of additive effects at each respective locus pair.
[2]Effects are reported as genotypic means for each of four 2-locus homozygous classes. Lower-case signifies the presence of 'Exeter' alleles and uppercase signifies 'Dal' alleles.

4.3.8 Discussion

We have identified eight QTL regions affecting total oil content and/or at least one FA component. The two QTLs with the largest effect on total oil content (near oPt-6135 and oPt-17489) also affected the FA components palmitic, stearic, oleic, and linoleic acid. Moreover, the multi-locus estimates of QTL effects for seven of these eight loci showed the same pattern of effects on all FA, regardless of whether the locus was significant for a specific FA (Figures 4-2a and b), and the QTL analysis for each FA produced highly parallel QTL scans (Figure 4-3). These results suggest that all or most of these QTLs have a similar mode of action, differing only in magnitude, and, therefore, in significance. That is, each QTL region may represent a single segregating locus that affects all oil-related traits pleiotropically through a common physiological or biochemical mechanism. Alternative explanations would imply the presence of clusters of linked QTLs with independent effects on each oil component. This seems improbable, because each QTL region behaves in a parallel manner, so the separate QTLs within each cluster would need to have the same linkage phase. In addition to this, there were no significant epistatic effects among loci affecting oil or its components. Therefore, our working hypothesis is that all or most QTLs affecting oil in this population operate additively through a single node of control.

The pattern of effects observed in this study is consistent with the hypothesis of a global upstream mechanism that affects total oil content as well as FA proportions. Given that TAG assembly is a separate downstream process from FA synthesis, this mechanism seems most likely to be operating at the level of *de novo* FA synthesis in the plastids or even further upstream. Despite our current understanding of the biochemistry of FA and lipid synthesis in plants, the signals and factors that direct the expression of genes in these pathways remain largely unknown, and it is not possible

from these results to determine which genes or regulatory elements are responsible. However, previous authors have speculated on the possible role of ACCase, a rate-limiting enzyme that affects

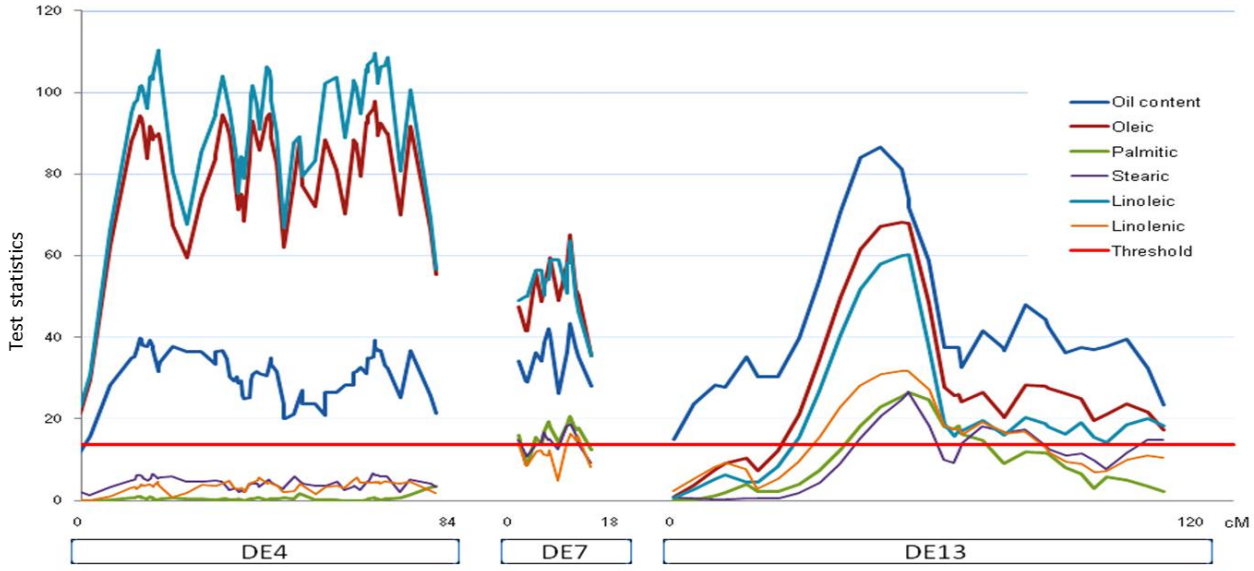


Figure 4- 3 *QTL scans of three linkage groups with parallel effects on oil content and FA components on DxE. The mean threshold value of the six scans was used in this figure.*

the availability of FA precursors, as a primary regulatory node in FA biosynthesis in oat (Kianian *et al.*, 1999). Our current results are consistent with the hypothesis that ACCase is a primary node of control. If, for example, each QTL affects the abundance of ACCase transcripts, then each QTL would have an additive effect on total oil content. The rates of conversion and/or transport among the downstream steps could remain genetically constant, but the relative proportions of FA in the groat would be affected by the abundance of precursors, resulting in our ability to detect QTL for each FA component. Specifically, more precursors could lead to a greater abundance of all FA, but rate-limiting steps could favour the accumulation of oleic acid that we observed. Alternately, White (2007) has suggested that the proportionately higher levels of oleic acid that are observed as total groat oil is increased are a result of increases in the storage of TAGs, where oleic acid is preferred, relative to structural lipids where it is less dominant. This explanation might be taken to suggest

that increased groat oil is a result of increased sink capacity for stored TAGs. Diacylglycerol acyltransferase (DGAT) activity has been proposed as the primary node of control for TAG synthesis. DGAT, which catalyzes the formation of TAGs from diacylglycerol (DAG) and acyl-CoA in the ER, exerts control over oil accumulation (Katavic *et al.,* 1995; Weselake *et al.,* 2008). Increased activity of DGAT may lead to a depletion of acyl-CoAs, which in turn could stimulate de *novo* fatty acid production to satisfy the demand for acyl-CoA. Metabolic control analysis experiments that showed DGAT has a flux control coefficient of 0.74 in olive provided some evidence for this (Ramli *et al.,* 2005). However the same experiment in oil palm showed a coefficient of 0.12 indicating that the control mechanism varies from one species to another. Determining which of these mechanisms (source vs sink) is responsible for genetic variation in groat oil content will be an important step in elucidating the molecular determinants of oil metabolism in oat.

To our knowledge, this is the first study to investigate QTL effects at the level of FA composition in oat. Because of this, the QTLs associated with FAs in this study cannot be compared directly to previous work. However, comparative mapping has shown that at least three of the eight oil-related QTL regions in DxE are coincident with QTLs affecting total oil content in other oat populations, so we speculate that their mode of action and pleiotropic effects on FA could be similar.

Although oil content in oat is a highly heritable trait that can be increased and decreased by selective breeding (Frey and Holland, 1999), the pleiotropic effects on FA composition observed in the current study suggest that there is limited opportunity for manipulation of the FA composition of the oil independently of the total oil concentration. The exception may be linolenic acid, for which QTLs were detected on DE 3 and DE 8 with no effect on other measured traits. However, the

individual effects of these two QTLs were very small, accounting for differences of only 7% and 9% of linolenic acid proportion in the oil or 9% and 13% of the linolenic acid proportion in the total seed by dry weight.

4.3.8.1 Segregation distortion

The proportion of markers exhibiting segregation distortion in the current study (27% at $p<0.05$) was higher than expected. O'Donoughue *et al.* (1995) reported that 8% of the markers tested deviated from the 1:1 ratio expected in a recombinant inbred line population, Zhu and Kaeppler (2003) reported 9%, and Portyanko *et al.* (2001) reported 13%. A combination of the type of population used and the dominant inheritance pattern of DArT markers may have contributed to the unusually high proportion of markers with distorted segregation. Both DH as well as RIL may exhibit extreme segregation distortion possibly due to the presence of recessive alleles that are lethal or strongly selected (Xian-Lang, 2006). It is also possible that dominant alleles segregating at two or more different genetic loci can masquerade as one over-represented allele at a single locus because of the nature of the marker assay. This phenomenon is known to affect some DArT markers in oat (Tinker *et al.,* 2009). However, the later phenomenon would only account for extreme (e.g. 3:1) genetic distortions, and such markers would generally be excluded from the map unless they clustered to produce a separate pseudo-LG comprised entirely of duplicated loci. It is also possible for segregation distortion to result from the loss of unbalanced gametes caused by translocation heterozygotes. For example, there is a major known translocation in oat between chromosomes 7C and 17 (Jellen and Beard, 2000) that has affected mapping in other crosses, but both Dal and Exeter are expected to contain the same version of this translocation and there is no evidence for segregation distortion on groups DE 27, 28, or 35 which correspond to KxO groups 1_3_38 known to be affected by this translocation (Figure S-3n and Table S-1c). It is possible that

other minor translocations or other chromosomal rearrangements have influenced segregation distortion. Markers with distortions greater than 2:1 were not included in the map and there were no entire LGs composed of markers distorted to this degree. Rather, the graduated regions of mild to moderate segregation distortion observed in the DxE map appear to represent regions where one or the other parental allele has been selected or has drifted to a majority state.

Specific regions of segregation distortion have been reported in other oat populations. O'Donoughue *et al.* (1995) reported a distorted region on KO 11. A distorted region was also reported on OM3, a LG in the Ogle/MAM17-5 population which is homologous to KO 11 (Zhu *et al.,* 2004). The distorted region on DE 7 is also potentially homologous to KO 11 (Figure S-3g). Genetic control of segregation distortion has been reported in other cereals such as rice (Devaux *et al.,* 1995) and barley (Li *et al.,* 2010). The presence of distortion in homologous regions across three different oat populations provides strong evidence for the presence of a heritable factor affecting segregation distortion. Since all three populations were developed through single seed descent, it is unlikely that there is a common mechanism affecting fitness after seed maturation. Possible mechanisms could include gametic or zygotic selection (Liu *et al.,* 2010). However, the loci on both KO 11 and OM3 were skewed toward the male parents, whereas the distortion on DE 7 favoured alleles from the female parent, Dal. Therefore, it seems unlikely that there is a common mechanism of gamete selection, but there could be a heritable mechanism affecting the survival or fitness of the zygotes. It is interesting that the same distorted region in all three oat populations contains major QTLs affecting oil content, suggesting that the mechanism responsible for this aberrant segregation pattern may directly or indirectly affect oil content.

4.3.8.2 QTLs affecting additional traits

The position of the single QTL identified for protein content in DxE corresponds with the QTL reported by De Koeyer *et al.* (2004) on LG TM 15 of the 'Terra' X 'Marion' map. Previous studies have identified additional QTLs affecting protein content in oats, but the genetic control of this trait does not appear to be consistent, ranging from numerous small effects (Zhu *et al.,* 2004) to relatively large effects (Tanhuanpää *et al.,* 2012). These inconsistencies may be attributed to the large magnitude of environmental influence on protein content. Crop production environments, and nitrogen supply in particular, affect protein percentages of oat (Welch and Leggett, 1997).

It is interesting that two QTL regions affecting heading date and height in both trials also coincided with two QTLs affecting lodging in both the hill plots and the row plots. The QTL oPt-11992 on DE 30, which had an effect on heading date in both trials, was also associated with lodging. Similarly, oPt-10256, which affected height in both trials, was located close to the QTL that had the largest effect on lodging (oPt-14057). These associations were consistent with the phenotypic correlations observed between the traits (Table S-1h), and the direction of the effects appears to have a developmental basis: Dal alleles that reduced plant height and days to heading also reduced lodging. These correlated responses could be due to linkage of the underlying QTLs or to pleiotropy. Some QTLs affecting plant height and heading date, however, were not consistent between the row plot and hill plot trials. This discrepancy might be attributed to differences in the growing conditions between the two sets of plots. The hill plots were planted two weeks earlier than the row plots, and they experienced a heavier weed infestation.

In summary, we have identified several chromosome regions and DArT markers associated with oil content, oil quality, and several agronomic traits, some of which have been linked to loci identified in other populations. Moreover, our study has demonstrated the relationship between genetic

factors controlling oil content and the concentration of individual FA components of oil in oats. Although QTL analysis lacks the power and resolution to elucidate metabolic pathways, by estimating the effects of all the alleles for each biosynthetic product along the metabolic pathway we have provided an overview of the genetic mechanism of oil biosynthesis in oats. This information will be useful not only for manipulating the composition and concentration of oil, but also in ongoing efforts to understand the molecular basis for the control of FA composition and oil accumulation in seeds. The results of this study will be used for further fine mapping and identification of suitable markers for molecular breeding.

Chapter 5 - General Discussion

There are numerous reports on the health benefits of oats for human diet. However, the high oil content of oat poses a challenge for the widespread adoption of oats as a main staple in the world. Enhancing the good nutritional profile, while at the same time reducing the caloric content, is a complex challenge faced by oat breeders. Most of the traits of value are controlled by multiple genetic factors which makes the task even more complex. Identification of the genetic factors controlling the myriad of traits in crop plants provides breeders with an important tool in their efforts to meet the constantly changing quality demands for crops.

The current study has identified important genomic regions contributing to the variation in oil content, oil composition, and protein content, as well as agronomic traits such as heading date, plant height, and lodging. The results show that there is little overlap in the position of the genetic factors associated with agronomic traits and oil-related traits. The QTLs for oil content and FAs were detected on linkage groups DE 1, DE 3, DE 4, DE 7, DE 8, DE 13, and DE 30, whereas QTLs for agronomic traits were detected on linkage groups DE 5, DE 11, DE 12, DE 14, DE 15, DE 17, DE 23, and DE 30. This apparently reduced linkage between the QTLs underlying oil-related traits and agronomic traits is useful from a breeding perspective, as it offers an opportunity to manipulate the traits independently with no need of decoupling the QTLs for linked traits.

In contrast to this, however, previous studies have indicated that increasing oil content in oats does impact agronomically important traits. Schipper and Frey (1992) reported unfavourable correlated responses in key agronomic traits after six cycles of recurrent selection in a genetically broad-based oat population. Lower biomass, grain yield, and test weight were observed in later cycles of

selection. After nine cycles of recurrent selection in the same population, Holland *et al*. (2001) reported reduced yield and disease resistance and a higher incidence of lodging in the high oil lines. There are several possible explanations for this. First, the number of QTLs involved in increasing oat oil content to such extreme levels would be greater than the QTLs identified in the present study, such that additional QTLs not present in 'Dal' x 'Exeter' are involved. Second, if there is linkage between some genetic factors affecting oil and agronomic performance, the QTLs with effects on both oil and agronomic traits may not have been detected in our study. In addition, by selecting for oil content and ignoring other QTLs for agronomic traits, deterioration in agronomic performance may have resulted from genetic drift. Further studies are needed to investigate the relationship among these traits and the developmental basis for these relationships.

One QTL near oPt-11992, with a pleiotropic effect on the oil-related traits palmitic acid, stearic acid, and oleic acid, did affect the agronomic traits lodging and heading date. Holland *et al.* (2001) observed that an increase in groat oil content was accompanied by correlated increases in stearic acid and oleic acid content, while palmitic acid, linoleic acid, and linolenic acid content decreased with further selection cycles. The results from the present study are consistent with this observation as all the loci associated with high oil content also affected the FAs in a parallel manner. The same authors also reported significant correlations between heading date and oil content, with early lines being associated with higher oil. Our results also show that the Dal version of oPt-11992 has a similar effect on the traits: it reduced heading date, increased stearic acid and oleic acid content, and reduced palmitic acid content (Table 4-1). The pattern of effects, as shown in Figures 4-2a and b, indicates that the locus also affects total oil content. These results suggest that there is genetic linkage between heading date and oil content and composition at least at one locus. There may be a developmental basis for this correlation between heading date and oil content. Oil synthesis in oat

starts early in seed formation at about 15 days from flowering (Brown *et al.,* 1970). Banās *et al*. (2007) compared a high oil cultivar with a low oil one and found that, while lipid accumulation started early in both cultivars, the accumulation continued until maturation in the high oil variety, while it started decreasing before maturation in the low oil variety.

The pattern of effects of QTL regions on all oil-related traits observed in this study suggests that the genetic control for oil content in oat may be a global upstream mechanism that affects total oil content as well as FA proportions (discussed in detail in Chapter 4). The extent of oil accumulation in oats could be controlled by the availability of source substrates and cofactors (FAs and enzymes), in which case the control may be exerted either at the level of *de novo* FA synthesis in the plastid or even further upstream. Alternatively, sink capacity may determine the rate of FA synthesis, in which case the control node may be at the level of DGAT, the enzyme which catalyzes the final step in TAG assembly in the ER. It could also be a combination of both mechanisms. The consistent collocation of the major QTLs for oil content with *Accase* clones in many populations suggests that the level of oil accumulation in oats is likely determined by the availability of FAs, but it remains possible that there are multiple genetic factors affecting oil-related traits that are collocated in the same regions of the genome.

The FA composition of oats is nutritionally important. Oats contain very high levels of oleic acid and linoleic acid. Oleic acid tends to increase high density lipoprotein (HDL) cholesterol in the human blood stream. Linoleic acid has no known effect on HDL cholesterol levels. Both oleic and linoleic acid however, lower LDL (low density lipoprotein) cholesterol (Erickson *et al.,* 1964). This combination has important health benefits and might be used to enhance the health food status of oats. The QTLs identified in this study could be useful starting points for breeding programs to

improve the FA composition of commercial varieties of oats. However, the pleiotropic effects of the QTLs we have observed indicate that manipulating the content of any individual FA will likely affect the oil content and the levels of the other FAs as well. To improve the composition of oat oil independent of oil content we will have to look for other sources of variation to utilize in oat cultivar improvement programs.

5.1 Suggestions for future research

The clone for the candidate gene *Accase* had been used previously as an RFLP marker and mapped near major QTLs for oil content in four different populations (KxO, KxM, OxM, and AxM) (Kianian *et al.*, 1999; Zhu *et al.*, 2004; Tanhuanpää *et al.*, 2010). Comparative mapping shows that the QTLs with the largest effect on oil content in the DxE population co-locate with the major QTLs for oil in these other oat populations (Fig S-3). Based on this consistent relationship between loci for *Accase* genes and major QTLs for oil, we can presume that *Accase* genes are also closely associated with two of the major QTLs for oil in the DxE population (those on DE 7 and DE 13). Screening the same clone (or a PCR-based marker derived from it) across the whole DxE population would confirm this.

The Iinkage map created in this study is not complete. While it provides a good framework for the identification of QTLs in the DxE population, it is possible that there are unmarked regions of the genome that contain QTLs. This map could be enhanced by the addition of other marker types such as SNP markers. Efforts are underway to produce high-throughput markers such as SNPs for oat (Tinker, personal communication). Using these markers to score the oat mapping populations used in previous studies would also make comparing the maps easier, and eliminate the need for a bridge map such as KxO. Having more common markers and not having to use a bridge map could also help

resolve conflicts between maps that can arise as a result of scoring errors or as a result of biological phenomena such as chromosomal rearrangements.

The present work could also be extended to identify the genes underlying QTLs associated with oil content and individual FAs in oat groats. Key genes involved in lipid metabolism have been cloned and characterised in the model species *A. thaliana* and orthologous genes have been identified in maize (Li *et al*., 2010). Comparison of the map positions of QTLs for oil-related traits and candidate genes involved in lipid metabolism in maize (Li *et al*., 2010; Yang *et al*., 2009) and rice (Ying *et al.,* 2012) showed that many of the QTLs co-located with the candidate genes. In oat, sequences from wheat, barley, and *Lolium* spp. have been used to identify PCR primer pairs that could be used to isolate and clone sequences corresponding to the vernalization genes *VRN1*, *VRN2* and *VRN3*. Using this approach, an oat *VRN1* fragment was mapped to a region that was homoeologous to a major QTL affecting response to vernalization (Nava *et al*., 2012). Available sequences for candidate genes involved in lipid metabolism in other cereals such as maize could be used in the same manner to develop markers which could be mapped and compared to QTLs affecting oil content and composition in oats. Such a study could be important not only for oat breeding purposes, but also for gaining useful insight into the mechanisms controlling oil biosynthesis in oats.

References

Ackerly, D. D., S. A. Dudley, S. E. Sultan, J. Schmitt, J. S. Coleman, C. R. Linder, D. R. Sandquist, M. A. Geber, A. S. Evans, T. E. Dawson and Martin J. Lechowicz. 2000. The evolution of plant ecophysiological traits: recent advances and future directions. BioScience 50 (11):979-995.

Agarwal, M., N. Shrivastava, and H. Padh. 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reports 27 (4):617-631.

Akbari, M., P. Wenzl, V. Caig, J. Carling, L. Xia, S. Yang, G. Uszynski, V. Mohler, A. Lehmensiek, H. Kuchel. M. J. Hayden, N. Howes, P. Sharp, P. Vaughan, B. Rathmell, E. Huttner and A. Kilian. 2006. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theoretical and Applied Genetics 113 (8):1409-1420.

Andre, C., R. P. Haslam, and J. Shanklin. 2012. Feedback regulation of plastidic acetyl-CoA carboxylase by 18: 1-acyl carrier protein in Brassica napus. Proceedings of the National Academy of Sciences 109 (25):10107-10112.

Banas, A., H. Debski, W. Banas, W. K. Heneen, A. Dahlqvist, M. Bafor, P. O. Gummeson, S. Marttila, Ã. Ekman, and A. S. Carlsson and S. Stymne. 2007. Lipids in grain tissues of oat (*Avena sativa*): differences in content, time of deposition, and fatty acid composition. Journal of Experimental Botany 58 (10):2463-2470.

Barker, G. C., T. R. Larson, I. A. Graham, J. R. Lynn, and G. J. King. 2007. Novel insights into seed fatty acid synthesis and modification pathways from genetic diversity and quantitative trait loci analysis of the Brassica C genome. Plant physiology 144 (4):1827-1842.

Baud, S., and L. Lepiniec. 2009. Regulation of de novo fatty acid synthesis in maturing oilseeds of Arabidopsis. Plant Physiology and Biochemistry 47 (6):448-455.

Baud, S., M. S. Mendoza, A. To, E. Harscoët, L. Lepiniec, and B. Dubreucq. 2007. WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. The Plant Journal 50 (5):825-838.

Beavis, W.D. 1998. QTL analyses: Power, precision and accuracy. *In*: A.H. PATERSON (ed). Molecular dissection of complex traits. CRC Press, Boca Raton, FL. pp. 145-161.

Beisson, F., A. J. K. Koo, S. Ruuska, J. Schwender, M. Pollard, J. J. Thelen, T. Paddock, J. J. Salas, L. Savage, A. Milcamps V. B. Mhaske, Y. Cho and J. B. Ohlrogge . 2003. Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. Plant physiology 132 (2): 681-697.

Brookes, A. J. 1999. The essence of SNPs. Gene 234 (2):177-186.

Broun, P., S. Gettner, and C. Somerville. 1999. Genetic engineering of plant lipids. Annual review of nutrition 19 (1):197-216.

Brown, C. M., E. J. Weber, and C. M. Wilson. 1970. Lipid and amino acid composition of developing oats (Avenasativa L. cultivar'Brave'). Crop Science 10 (5):488-491.

Brugmans, B., R. G. M. van der Hulst, R. G. F. Visser, P. Lindhout, and H. J. van Eck. 2003. A new and versatile method for the successful conversion of AFLP™ markers into simple single locus markers. Nucleic acids research 31 (10):e55.

Burns, M. J., S. R. Barnes, J. G. Bowman, M. H. E. Clarke, C. P. Werner, and M. J. Kearsey. 2003. QTL analysis of an intervarietal set of substitution lines in Brassica napus: (i) Seed oil content and fatty acid composition. Heredity 90 (1):39-48.

Burr, B., F. A. Burr, K. H. Thompson, M. C. Albertson, and C. W. Stuber. 1988. Gene mapping with recombinant inbreds in maize. Genetics 118 (3):519-526.

Chalmers, K. J., A. W. Campbell, J. Kretschmer, A. Karakousis, P. H. Henschke, S. Pierens, N. Harker, M. Pallotta, G. B. Cornish, M. R. Shariflou, L. R. Rampling, A. Mclauchlan, G. Daggard, P. J. Sharp, T. A. Holton, M. W. Sutherland, R. Appels and P. Langridge. 2001. Construction of three linkage maps in bread wheat, Triticumaestivum. Australian Journal of Agricultural Research 52 (11-12):1089-1119.

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

Collins, F. S., M. S. Guyer, and A. Chakravarti. 1997. Variations on a theme: cataloging human DNA sequence variation. Science 278 (5343):1580-1581.

De Koeyer, D. L., N. A. Tinker, C. P. Wight, J. Deyl, V. D. Burrows, L. S. O'Donoughue, A. Lybaert, S. J. Molnar, K. C. Armstrong, G. Fedak, D. M. Wesenberg, B. G. Rossnagel, and A. R. McElroy. 2004. A molecular linkage map with associated QTLs from a hulless x covered spring oat population. Theoretical and Applied Genetics 108 (7):1285-1298.

Devaux, P., A. Kilian, and A. Kleinhofs. 1995. Comparative mapping of the barley genome with male and female recombination-derived, doubled haploid populations. Molecular and General Genetics 249 (6):600-608.

Ecke, W., M. Uzunova, and K. Weissleder. 1995. Mapping the genome of rapeseed (Brassica napus L.). II. Localization of genes controlling erucic acid synthesis and seed oil content. Theoretical and Applied Genetics 91 (6):972-977.

Erickson, B. A., R. H. Coots, F. H. Mattson, and A. M. Kligman. 1964. The effect of partial hydrogenation of dietary fats, of the ratio of polyunsaturated to saturated fatty acids, and of dietary cholesterol upon plasma lipids in man. Journal of Clinical Investigation 43 (11):2017 –2025.

FAOSTAT (2012) Food and Agriculture Organization of the United Nations. Available at: http://faostat.fao.org/. Accessed July 14, 2012

Farooq, S., and F. Azam. 2002. Molecular markers in plant breeding: Concepts and characterization. Pakistan Journal of Biological Sciences 5 (10):1135-1140.

Food and Drug Administration, U. S. 1997. FDA final rule for federal labeling: health claims: oats and coronary heart disease. Federal Registry 62 (15):3584-3601.

Foolad, M. R., L. P. Zhang, and G. Y. Lin. 2001. Identification and validation of QTLs for salt tolerance during vegetative growth in tomato by selective genotyping. Genome 44 (3):444-454.

Frentzen, M. 1998. Acyltransferases from basic science to modified seed oils. European Journal of Lipid Science and Technology 100 (4/5):161-166.

Frey, K. J., and E. G. Hammond. 1975. Genetics, characteristics, and utilization of oil in caryopses of oat species. Journal of the American Oil Chemists' Society 52 (9):358-362.

Frey, K. J., and J. B. Holland. 1999. Nine cycles of recurrent selection for increased groat-oil content in oat. Crop Science 39 (6):1636-1641.

Gibson, L., and G. Benson. 2002. Origin, History, and Uses of Oat (*Avena sativa*) and Wheat (*Triticum aestivum*). Iowa State University, Department of Agronomy. http://www.agron.iastate.edu/courses/agron212/readings/Oat_wheat_history.htm . Accessed August 20, 2012.

Groh, S., A. Zacharias, S. F. Kianian, G. A. Penner, J. Chong, H. W. Rines, and R. L. Phillips. 2001. Comparative AFLP mapping in two hexaploid oat populations. Theoretical and Applied Genetics 102 (6):876-884.

Gupta, P. K., R. K. Varshney, P. C. Sharma, and B. Ramesh. 1999. Molecular markers and their applications in wheat breeding. Plant Breeding 118 (5):369-390.

Haldane, J. B. S. 1919. The combination of linkage values and the calculation of distances between the loci of linked factors. Journal of Genetics 8 (29):299-309.

Harwood, J. L. 1996. Recent advances in the biosynthesis of plant fatty acids. Biochimica et biophysica acta 1301(1-2):7-56.

Helentjaris, T., M. Slocum, S. Wright, A. Schaefer, and J. Nienhuis. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theoretical and Applied Genetics 72 (6):761-769.

Holland, J. B., K. J. Frey, and E. G. Hammond. 2001. Correlated responses of fatty acid composition, grain quality, and agronomic traits to nine cycles of recurrent selection for increased oil content in oat. Euphytica 122 (1):69-79.

Jaccoud, D., K. Peng, D. Feinstein, and A. Kilian. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic acids research 29 (4):e25.

Jansen, R. C., and J. P. Nap. 2001. Genetical genomics: the added value from segregation. Trends in Genetics 17 (7):388-391.

Jansen, R. C., and P. Stam. 1994. High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136 (4):1447-1455.

Jellen, E. N., and J. Beard. 2000. Geographical Distribution of a Chromosome 7C and 17 Intergenomic Translocation in Cultivated Oat. Crop Science. 40 (1):256-263.

Katavic, V., D. W. Reed, D. C. Taylor, E. M. Giblin, D. L. Barton, J. Zou, S. L. MacKenzie, P. S. Covello, and L. Kunst. 1995. Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in Arabidopsis thaliana affecting diacylglycerolacyltransferase activity. Plant physiology 108 (1):399-409.

Kearsey, M. J., and H. S. Pooni. 1996. The genetical analysis of quantitative traits. London: Chapman and Hall.

Keurentjes, J. J. B., J. Fu, C. H. R. De Vos, A. Lommen, R. D. Hall, R. J. Bino, L. H. W. van der Plas, R. C. Jansen, D. Vreugdenhil, and M. Koornneef. 2006. The genetics of plant metabolism. Nature genetics 38 (7):842-849.

Kianian, S. F., M. A. Egli, R. L. Phillips, H. W. Rines, D. A. Somers, B. G. Gengenbach, F. H. Webster, S. M. Livingston, S. Groh, and L. S. O'Donoughue, M. E. Sorrells, D. M. Wesenberg, D. D. Stuthman and R. G. Fulcher. 1999. Association of a major groat oil content QTL and an acetyl-CoA carboxylase gene in oat. Theoretical and Applied Genetics 98 (6):884-894.

Kianian, S. F., B. C. Wu, S. L. Fox, H. W. Rines, and R. L. Phillips. 1997. Aneuploid marker assignment in hexaploid oat with the C genome as a reference for determining remnant homoeology. Genome 40 (3):386-396.

Kosambi, D. D. 1943. The estimation of map distances from recombination values. Annals of Human Genetics 12 (1):172-175.

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

Langridge, P., and K. Chalmers. 2005. Molecular Marker Systems in Plant Breeding and Crop Improvement, edited by H. Lörz and G. Wenzel: Springer Berlin Heidelberg, 3-22.

Langridge, P., E. S. Lagudah, T. A. Holton, R. Appels, P. J. Sharp, and K. J. Chalmers. 2001. Trends in genetic and genome analyses in wheat: a review. Australian Journal of Agricultural Research 52 (11-12):1043-1077.

Lee, J. M., M. E. Williams, S. V. Tingey, and A. J. Rafalski. 2002. DNA array profiling of gene expression changes during maize embryo development. Functional and integrative genomics 2 (1):13-27.

Lemieux, B., M. Miquel, C. Somerville, and J. Browse. 1990. Mutants of Arabidopsis with alterations in seed lipid fatty acid composition. Theoretical and Applied Genetics 80 (2):234-240.

Leon, A. J., F. H. Andrade, and M. Lee. 2003. Genetic analysis of seed-oil concentration across generations and environments in sunflower. Crop Science 43:135-140.

Li, H., A. Kilian, M. Zhou, P. Wenzl, E. Huttner, N. Mendham, L. McIntyre, and R. E. Vaillancourt. 2010. Construction of a high-density composite map and comparative mapping of segregation distortion regions in barley. Molecular Genetics and Genomics 284 (5):319-331.

Li, L., H. Li, J. Y. Li, S. T. Xu, X. H. Yang, J. S. Li, and J. B. Yan. 2010. A genome-wide survey of maize lipid-related genes: candidate genes mining, digital gene expression profiling and co-location with QTL for maize kernel oil. Science China Life Sciences 53 (6):690-700.

Lionneton, E., S. Ravera, L. Sanchez, G. Aubert, R. Delourme, and S. Ochatt. 2002. Development of an AFLP-based linkage map and localization of QTLs for seed fatty acid content in condiment mustard (Brassica juncea). Genome 45 (6):1203-1215.

Liu, B. H. 1998. Statistical genomics: linkage, mapping, and QTL analysis: CRC Press, Boca Raton.

Liu, X., L. Guo, J. You, X. Liu, Y. He, J. Yuan, G. Liu, and Z. Feng. 2010. Progress of Segregation Distortion in Genetic Mapping of Plants. Research Journal of Agronomy 4 (4):78-83.

Mace, E., L. Xia, D. Jordan, K. Halloran, D. Parh, E. Huttner, P. Wenzl, and A. Kilian. 2008. DArT markers: diversity analyses and mapping in Sorghum bicolor. BMC genomics 9:26.

Manly, K. F., and J. M. Olson. 1999. Overview of QTL mapping software and introduction to Map Manager QT. Mammalian Genome 10 (4):327-334.

Martinez, O., and R. N. Curnow. 1992. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theoretical and Applied Genetics 85 (4):480-488.

Melchinger, A. E., H. F. Utz, and C. C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149 (1):383-403.

Meydani, M. 2009. Potential health benefits of avenanthramides of oats. Nutrition reviews 67 (12):731-735.

Mokrani, L., L. Gentzbittel, F. Azanza, L. Fitamant, G. Al-Chaarani, and A. Sarrafi. 2002. Mapping and analysis of quantitative trait loci for grain oil content and agronomic traits using AFLP and SSR in sunflower (*Helianthus annuus* L.). Theoretical and Applied Genetics 106 (1):149-156.

Molnar, S. J., N. A. Tinker, H. F. Kaeppler and H. W. Rines. 2011. Molecular genetics of quality in oats. *In*: F. H. Webster and P.J. Wood (eds.), Oats: Chemistry and Technology. 2nd Edition, American Association of Cereal Chemists (AACC), St Paul, MN. pp. 51-75.

Nava, I., C. Wight, M. Pacheco, L. Federizzi, and N. Tinker. 2012. Tagging and mapping candidate loci for vernalization and flower initiation in hexaploid oat. Molecular Breeding. DOI 10.1007/s11032-012-9715-x.

Nguyen, H.T., and X. Wu. 2005. Molecular Marker Systems for Genetic Mapping, *In* K. Meksem and G. Kahl, eds. The Handbook of Plant Genome Mapping. Wiley-VCH Verlag GmbH and Co., Weinheim, Germany. pp. 23-52

O'Donoughue, L. S., M. E. Sorrells, S. D. Tanksley, E. Autrique, A. V. Deynze, S. F. Kianian, R. L. Phillips, B. Wu, H. W. Rines, P. J. Rayapati, M. Lee, G.A Penner, G. Fedak, S.J. Molnar, D. Hoffman, and C.A. Salas. 1995. A molecular linkage map of cultivated oat. Genome 38 (2):368- 380.

O'Donoughue, L. S., Z. Wang, M. Röder, B. Kneen, M. Leggett, M. E. Sorrells, and S. D. Tanksley. 1992. An RFLP-based linkage map of oats based on a cross between two diploid taxa (*A. atlantica× A. hirtula*). Genome 35 (5):765-771.

Ohlrogge, J., V. B. Mhaske, F. Beisson, and S. Ruuska. 2004. Genomics approaches to lipid biosynthesis. Proceedings of the 4th International Crop Science Congress. 26 September – 1 October Brisbane, Australia

Ohlrogge, J. B., and J. G. Jaworski. 1997. Regulation of fatty acid synthesis. Annual Review of Plant Biology 48 (1):109-136.

Okuley, J., J. Lightner, K. Feldmann, N. Yadav, and E. Lark. 1994. Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. The Plant Cell Online 6 (1):147-158.

Oliver, R. E., E. N. Jellen, G. Ladizinsky, A. B. Korol, A. Kilian, J. L. Beard, Z. Dumlupinar, N. H. Wisniewski-Morehead, E. Svedin, M. Coon, R.R. Redman, P.J. Maughan, D.E. Obert, and E.W. Jackson. 2011. New Diversity Arrays Technology (DArT) markers for tetraploid oat (*Avena magna* Murphy et Terrell) provide the first complete oat linkage map and markers linked to domestication genes from hexaploid *A. sativa* L. Theoretical and Applied Genetics 123:1159-1171.

Peterson, D. M. 1992. Composition and nutritional characteristics of oat grain and products. *In:* Oat science and technology, edited by H. G. Marshall and M. E. Sorrells.: American Society of Agronomy, Madison, WI. pp.265-292.

Polanco, C., A. I. González, R. de la Puente, S. Somalo, and M. L. Ruiz. 2005. Cloning of AFLP markers detected by fluorescence capillary electrophoresis. Plant Molecular Biology Reporter 23 (3):271-277.

Portyanko, V. A., D. L. Hoffman, M. Lee, and J. B. Holland. 2001. A linkage map of hexaploid oat based on grass anchor DNA clones and its relationship to other oat maps. Genome 44 (2):249-265.

Qiu, D., C. Morgan, J. Shi, Y. Long, J. Liu, R. Li, X. Zhuang, Y. Wang, X. Tan, and E. Dietrich. 2006. A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. Theoretical and Applied Genetics 114 (1):67-80.

Ramli, U. S., J. J. Salas, P. A. Quant, and J. L. Harwood. 2005. Metabolic control analysis reveals an important role for diacylglycerolacyltransferase in olive but not in oil palm lipid accumulation. Federation of European Biochemical Societies Journal 272 (22):5764-5770.

Ranhotra G.S., J.A. Gelroth. 1995. Food uses of oats. *In*: R.W. Welch (ed.) The oat crop: Production and utilization. 1st Edition, Chapman and Hall, London, UK. pp. 409-430

Ribaut, J. M., and J. Betrán. 1999. Single large-scale marker-assisted selection (SLS-MAS). Molecular Breeding 5 (6):531-541.

Rines, H.W., S.J. Molnar, N.A. Tinker, and R.L. Phillips. 2006. Oat. *In*: C. Kole (ed.) Genome mapping and molecular breeding in plants: Cereals and millets Vol. 1. Springer, Inc., NY, USA. pp. 211-242

Ruuska, S. A., T. Girke, C. Benning, and J. B. Ohlrogge. 2002. Contrapuntal networks of gene expression during Arabidopsis seed filling. The Plant Cell Online 14 (6):1191-1206.

Saastamoinen, M., J. Kumpulainen, and S. Nummela. 1989. Genetic and environmental variation in oil content and fatty acid composition of oats. Cereal Chemistry 66 (4):296-300. Available at http://www.aaccnet.org/publications/cc/backissues/1989/Documents/66_296.pdf

Schipper, H., and K. J. Frey. 1991. Selection for groat-oil content in oat grown in field and greenhouse. Crop Science 31 (3):661-665.

Schipper, H., and K. J. Frey. 1992. Responses in agronomic traits associated with three recurrent selection regimes for groat-oil content in oat. Plant Breeding 108 (4):302-313.

Schön, C. C., H. F. Utz, S. Groh, B. Truberg, S. Openshaw, and A. E. Melchinger. 2004. Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167 (1):485-498.

Semagn, K., Å. Bjørnstad, and M. N. Ndjiondjop. 2006. Principles, requirements and prospects of genetic mapping in plants. African Journal of Biotechnology 5 (25) 2569-2587.

Semagn, K., Å. Bjørnstad, and Y. Xu. 2010. The genetic dissection of quantitative traits in crops. Electronic Journal of Biotechnology 13 (5):16-17.

Shinnick, F. L., R. Mathews, and S. Ink. 1991. Serum cholesterol reduction by oats and other fiber sources. Cereal Foods World 36:815-821.

Shintani, D. K., and J. B. Ohlrogge. 1995. Feedback inhibition of fatty acid synthesis in tobacco suspension cells. The Plant Journal 7 (4):577-587.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of molecular biology 98 (3):503-517.

Squirrell, J., P. M. Hollingsworth, M. Woodhead, J. Russell, A. J. Lowe, M. Gibby, and W. Powell. 2003. How much effort is required to isolate nuclear microsatellites from plants? Molecular Ecology 12 (6):1339-1348.

Stich, B., and A. E. Melchinger. 2010. An introduction to association mapping in plants. CAB Reviews. Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources, 5 (039) 1-9

Strychar, R. 2011. World oat production, trade, and usage. *In:* F. H. Webster and P.J. Wood (eds.). Oats: Chemistry and Technology. 2nd Edition, American Association of Cereal Chemists (AACC), St Paul, MN. pp.1-10

Tanhuanpää, P., O. Manninen, A. Beattie, P. Eckstein, G. Scoles, B. Rossnagel, E. Kiviharju, and D. Somers. 2012. An updated doubled haploid oat linkage map and QTL mapping of agronomic and grain quality traits from Canadian field trials. Genome 55 (4):289-301.

Tanhuanpää, P., O. Manninen, and E. Kiviharju. 2010. QTLs for important breeding characteristics in the doubled haploid oat progeny. Genome 53 (6):482-493.

Tanksley, S. D. 1993. Mapping polygenes. Annual review of genetics 27 (1):205-233.

Tautz, D., and M. Renz. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic acids research 12 (10):4127-4138.

Thro, A. M., and K. J. Frey. 1985. Inheritance of groat-oil content and high-oil selection in oats (*Avena sativa* L.). Euphytica 34 (2):251-263.

Thro, A. M., K. J. Frey, and E. G. Hammond. 1983. Inheritance of fatty acid composition in oat (*Avena sativa* L.) caryopses. Plant Foods for Human Nutrition 32 (1):29-36.

Tinker, N. A., A. Kilian, C. P. Wight, K. Heller-Uszynska, P. Wenzl, H. W. Rines, A. Bjornstad, C. J. Howarth, J. L. Jannink, J. M. Anderson, B. G. Rossnagel, D. D. Stuthman, M. E. Sorrells, E. W. Jackson, S. Tuvesson, F. L. Kolb, O. Olsson, L. C. Federizzi, M. L. Carson, H. H. Ohm, S. J. Molnar, G. J. Scoles, P. E. Eckstein, J. M. Bonman, A. Ceplitis, and T. Langdon. 2009. New DArT markers for oat provide enhanced map coverage and global germplasm characterization. BMC genomics 10 (1):39.

Tinker, N. A., and D. E. Mather. 1993. KIN: Software for computing kinship coefficients. Journal of Heredity 84 (3):238-238.

Tinker, N. A. and D.E. Mather. 1995. MQTL: software for simplified composite interval mapping of QTL in multiple environments. Journal of Agricultural genomics 1 (2):1-5.

Turnham, E., and D. H. Northcote. 1983. Changes in the activity of acetyl-CoA carboxylase during rape-seed formation. Biochemical Journal 212 (1):223.

Van Horn, L., L. A. Emidy, K. Liu, Y. Liao, C. Ballew, J. King, and J. Stamler. 1988. Serum lipid response to a fat-modified, oatmeal-enhanced diet. Preventive medicine 17 (3):377-386.

Van, K., E. Y. Hwang, M. Y. Kim, H. J. Park, S. H. Lee, and P. B. Cregan. 2005. Discovery of SNPs in soybean genotypes frequently used as the parents of mapping populations in the United States and Korea. Journal of Heredity 96 (5):529-535.

Van Ooijen, J. W. 2006. JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Netherlands.

Voelker, T., and A. J. Kinney. 2001. Variations in the biosynthesis of seed-storage lipids. *Annual* Review of Plant Biology 52 (1):335-361.

Voelker, T. A., T. R. Hayes, A. M. Cranmer, J. C. Turner, and H. M. Davies. 1996. Genetic engineering of a quantitative trait: metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. The Plant Journal 9 (2):229-241.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van De Lee, M. Hornes, A. Friters, J. Pot, J. Paleman, and M. Kuiper. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic acids research 23 (21):4407-4414.

Welch, R. W., and J. M. Leggett. 1997. Nitrogen Content, Oil Content and Oil Composition of Oat Cultivars (A. sativa) and Wild *Avena Species* in Relation to Nitrogen Fertility, Yield and Partitioning of Assimilates. Journal of cereal science 26 (1):105-120.

Welsh, J., and M. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic acids research 18 (24):7213-7218.

Weselake, R. J., S. Shah, M. Tang, P. A. Quant, C. L. Snyder, T. L. Furukawa-Stoffer, W. Zhu, D. C. Taylor, J. Zou, A. Kumar, L. Hall, A. Laroche, G. Rakow, P. Raney, M. M. Moloney, and J. L. Harwood. 2008. Metabolic control analysis is helpful for informed genetic manipulation of oilseed rape (*Brassica napus*) to increase seed oil content. Journal of Experimental Botany 59 (13):3543-3549.

White, P. 2007. Fatty Acids in Oilseeds (Vegetable Oils). *In* Fatty Acids in Foods and their Health Implications, 3rd Edition. CRC Press, 227-262.

Wight, C.P., N. A. Tinker, S. F. Kianian, M. E. Sorrells, L. S. O'Donoughue, D. L. Hoffman, S. Groh, G. J. Scoles, C. D. Li, F. H. Webster, R. L. Phillips, H. W. Rines, S.M. Livingston, K. C. Armstrong, G. Fedak, and S. J. Molnar. 2003. A molecular marker map in 'Kanota'x'Ogle' hexaploid oat

(*Avena* spp.) enhanced by additional markers and a robust framework. Genome 46(1):28-47.

Williams, J. G. K., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic acids research 18 (22):6531-6535.

Xian-Liang, S., S. Xue-Zhen, and Z. Tian-Zhen. 2006. Segregation distortion and its effect on genetic mapping in plants. *Chinese Journal of Agricultural Biotechnology* 3 (3):163-170.

Xu, Y., and J. H. Crouch. 2008. Marker-assisted selection in plant breeding: from publications to practice. Crop Science 48 (2):391-407.

Yang, X., Y. Guo, J. Yan, J. Zhang, T. Song, T. Rocheford, and J. S. Li. 2010. Major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. Theoretical and Applied Genetics 120 (3):665-678.

Ying, J. Z., J. X. Shan, J. P. Gao, M. Z. Zhu, M. Shi, and H. X. Lin. 2011. Identification of Quantitative Trait Loci for Lipid Metabolism in Rice Seeds. Molecular Plant 5(4):865-875.

Youngs, V. L. 1986. Oat lipids and lipid-related enzymes. *In:* F. H. Webster (ed.). Oats: Chemistry and Technology. 2nd Edition, American Association of Cereal Chemists (AACC), St Paul, MN. pp. 205-226.

Youngs, V. L. 1978. Oat lipids. Cereal Chemistry 55 (5):591-597

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.

Zhou, M., K. Robards, M. Glennie-Holmes, and S. Helliwell. 1999. Oat lipids. Journal of the American Oil Chemists' Society 76 (2):159-169.

Zhu, S., and H. F. Kaeppler. 2003. A genetic linkage map for hexaploid, cultivated oat (Avenasativa L.) based on an intraspecific cross' Ogle/MAM17-5'. Theoretical and Applied Genetics 107 (1):26-35.

Zhu, S., B. G. Rossnagel, and H. F. Kaeppler. 2004. Genetic analysis of quantitative trait loci for groat protein and oil content in oat. Crop Science 44 (1):254-260.

Zou, J., Y. Wei, C. Jako, A. Kumar, G. Selvaraj, and D. C. Taylor. 1999. The Arabidopsis thaliana TAG1 mutant has a mutation in a diacylglycerolacyltransferase gene. The Plant Journal 19 (6):645-653.

Appendix I – Linkage map of Dal x Exeter

Figure S-2 Linkage map of the hexaploid cultivated oat (*A. sativa* L.) based on 152 recombinant inbred lines from the cross Dal X Exeter population cross. Map distances are given in centi-Morgans

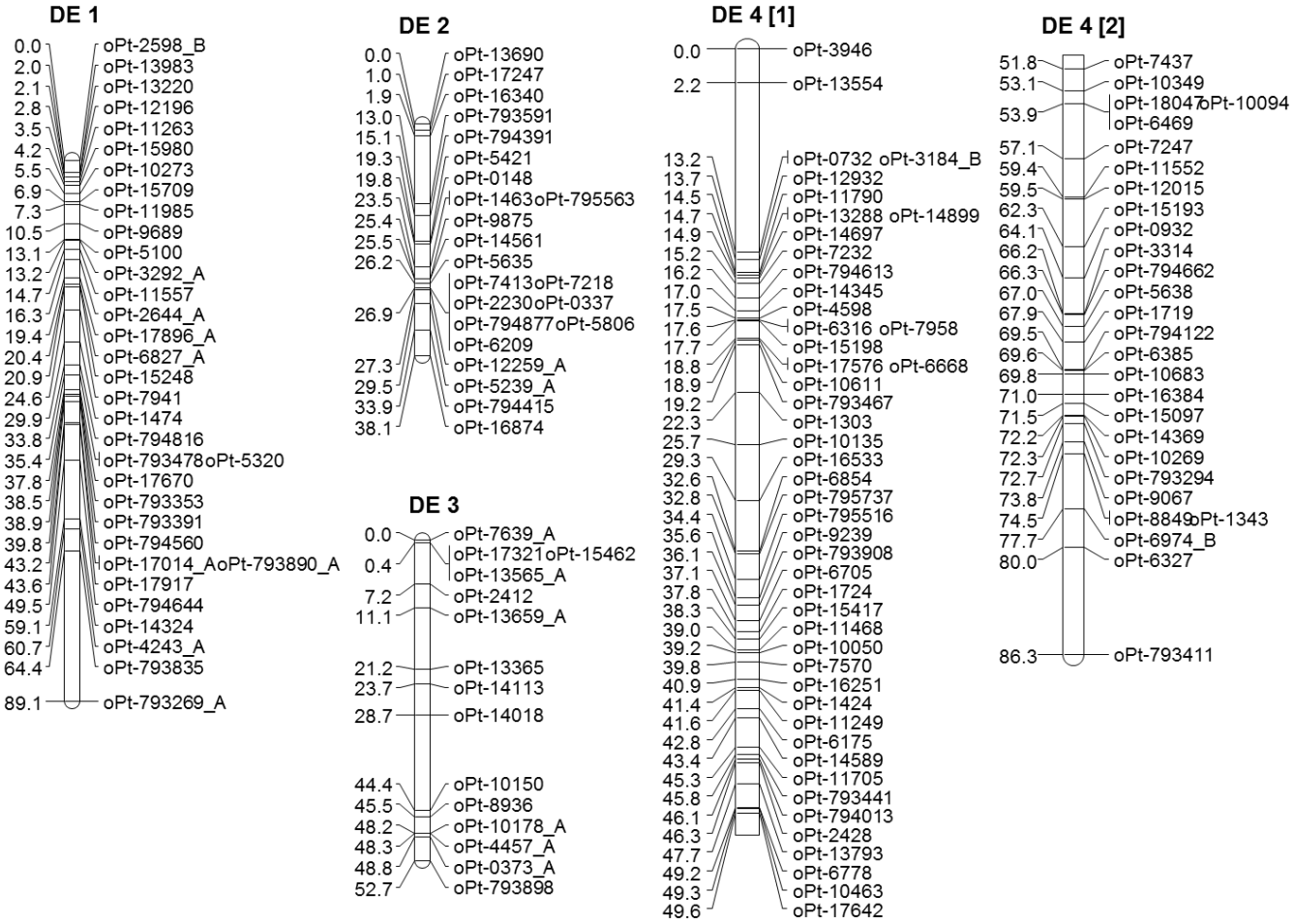


Figure S-2 Linkage map of the Dal X Exeter population continued…

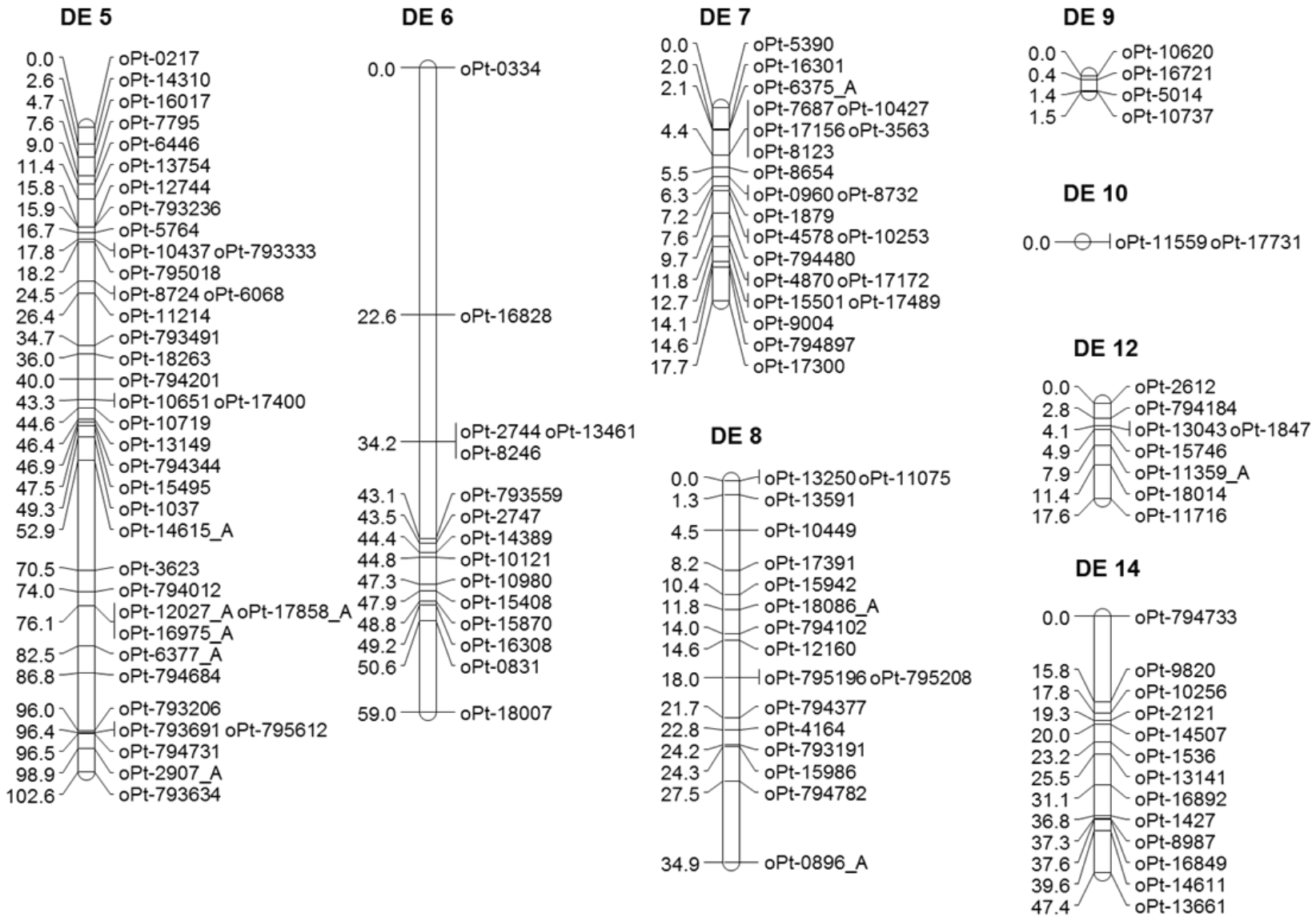


Figure S-2 Linkage map of the Dal X Exeter population continued…

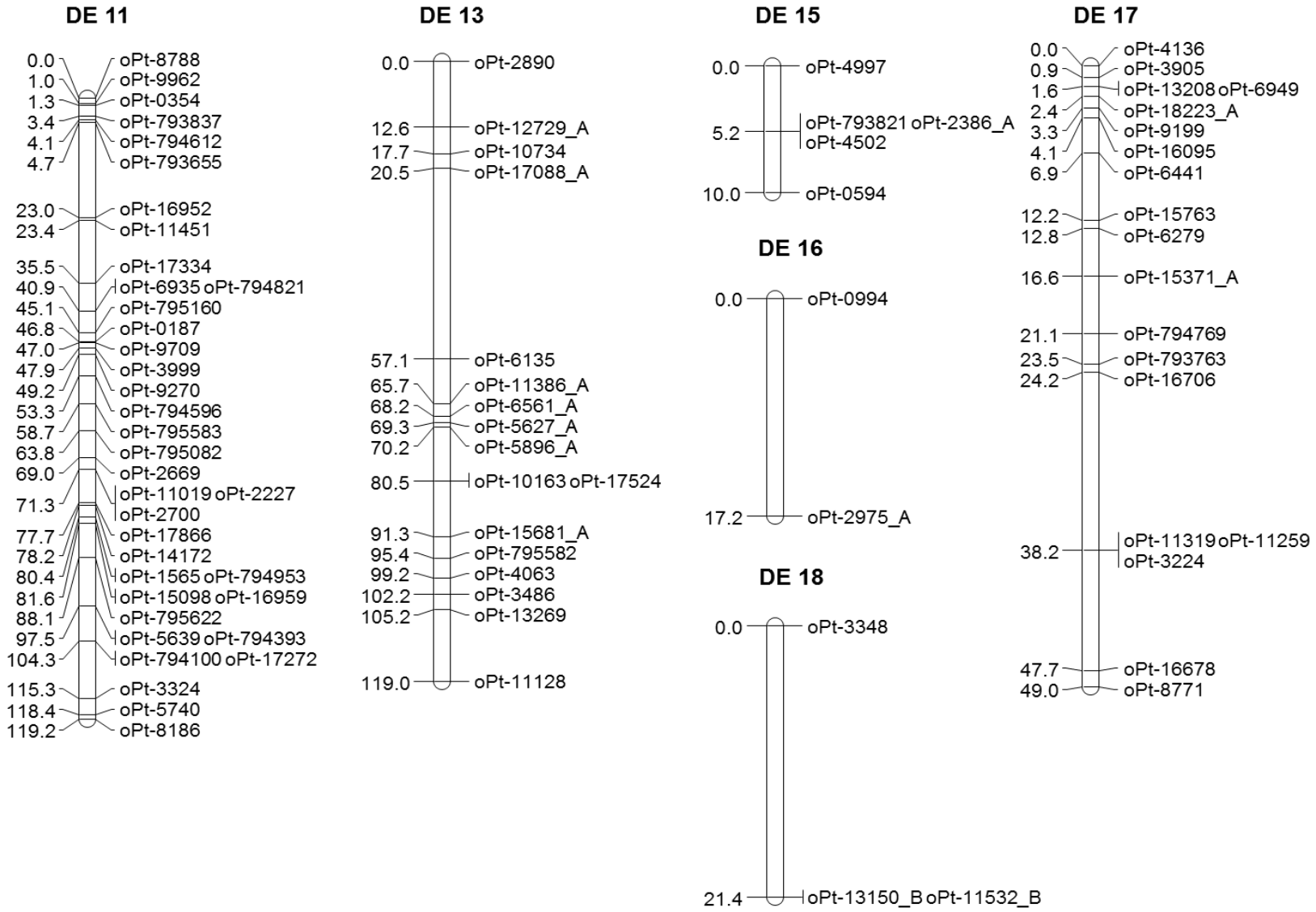
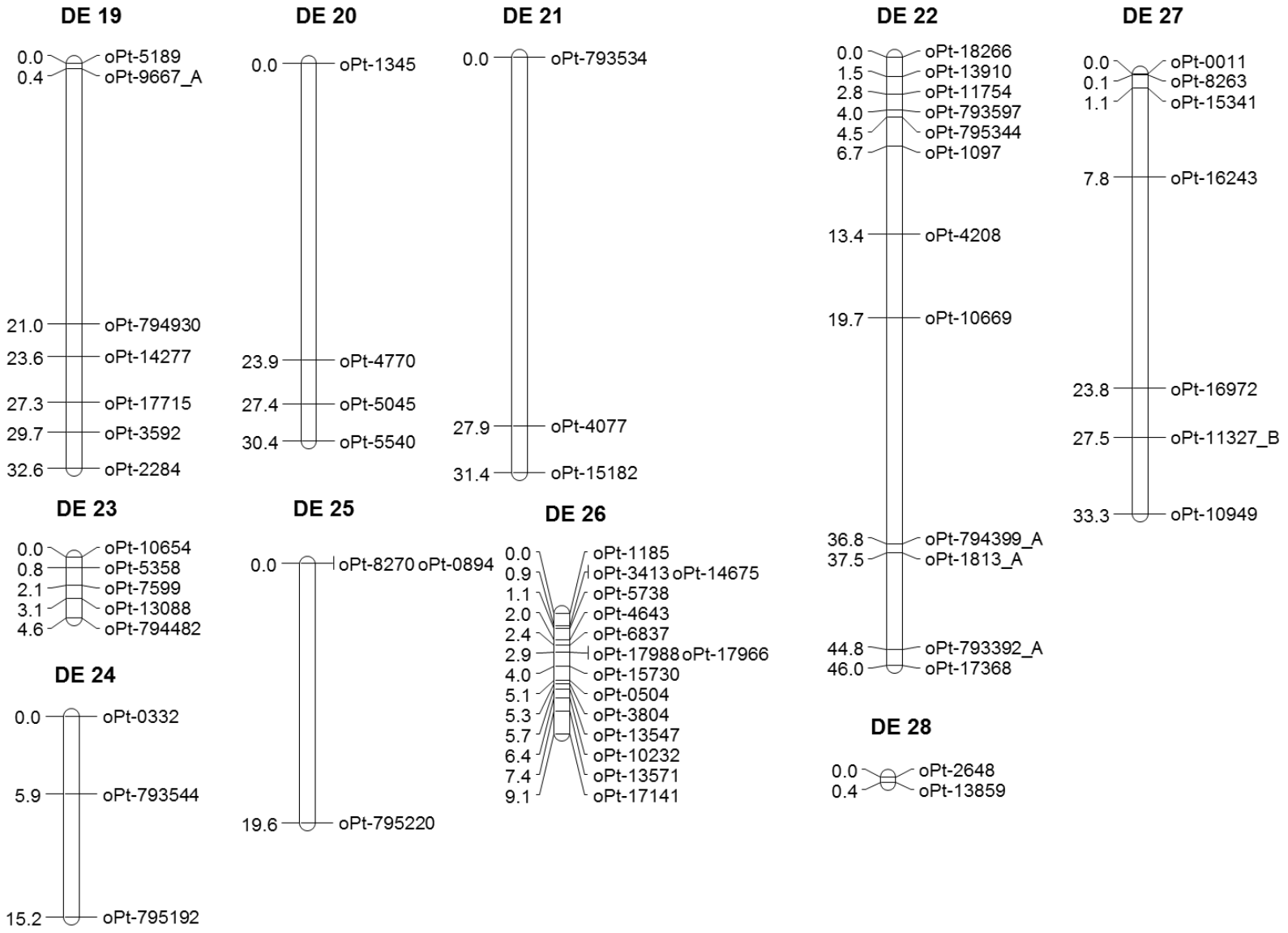
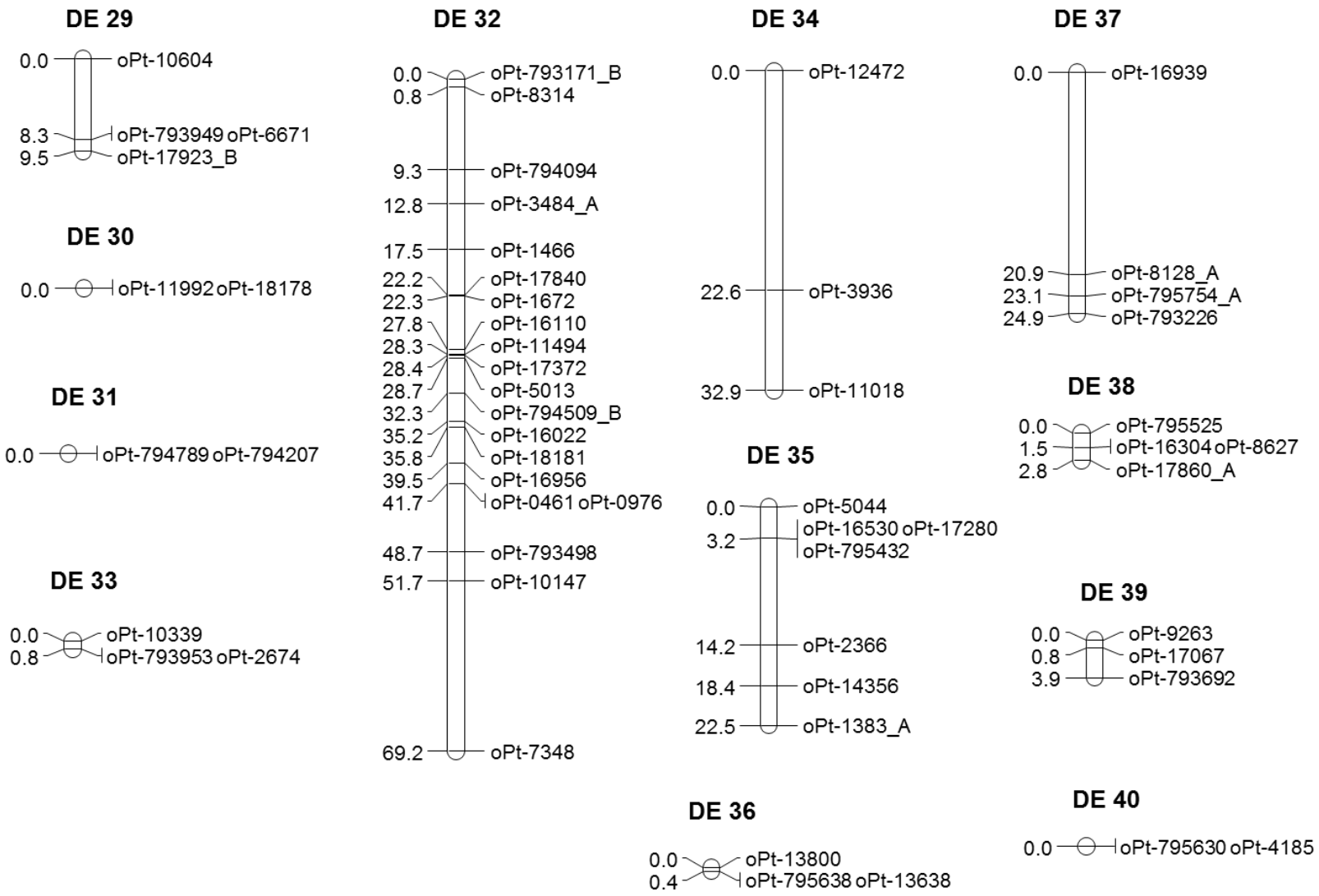


Figure S-2 Linkage map of the Dal X Exeter population continued…



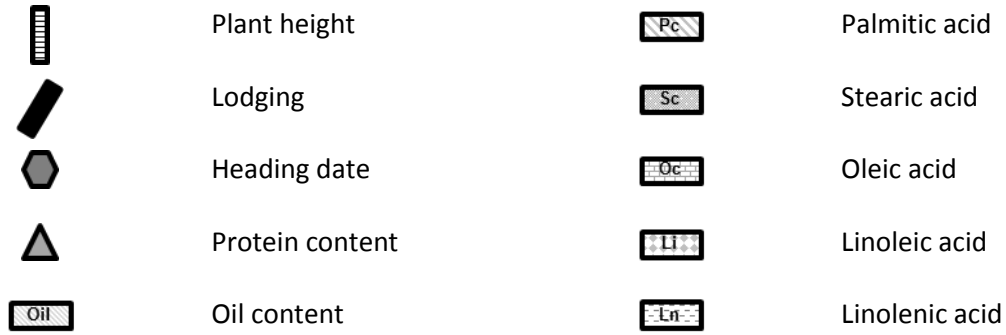
FigureS-2 Linkage map of the Dal X Exeter population continued….



FigureS-2 Linkage map of the Dal X Exeter population concluded

Appendix II- Comparative map of KxO vs DxE

Figure S-3 Comparative map of linkage maps of KxO vs DxE mapping populations of hexaploid oat.



NB. Stacked horizontal lines show the location of additional markers placed on the map framework. Solid lines are used to join markers that are named in the figure, while dotted lines join placed markers with hidden names. Boxes beside the KOD LGs show approximate locations of QTLs identified previous studies, some of which are placed on the map using comparative mapping (Wight et al. 2006).Numbers beside icons representing QTLs indicate the reference in which the QTL was reported, which are provided on the last page

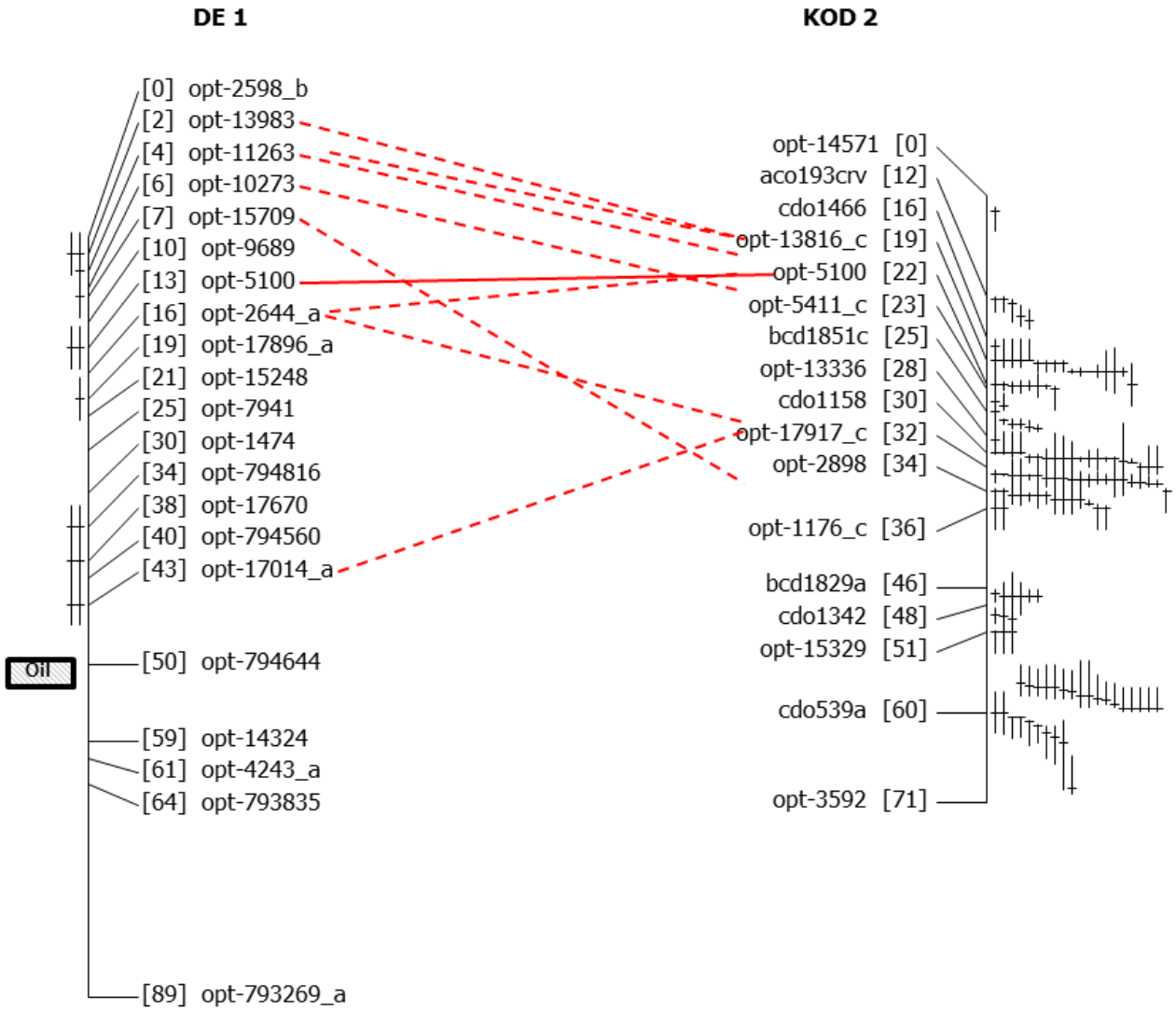

DE 1
KOD 2
[0] opt-2598_b
[2] opt-13983
[4] opt-11263
[6] opt-10273
[7] opt-15709
[10] opt-9689
[13] opt-5100
[16] opt-2644_a
[19] opt-17896_a
[21] opt-15248
[25] opt-7941
[30] opt-1474
[34] opt-794816
[38] opt-17670
[40] opt-794560
[43] opt-17014_a
Oil
[50] opt-794644
[59] opt-14324
[61] opt-4243_a
[64] opt-793835
[89] opt-793269_a
opt-14571 [0]
aco193crv [12]
cdo1466 [16]
opt-13816_c [19]
opt-5100 [22]
opt-5411_c [23]
bcd1851c [25]
opt-13336 [28]
cdo1158 [30]
opt-17917_c [32]
opt-2898 [34]
opt-1176_c [36]
bcd1829a [46]
cdo1342 [48]
opt-15329 [51]
cdo539a [60]
opt-3592 [71]

Figure S-3a

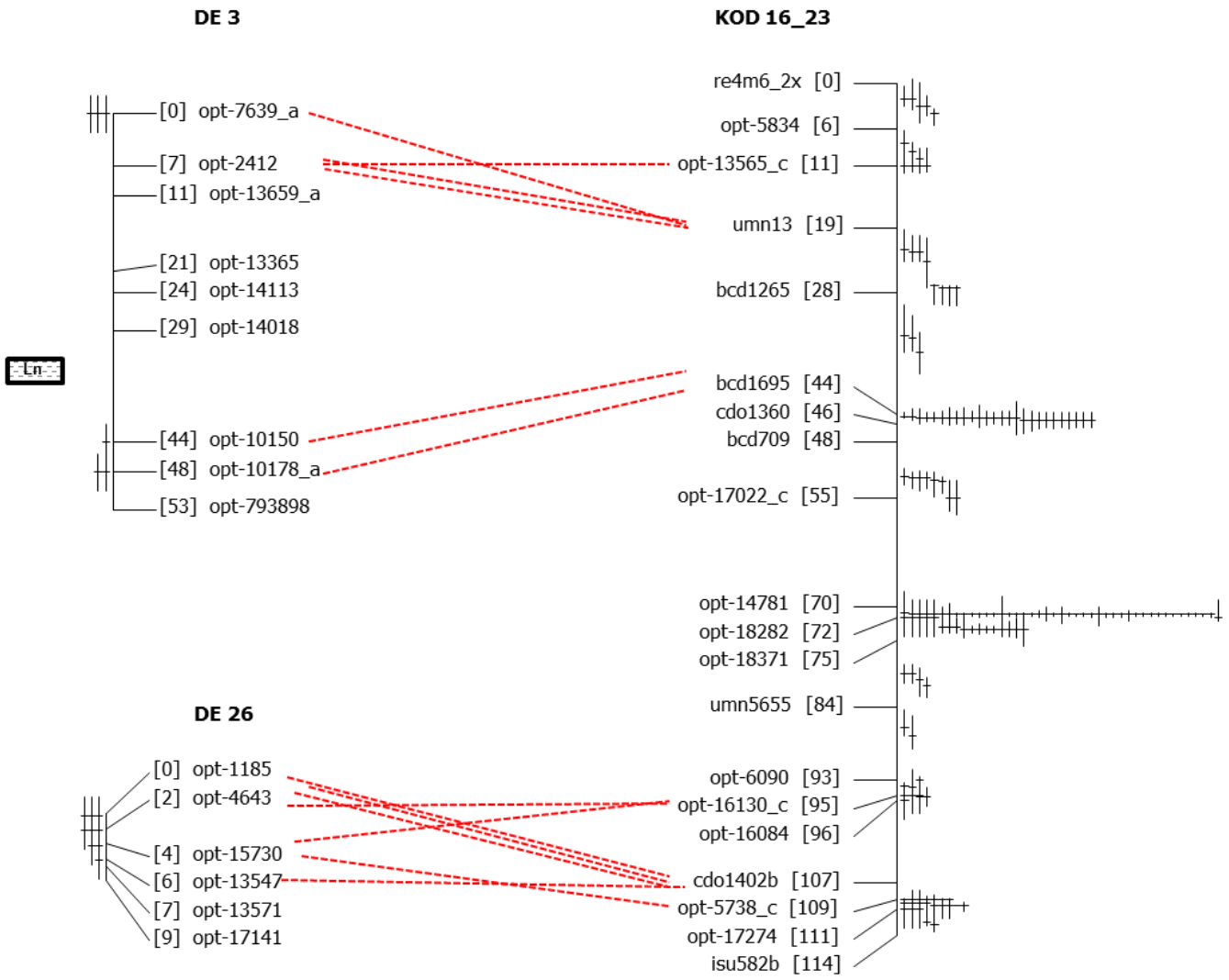

DE 3
KOD 16_23
[0] opt-7639_a
[7] opt-2412
[11] opt-13659_a
[21] opt-13365
[24] opt-14113
[29] opt-14018
[44] opt-10150
[48] opt-10178_a
[53] opt-793898
re4m6_2x [0]
opt-5834 [6]
opt-13565_c [11]
umn13 [19]
bcd1265 [28]
bcd1695 [44]
cdo1360 [46]
bcd709 [48]
opt-17022_c [55]
opt-14781 [70]
opt-18282 [72]
opt-18371 [75]
umn5655 [84]
DE 26
[0] opt-1185
[2] opt-4643
[4] opt-15730
[6] opt-13547
[7] opt-13571
[9] opt-17141
opt-6090 [93]
opt-16130_c [95]
opt-16084 [96]
cdo1402b [107]
opt-5738_c [109]
opt-17274 [111]
isu582b [114]

Figure S-3b

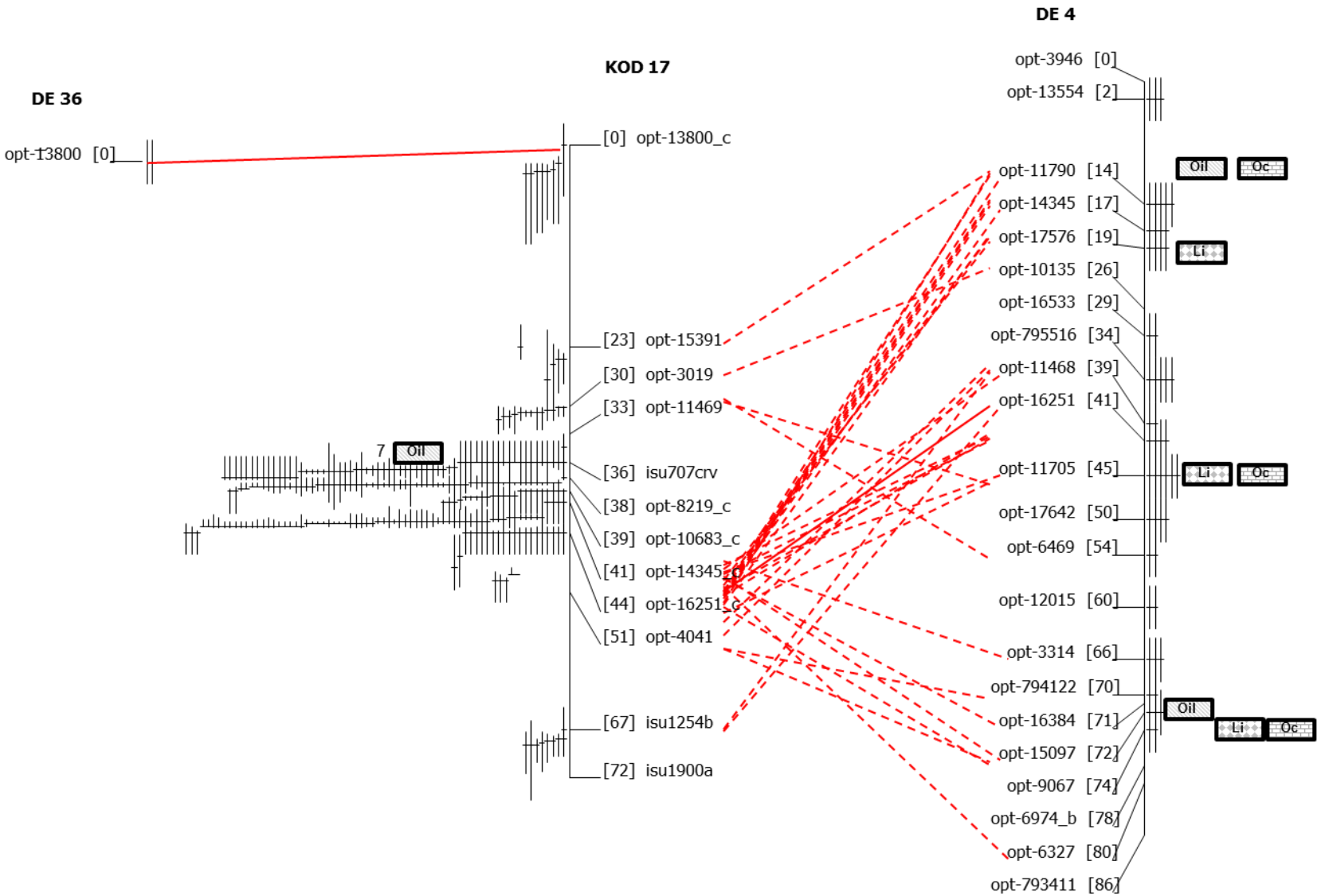

DE 36
KOD 17
DE 4
opt-13800 [0]
[0] opt-13800_c
[23] opt-15391
[30] opt-3019
[33] opt-11469
7 Oil
[36] isu707crv
[38] opt-8219_c
[39] opt-10683_c
[41] opt-14345_c
[44] opt-16251_c
[51] opt-4041
[67] isu1254b
[72] isu1900a
opt-3946 [0]
opt-13554 [2]
opt-11790 [14]
opt-14345 [17]
opt-17576 [19]
opt-10135 [26]
opt-16533 [29]
opt-795516 [34]
opt-11468 [39]
opt-16251 [41]
opt-11705 [45]
opt-17642 [50]
opt-6469 [54]
opt-12015 [60]
opt-3314 [66]
opt-794122 [70]
opt-16384 [71]
opt-15097 [72]
opt-9067 [74]
opt-6974_b [78]
opt-6327 [80]
opt-793411 [86]
Oil
Oc
Li
Oil
Oc
Li
Oil
Li
Oc

Figure S-3c

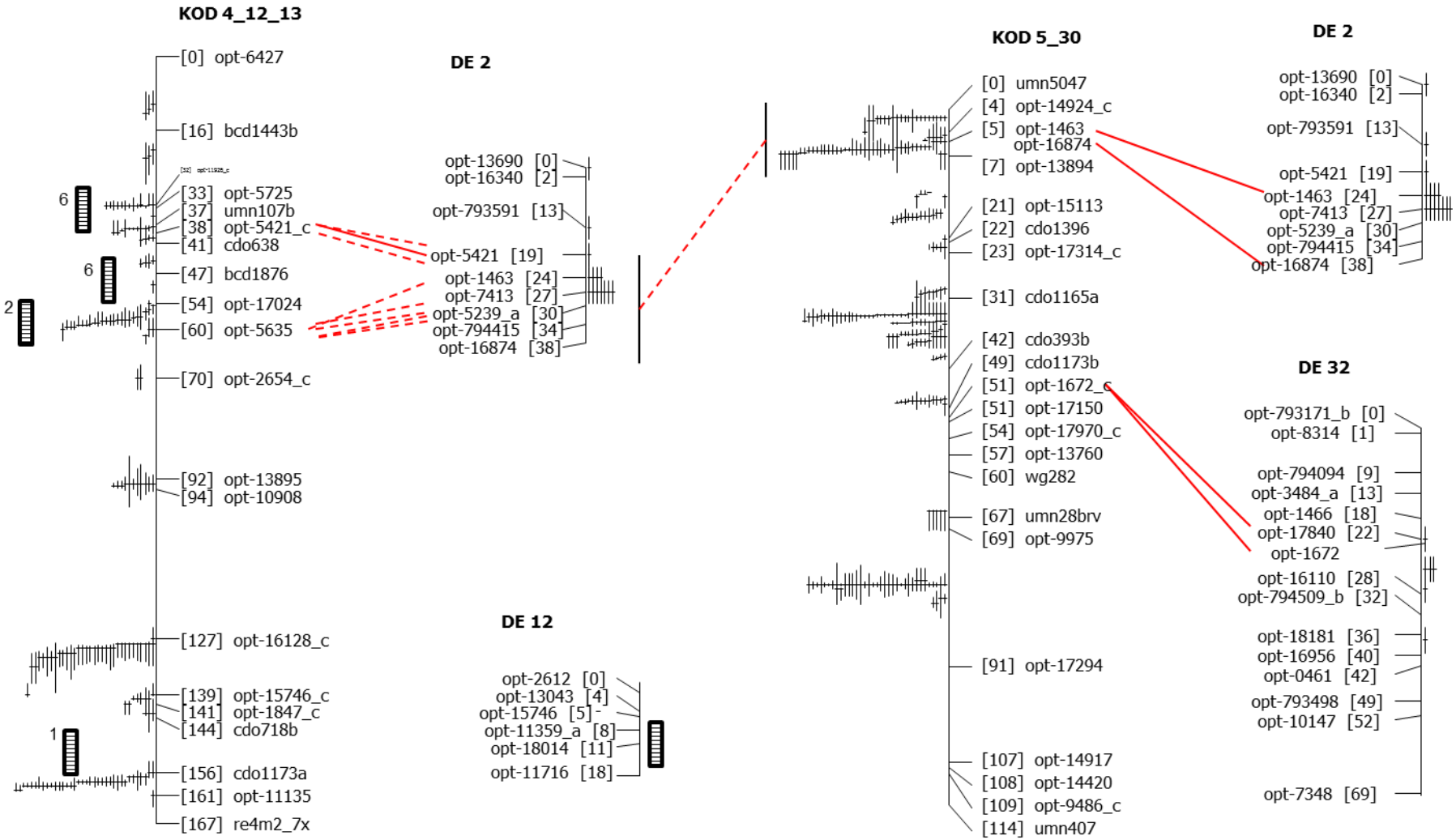

KOD 4_12_13
[0] opt-6427
[16] bcd1443b
[33] opt-5725
[37] umn107b
[38] opt-5421_c
[41] cdo638
[47] bcd1876
[54] opt-17024
[60] opt-5635
[70] opt-2654_c
[92] opt-13895
[94] opt-10908
[127] opt-16128_c
[139] opt-15746_c
[141] opt-1847_c
[144] cdo718b
[156] cdo1173a
[161] opt-11135
[167] re4m2_7x
6
6
2
1
DE 2
opt-13690 [0]
opt-16340 [2]
opt-793591 [13]
opt-5421 [19]
opt-1463 [24]
opt-7413 [27]
opt-5239_a [30]
opt-794415 [34]
opt-16874 [38]
DE 12
opt-2612 [0]
opt-13043 [4]
opt-15746 [5]
opt-11359_a [8]
opt-18014 [11]
opt-11716 [18]
KOD 5_30
[0] umn5047
[4] opt-14924_c
[5] opt-1463
opt-16874
[7] opt-13894
[21] opt-15113
[22] cdo1396
[23] opt-17314_c
[31] cdo1165a
[42] cdo393b
[49] cdo1173b
[51] opt-1672_c
[51] opt-17150
[54] opt-17970_c
[57] opt-13760
[60] wg282
[67] umn28brv
[69] opt-9975
[91] opt-17294
[107] opt-14917
[108] opt-14420
[109] opt-9486_c
[114] umn407
DE 2
opt-13690 [0]
opt-16340 [2]
opt-793591 [13]
opt-5421 [19]
opt-1463 [24]
opt-7413 [27]
opt-5239_a [30]
opt-794415 [34]
opt-16874 [38]
DE 32
opt-793171_b [0]
opt-8314 [1]
opt-794094 [9]
opt-3484_a [13]
opt-1466 [18]
opt-17840 [22]
opt-1672
opt-16110 [28]
opt-794509_b [32]
opt-18181 [36]
opt-16956 [40]
opt-0461 [42]
opt-793498 [49]
opt-10147 [52]
opt-7348 [69]

Figure S-3d

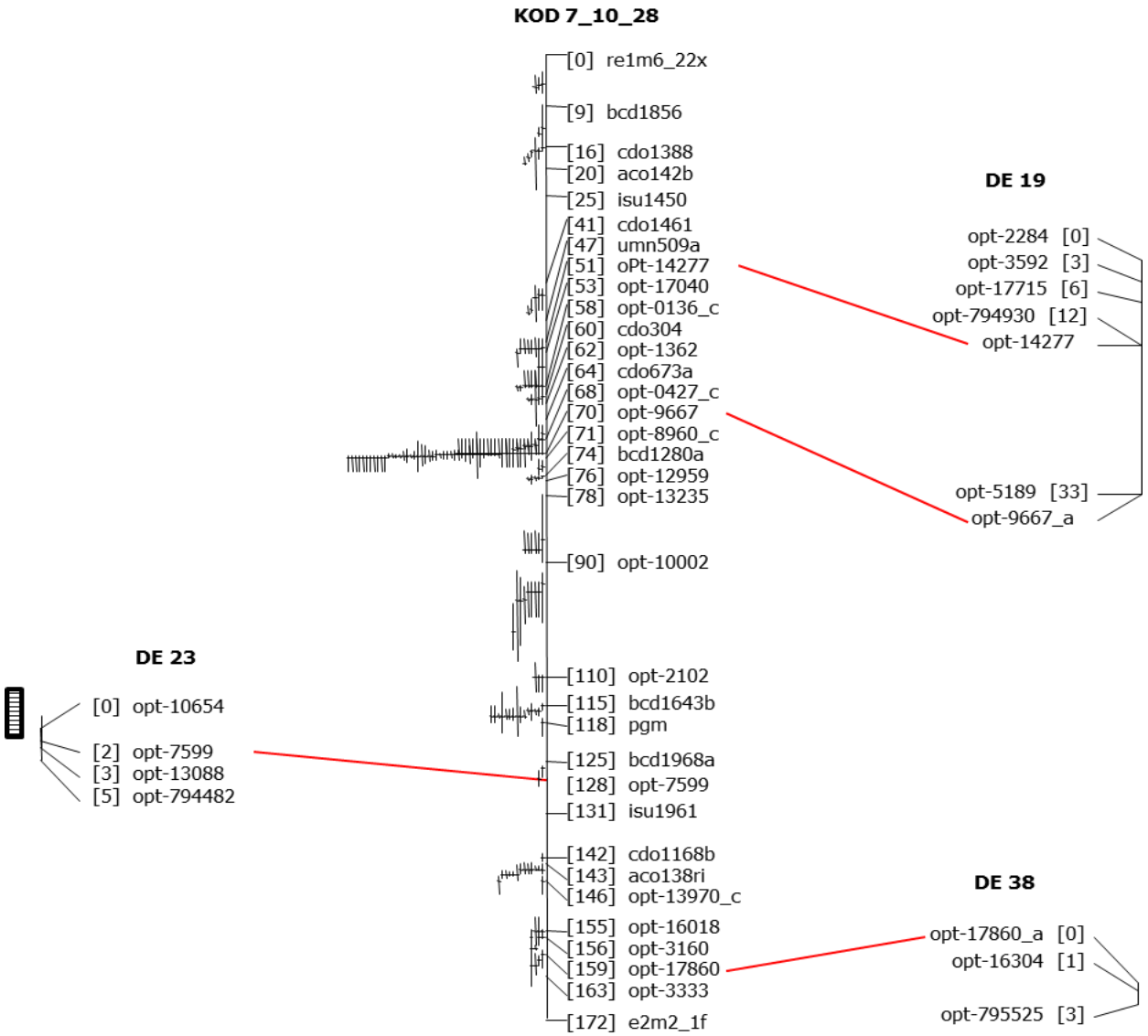

KOD 7_10_28
[0] re1m6_22x
[9] bcd1856
[16] cdo1388
[20] aco142b
[25] isu1450
[41] cdo1461
[47] umn509a
[51] oPt-14277
[53] opt-17040
[58] opt-0136_c
[60] cdo304
[62] opt-1362
[64] cdo673a
[68] opt-0427_c
[70] opt-9667
[71] opt-8960_c
[74] bcd1280a
[76] opt-12959
[78] opt-13235
[90] opt-10002
[110] opt-2102
[115] bcd1643b
[118] pgm
[125] bcd1968a
[128] opt-7599
[131] isu1961
[142] cdo1168b
[143] aco138ri
[146] opt-13970_c
[155] opt-16018
[156] opt-3160
[159] opt-17860
[163] opt-3333
[172] e2m2_1f
DE 19
opt-2284 [0]
opt-3592 [3]
opt-17715 [6]
opt-794930 [12]
opt-14277
opt-5189 [33]
opt-9667_a
DE 23
[0] opt-10654
[2] opt-7599
[3] opt-13088
[5] opt-794482
DE 38
opt-17860_a [0]
opt-16304 [1]
opt-795525 [3]

Figure S-3e

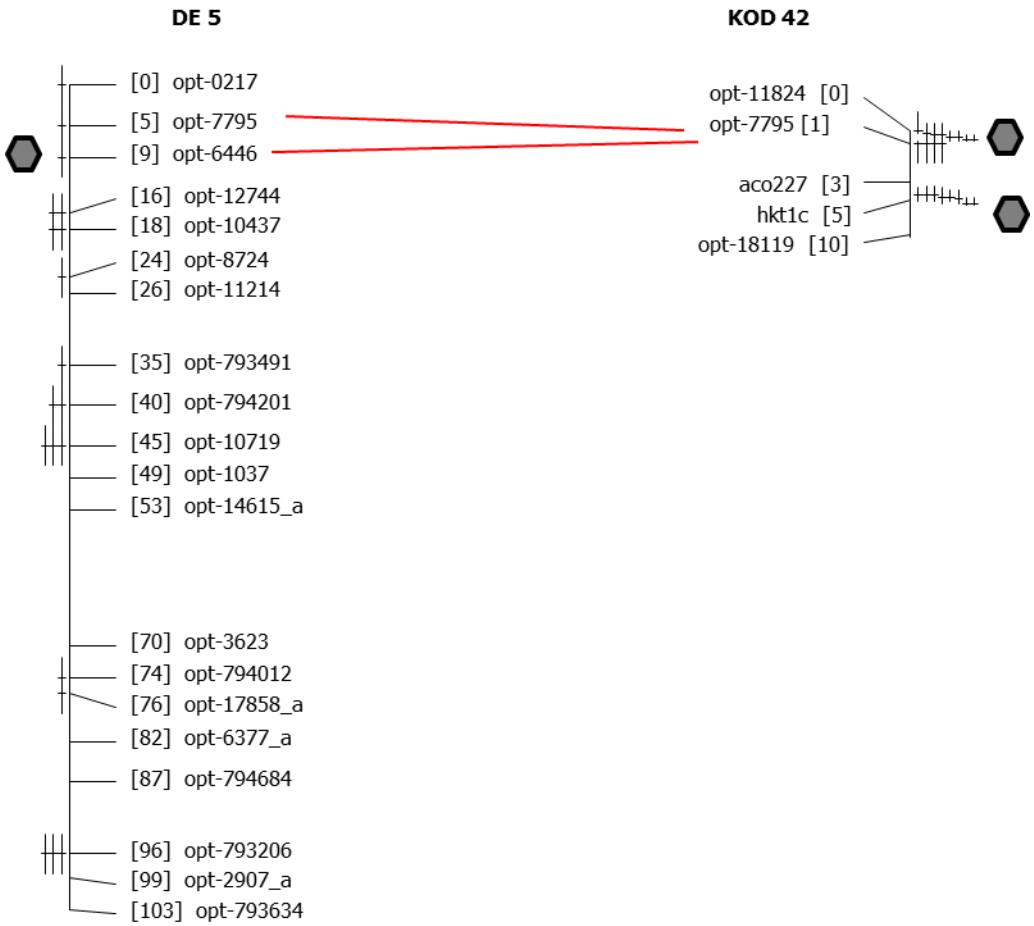

DE 5
KOD 42
[0] opt-0217
[5] opt-7795
[9] opt-6446
[16] opt-12744
[18] opt-10437
[24] opt-8724
[26] opt-11214
[35] opt-793491
[40] opt-794201
[45] opt-10719
[49] opt-1037
[53] opt-14615_a
[70] opt-3623
[74] opt-794012
[76] opt-17858_a
[82] opt-6377_a
[87] opt-794684
[96] opt-793206
[99] opt-2907_a
[103] opt-793634
opt-11824 [0]
opt-7795 [1]
aco227 [3]
hkt1c [5]
opt-18119 [10]

Figure S-3f

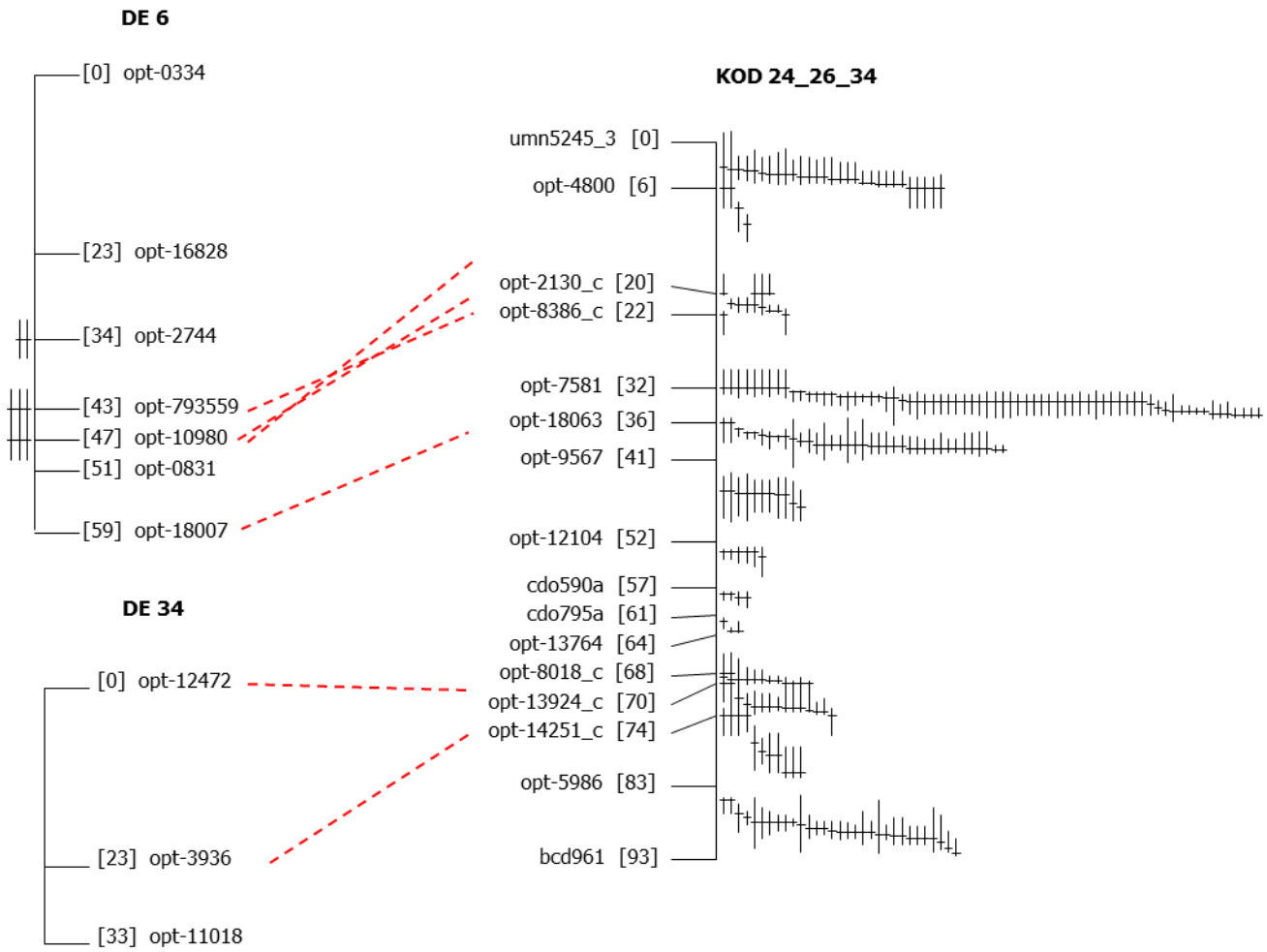

DE 6
[0] opt-0334
[23] opt-16828
[34] opt-2744
[43] opt-793559
[47] opt-10980
[51] opt-0831
[59] opt-18007
DE 34
[0] opt-12472
[23] opt-3936
[33] opt-11018
KOD 24_26_34
umn5245_3 [0]
opt-4800 [6]
opt-2130_c [20]
opt-8386_c [22]
opt-7581 [32]
opt-18063 [36]
opt-9567 [41]
opt-12104 [52]
cdo590a [57]
cdo795a [61]
opt-13764 [64]
opt-8018_c [68]
opt-13924_c [70]
opt-14251_c [74]
opt-5986 [83]
bcd961 [93]

Figure S-3g

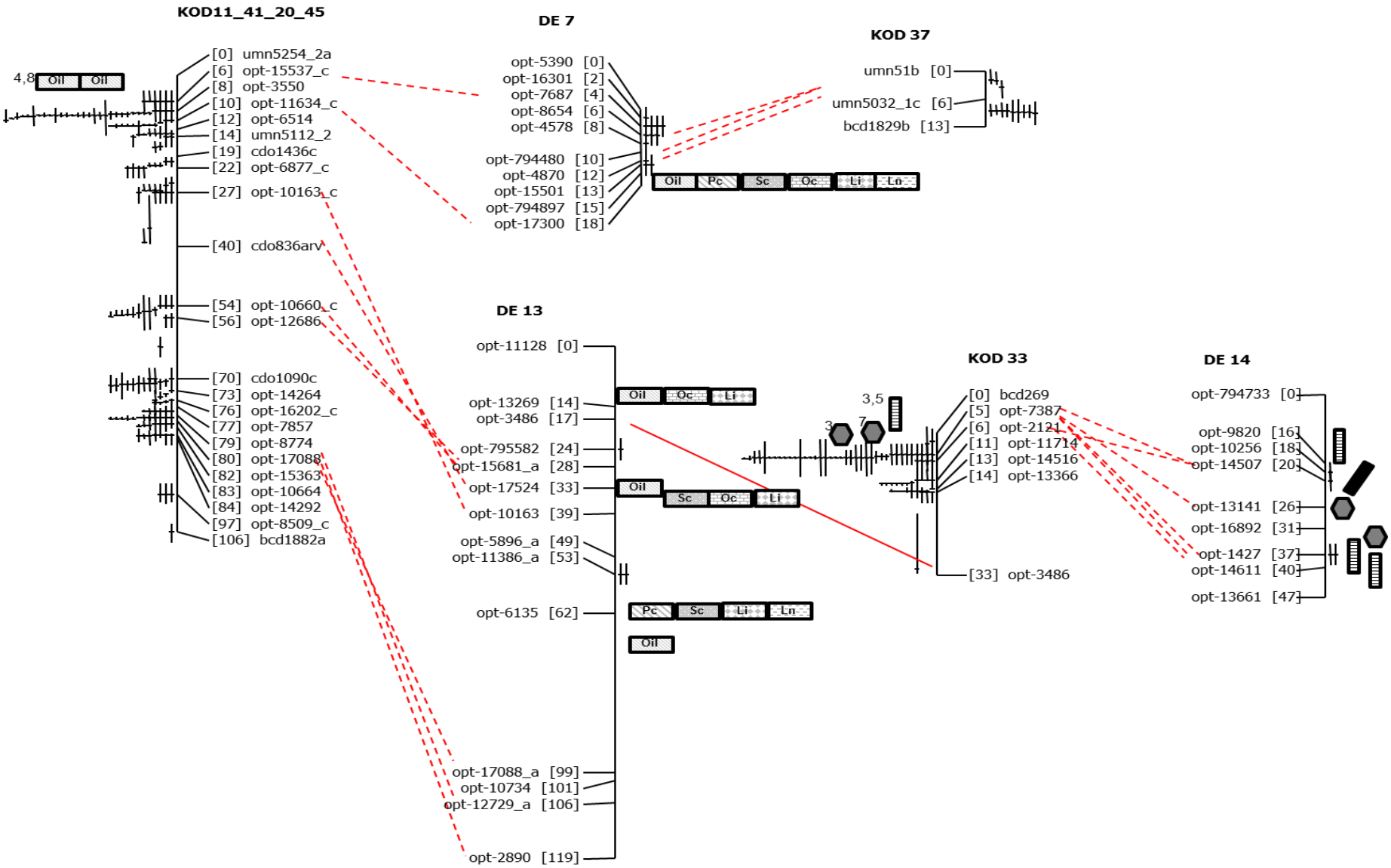

KOD11_41_20_45
4,8
Oil
Oil
[0] umn5254_2a
[6] opt-15537_c
[8] opt-3550
[10] opt-11634_c
[12] opt-6514
[14] umn5112_2
[19] cdo1436c
[22] opt-6877_c
[27] opt-10163_c
[40] cdo836arv
[54] opt-10660_c
[56] opt-12686
[70] cdo1090c
[73] opt-14264
[76] opt-16202_c
[77] opt-7857
[79] opt-8774
[80] opt-17088
[82] opt-15363
[83] opt-10664
[84] opt-14292
[97] opt-8509_c
[106] bcd1882a
DE 7
opt-5390 [0]
opt-16301 [2]
opt-7687 [4]
opt-8654 [6]
opt-4578 [8]
opt-794480 [10]
opt-4870 [12]
opt-15501 [13]
opt-794897 [15]
opt-17300 [18]
Oil
Pc
Sc
Oc
Li
Ln
KOD 37
umn51b [0]
umn5032_1c [6]
bcd1829b [13]
DE 13
opt-11128 [0]
opt-13269 [14]
opt-3486 [17]
opt-795582 [24]
opt-15681_a [28]
opt-17524 [33]
opt-10163 [39]
opt-5896_a [49]
opt-11386_a [53]
opt-6135 [62]
opt-17088_a [99]
opt-10734 [101]
opt-12729_a [106]
opt-2890 [119]
Oil
Oc
Li
Oil
Sc
Oc
Li
Pc
Sc
Li
Ln
Oil
KOD 33
3,5
3
7
[0] bcd269
[5] opt-7387
[6] opt-2121
[11] opt-11714
[13] opt-14516
[14] opt-13366
[33] opt-3486
DE 14
opt-794733 [0]
opt-9820 [16]
opt-10256 [18]
opt-14507 [20]
opt-13141 [26]
opt-16892 [31]
opt-1427 [37]
opt-14611 [40]
opt-13661 [47]

Figure S-3h

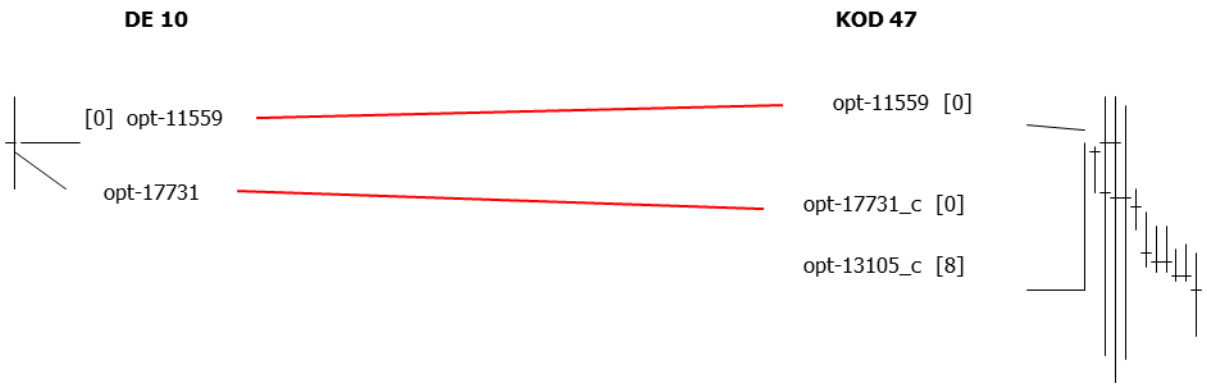

DE 10
KOD 47
[0] opt-11559
opt-17731
opt-11559 [0]
opt-17731_c [0]
opt-13105_c [8]

Figure S-3i

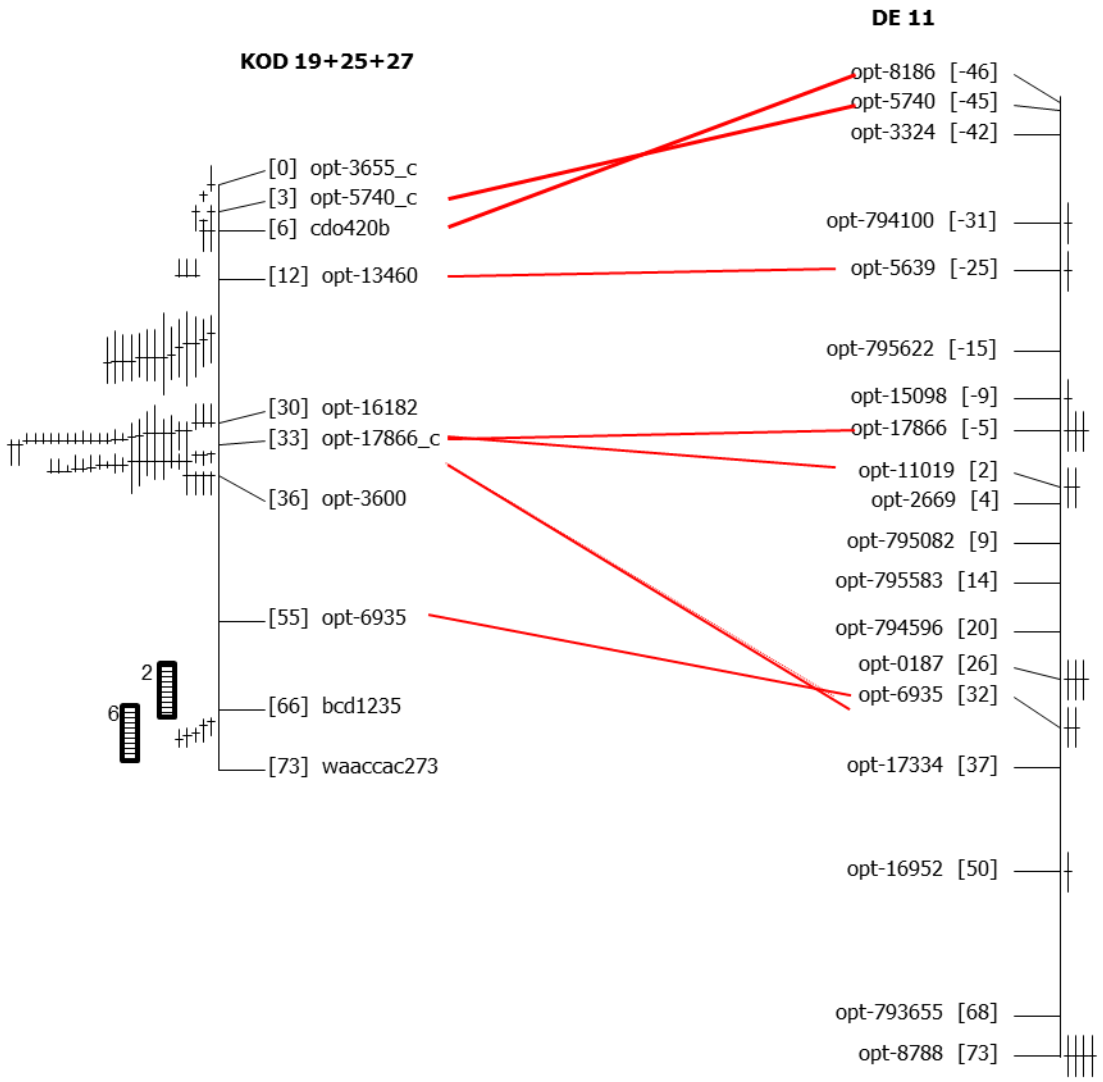

KOD 19+25+27
DE 11
[0] opt-3655_c
[3] opt-5740_c
[6] cdo420b
[12] opt-13460
[30] opt-16182
[33] opt-17866_c
[36] opt-3600
[55] opt-6935
[66] bcd1235
[73] waaccac273
2
6
opt-8186 [-46]
opt-5740 [-45]
opt-3324 [-42]
opt-794100 [-31]
opt-5639 [-25]
opt-795622 [-15]
opt-15098 [-9]
opt-17866 [-5]
opt-11019 [2]
opt-2669 [4]
opt-795082 [9]
opt-795583 [14]
opt-794596 [20]
opt-0187 [26]
opt-6935 [32]
opt-17334 [37]
opt-16952 [50]
opt-793655 [68]
opt-8788 [73]

Figure S-3j

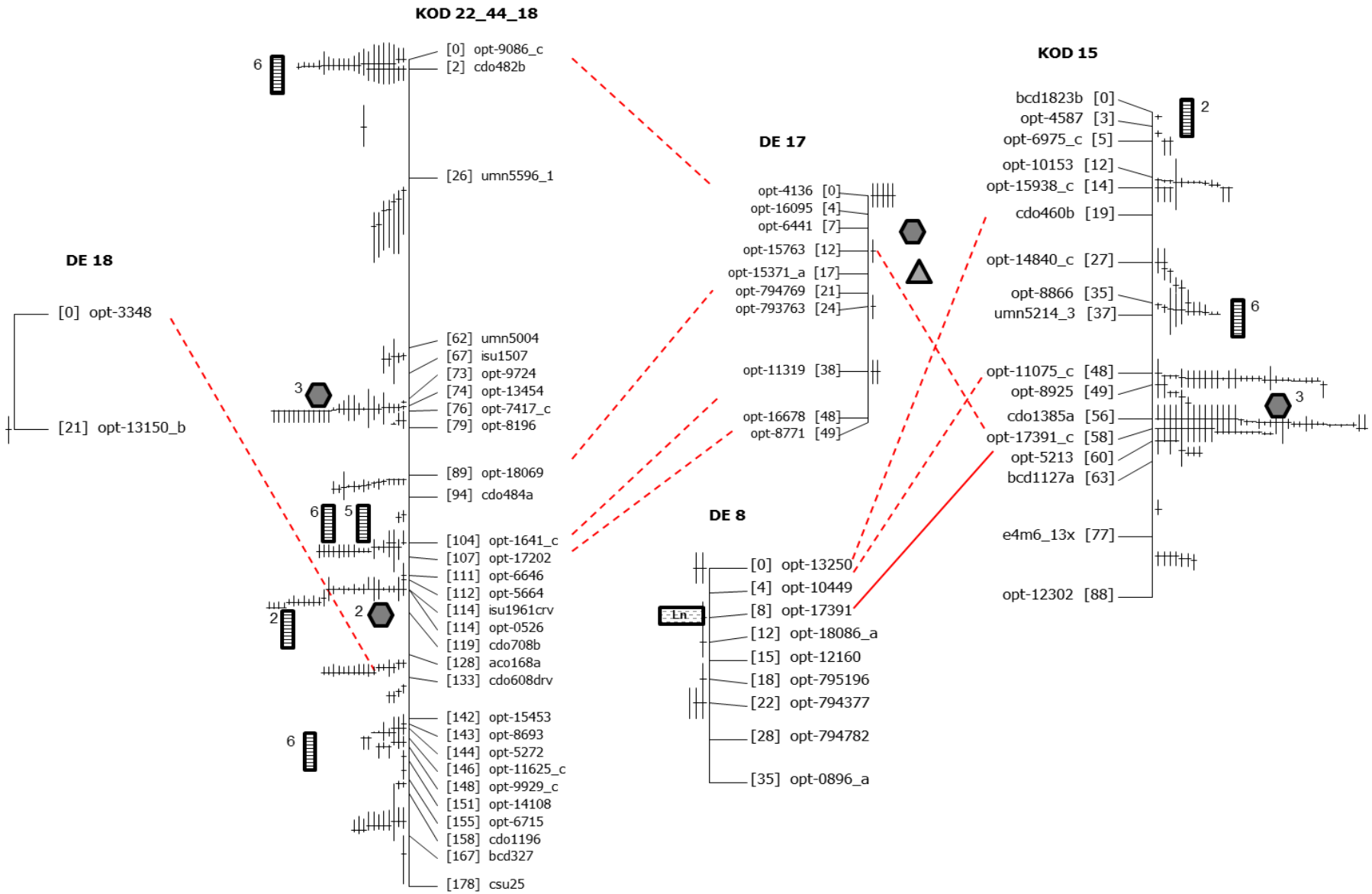

KOD 22_44_18
[0] opt-9086_c
[2] cdo482b
[26] umn5596_1
[62] umn5004
[67] isu1507
[73] opt-9724
[74] opt-13454
[76] opt-7417_c
[79] opt-8196
[89] opt-18069
[94] cdo484a
[104] opt-1641_c
[107] opt-17202
[111] opt-6646
[112] opt-5664
[114] isu1961crv
[114] opt-0526
[119] cdo708b
[128] aco168a
[133] cdo608drv
[142] opt-15453
[143] opt-8693
[144] opt-5272
[146] opt-11625_c
[148] opt-9929_c
[151] opt-14108
[155] opt-6715
[158] cdo1196
[167] bcd327
[178] csu25
DE 18
[0] opt-3348
[21] opt-13150_b
DE 17
opt-4136 [0]
opt-16095 [4]
opt-6441 [7]
opt-15763 [12]
opt-15371_a [17]
opt-794769 [21]
opt-793763 [24]
opt-11319 [38]
opt-16678 [48]
opt-8771 [49]
DE 8
[0] opt-13250
[4] opt-10449
[8] opt-17391
[12] opt-18086_a
[15] opt-12160
[18] opt-795196
[22] opt-794377
[28] opt-794782
[35] opt-0896_a
KOD 15
bcd1823b [0]
opt-4587 [3]
opt-6975_c [5]
opt-10153 [12]
opt-15938_c [14]
cdo460b [19]
opt-14840_c [27]
opt-8866 [35]
umn5214_3 [37]
opt-11075_c [48]
opt-8925 [49]
cdo1385a [56]
opt-17391_c [58]
opt-5213 [60]
bcd1127a [63]
e4m6_13x [77]
opt-12302 [88]

Figure S-3k

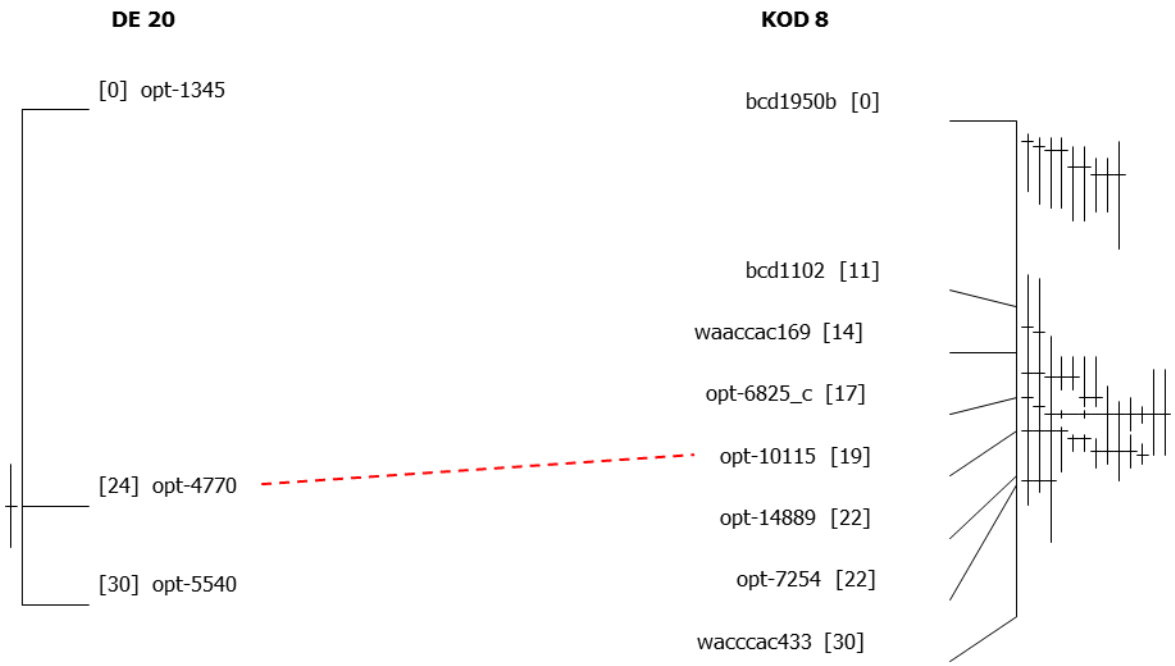

DE 20
KOD 8
[0] opt-1345
[24] opt-4770
[30] opt-5540
bcd1950b [0]
bcd1102 [11]
waaccac169 [14]
opt-6825_c [17]
opt-10115 [19]
opt-14889 [22]
opt-7254 [22]
wacccac433 [30]

Figure S-3l

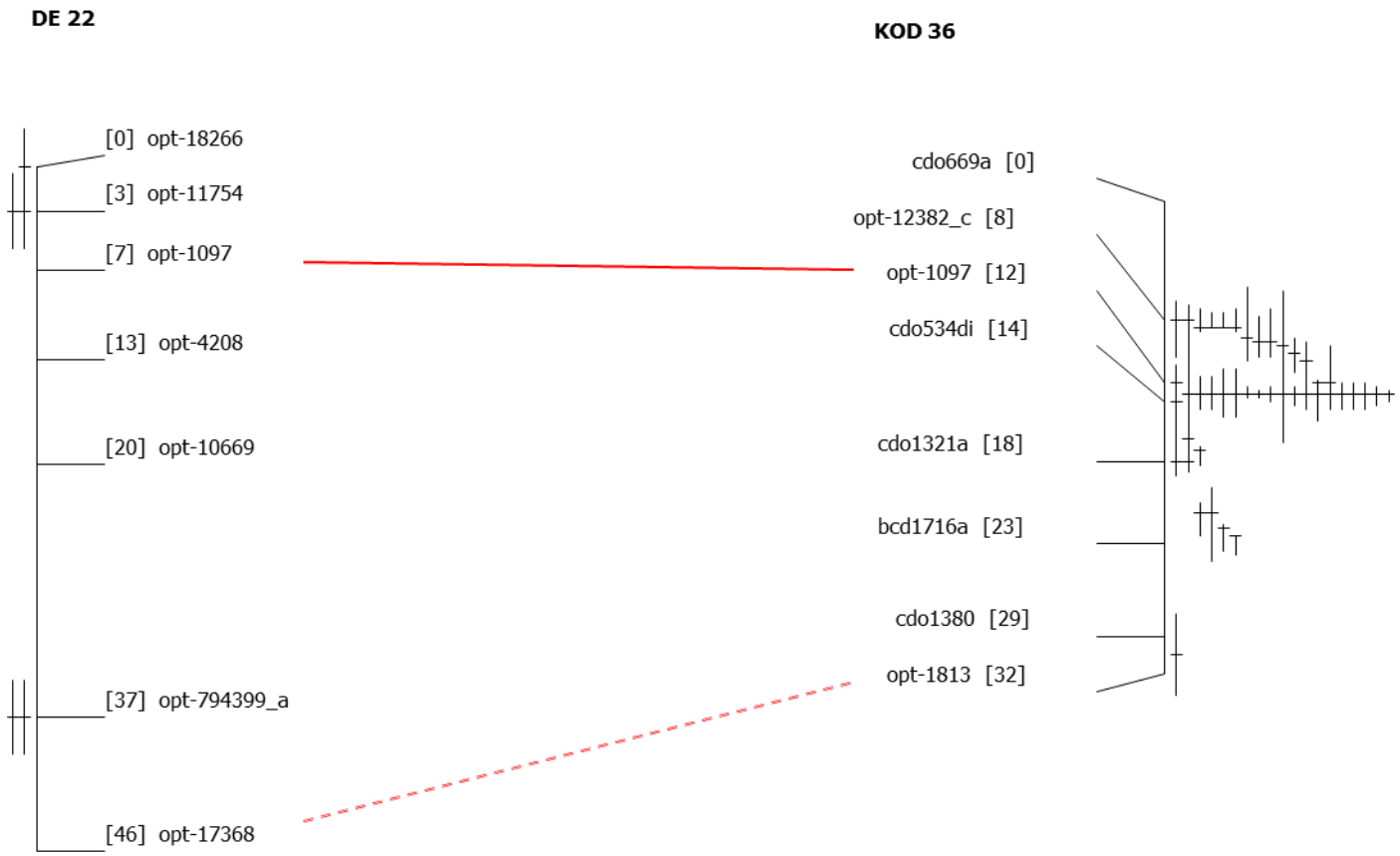

DE 22
KOD 36
[0] opt-18266
[3] opt-11754
[7] opt-1097
[13] opt-4208
[20] opt-10669
[37] opt-794399_a
[46] opt-17368
cdo669a [0]
opt-12382_c [8]
opt-1097 [12]
cdo534di [14]
cdo1321a [18]
bcd1716a [23]
cdo1380 [29]
opt-1813 [32]

Figure S-3m

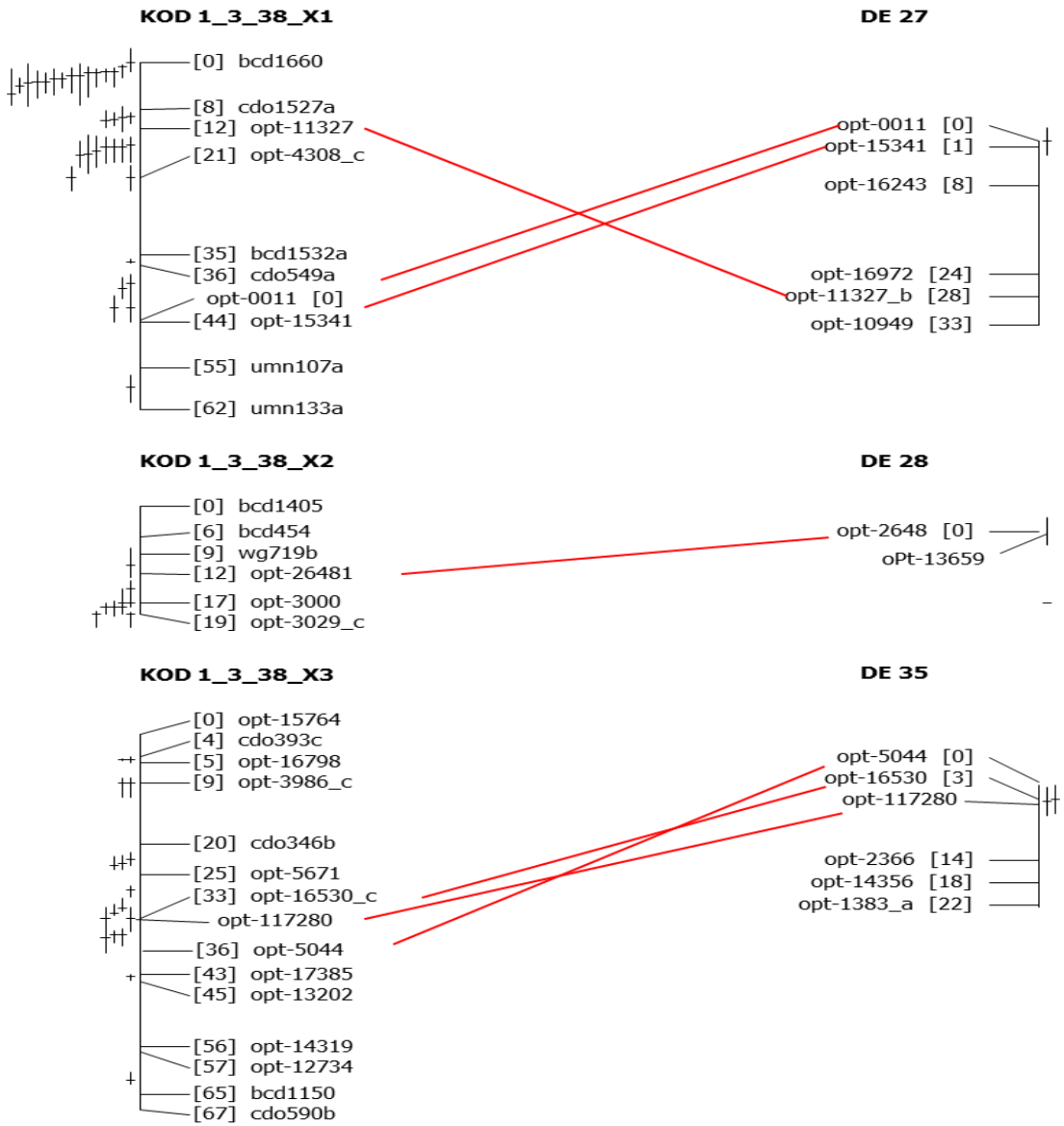

KOD 1_3_38_X1
[0] bcd1660
[8] cdo1527a
[12] opt-11327
[21] opt-4308_c
[35] bcd1532a
[36] cdo549a
opt-0011 [0]
[44] opt-15341
[55] umn107a
[62] umn133a
DE 27
opt-0011 [0]
opt-15341 [1]
opt-16243 [8]
opt-16972 [24]
opt-11327_b [28]
opt-10949 [33]
KOD 1_3_38_X2
[0] bcd1405
[6] bcd454
[9] wg719b
[12] opt-26481
[17] opt-3000
[19] opt-3029_c
DE 28
opt-2648 [0]
oPt-13659
KOD 1_3_38_X3
[0] opt-15764
[4] cdo393c
[5] opt-16798
[9] opt-3986_c
[20] cdo346b
[25] opt-5671
[33] opt-16530_c
opt-117280
[36] opt-5044
[43] opt-17385
[45] opt-13202
[56] opt-14319
[57] opt-12734
[65] bcd1150
[67] cdo590b
DE 35
opt-5044 [0]
opt-16530 [3]
opt-117280
opt-2366 [14]
opt-14356 [18]
opt-1383_a [22]

Figure S-3n

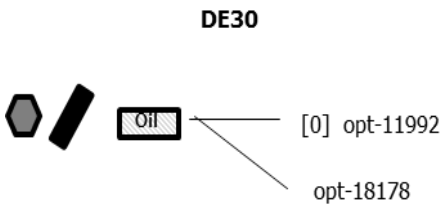


Figure S-3o

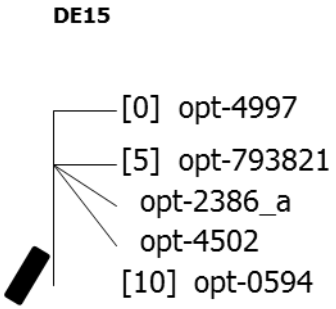


Figure S-3p

References for Figure S-3

1. Beer S.C., Siripoonwiwat W., O'donoughue L.S., Souza E., Matthews D., Sorrells M.E. (1997) Associations between molecular markers and quantitative traits in an oat germplasm pool: can we infer linkages. J. Agricultural Genomics 3:1-16.

2. Holland J., Portyanko V., Hoffman D., Lee M. (2002) Genomic regions controlling vernalization and photoperiod responses in oat. TAG Theoretical and Applied Genetics 105:113-126.

3. Holland J.B., Moser H.S., O'Donoughue L.S., Lee M. (1997) QTLs and epistasis associated with vernalization responses in oat. Crop Science 37:1306-1316.

4. Kianian, S. F., M. A. Egli, R. L. Phillips, H. W. Rines, D. A. Somers, B. G. Gengenbach, F. H. Webster, S. M..Livingston, S. Groh, and L. S. O'Donoughue, M. E. Sorrells, D. M. Wesenberg, D. D. Stuthman and R. G. Fulcher. 1999. Association of a major groat oil content QTL and an acetyl-CoA carboxylase gene in oat. Theoretical and Applied Genetics 98:884-894

5. Milach S.C.K., Rines H.W., Phillips R.L. (1997) Molecular genetic mapping of dwarfing genes in oat. Theoretical and Applied Genetics 95:783-790.

6. Siripoonwiwat, W., L. S. O'Donoughue, D. Wesenberg, D. L. Hoffman, J. F. Barbosa-Neto, and M. E. Sorrells. 1996. Chromosomal regions associated with quantitative traits in oat. Journal of Agricultural Genomics 2, Available at http://wheat.pw.usda.gov/jag/papers96/paper396/indexp396.html (verified 4 July, 2012)

7. Tanhuanpää P., Manninen O., Kiviharju E. (2010) QTLs for important breeding characteristics in the doubled haploid oat progeny. Genome 53:482-493.

8. Zhu S., Rossnagel B.G., Kaeppler H.F. (2004) Genetic analysis of quantitative trait loci for groat protein and oil content in oat. Crop Science 44:254-260.

Appendix III – Copyright permission

17 Aug. 2012

Dear Dr. Hizbai,

Permission for re-use of the following article:

"QTLs affecting oil content, oil composition, and other agronomically important traits in Oat (Avena sativa L.)" from *The Plant Genome* (in review),

is granted for use in your thesis, "Comparative mapping of QTLs affecting oil content, oil composition, and other agronomically important traits in Oat (*Avena sativa* L.)".

Sincerely,

Elizabeth Gebhardt

Managing Editor, *Crop Science*
ASA-SSSA-CSSA