Identification of glucosinolate profile in *Brassica oleracea* for Quantitative Trait Locus mapping

Reem Adnan Issa

A thesis submitted for the degree of Doctor of Philosophy in Chemistry

Department of Chemistry, University of Warwick

April 2010

THE UNIVERSITY OF
WARWICK

LIST OF CONTENT

I

LIST OF FIGURES

VII

LIST OF TABLES

XI

LIST OF EQUATIONS

LIST OF DIAGRAMS

LIST OF MAPS

Map 5 *Brassica oleracea* linkage map based on AGDH population with QTLs for aliphatic glucosinoltes, that were expected to be under the control of **major gene** effect. Alignment of conserved areas between the *B. oleracea* linkage map and *B. rapa* map with the *A. thaliana* map.

LIST OF ABBREVIATIONS

APCI	Atmospheric pressure chemical ionization
A. thaliana	*Arabidopsis thaliana*
A	A genome of *B. rapa* (n=10)
At	At genome of *A. thaliana* (n=5)
B. oleracea	*Brassica oleracea*
B. rapa	*Brassica rapa*
CIM	Composite Interval Mapping
cM	Centi-Morgan
CYP	Cytochromes P450
C	C genome of *B. oleracea* (n=9)
Da	Dalton, units used for measuring the *m/z* value in the mass spectrum
DH	Double Haploid
DIM	3,3'-diindolymethane, a catabolic product of I3C
DRs	Double Recombinant scores
ESI	Electro-spray Ionization
GC	Gas Chromatography
GST	Glutathione S-transferase
HPLC	reverse phase High Performance Liquid Chromatography
IC$_{50}$	The half-maximal inhibitory concentration
IM	Interval Mapping
IPM	Isoropyl malate
IPMS	Isopropyl malate synthase
IS	Internal Standard material used prior to the extraction of the desulfated glucosinolates for the validation of the extraction, desulfation, separation and quantification proposes
IS1	Glucosinolate naturally absent in plant material (glucotropaeolin), used as an Internal Standard

IS2	Internal standard added to each sample at a fixed concentration prior to sample injection onto the *HPLC* system to reduce the variability in the quantitative measurements due to variation in the injection volume performed by the auto sampler
I3C	Indole-3-Carbinol
LOD	Log of Odds
LG	Linkage group of the C genome, referred to the *B. oleracea* chromosomes (n=9)
MAMs	Methylthio Alkyl Malate synthase
M. brassicae	*Mamestra brassicae*
MQM	Multiple QTL Mapping
MS	Mass Spectrometry
MS/MS	Tandem *MS*
m/z	Mass over charge value for the ions detected in the mass spectrum
NIRS	Near Infrared Spectroscopy
NMR	Nuclear Magnetic Resonance
PDA	Photo Diode Array detector
PPF	Photosynthetic Photon Flux
P .rapea	*Pieris rapae*
QC	Quality Control
QTL	Quantitative Trait Locus
R	Side-chain structure of the glucosinolates
RRF	Relative Response Factor
RT	Retention Time
SF	Sulforaphane
SPE	Solid Phase Extraction
U	Units of sulfatase enzyme
UV	Ultra Violet
Win QTL Cart	Windows QTL Cartographer ver 2.5 program

GLOSSARY

Additive effect	Calculated as half the difference between the trait means for alleles homozygous to each parent at each marker
Allele	Any one of a series of two or more different genes that occupy the same position (locus) on a chromosome
Back cross	To mate the progeny of a cross with one of its parents
CentiMorgans (cM)	A measure of recombination frequency. This unit of linkage refers to the distance between two loci based on the number of recombination events occurring between them. Two loci are said to be 1 cM apart if recombination is observed between them in 1% of meiosis events
Chromosome	A linear end-to-end arrangement of genes and other DNA that is packaged with nucleoproteins in the cell nucleus and contains genetic information. Independent assortment among, and recombination between chromosomes are responsible for the pattern of transmission of hereditary characteristics
Confidence interval	An interval of values bonded by confidence limits within which the true value of a population parameter is stated to lie with a specified probability
Locus (pl. Loci)	The position that a given gene or genes occupy on a chromosome
Linkage group (LG)	The set of genes on a chromosome that tend to be transmitted together

LOD score	The $\log_{10}$ of the likelihood ratio (variance ratio), statistically determined as the Log of Odds
Mapping population	A collection of progeny generated from the cross between two parental lines that is used for genetic analysis to create a genetic map, or to identify the positions of genetic loci that influence a particular trait
Marker	An identifying factor, a gene or other DNA of know location and effect which is used to track the inheritance etc. of other genes whose exact location is not yet known
Phenotype	The constitution of organism as determined by the interaction of its genetics constitution and the environment
Polygene	Pertaining to the combined action of alleles of more than one gene
Progeny	Offspring
Quantitative Trait Loci (QTL)	Genetic loci that influence quantitative traits. QTL analysis integrates molecular marker linkage maps with data for quantitative traits to give information on the effects and locations of the loci controlling them
Trait	A characteristic feature or quality that can be scored or measured
Variance	A measure of the variation shown by a set of observations; defined as the sum of squares of deviations from the mean, divided by the of degrees of freedom in the set of observations
The standard procedures	described in the official Journal of the European communities (EEC, 1990)

ACKNOWLEDGEMENTS

ABSTRACT

Glucosinolates are a group of secondary plant metabolites, which have been shown to play important roles in human health and nutrition. Identification of novel genes or regulators of expression are important for optimising the glucosinolate composition of *Brassica* crops. This project aimed to develop a HPLC based methodology for quantifying these compounds within *Brassica* leaf material and to use this to map Quantitative Trait Loci for individual glucosinolates within *Brassica oleracea* mapping populations. Glucosinolates were analysed using an *optimized HPLC-UV* method developed in this study for complete separation of desulfated glucosinolates with high resolution for quantification measurements. *The reproducibility of the desulfation reaction was improved* for robust enzymatic reaction of sulfatase. *A data dependent MS and MS/MS* methodology was *developed* to confidently identify seven glucosinolates in the 89 AGDH plant lines distributed between aliphatic and indolic glucosinolate, with different combinations from the parental plants A12DHd and GDDH33. For the quantitative measurements of glucosinolates, an optimized level of glucotropaeolin was used as an internal standard (IS1). In addition, we have demonstrated the *first use of a second internal standard* (IS2) to significantly improve the reproducibility of the quantitative measurements. Aliphatic glucosinolates were predominant over indolic glucosinolates, where progoitrin has the highest abundance.

This methodology was then used to identify Quantitative Trait Loci for individual glucosinolates and for key points in their biosynthesis. A major gene effect was found near the top of *B. oleracea* LG9 associated with aliphatic glucosinolate synthesis. In addition other Quantitative Trait Loci were identified which corresponded with previous work by other groups and to which individual gene function could be attributed. A number of *novel* Quantitative Trait Loci were also found which control the synthesis of glucosinolates distributed on the nine chromosomes of C genome. A combination of the quantitative data and genetic analysis of glucosinolate profiles was used to infer the existence of factors at distinct loci and associated these with specific steps in the biosynthesis pathway of glucosinolates in *B. oleracea.* The assignment of genes or gene regulator functions to Quantitative Trait Loci identified in this study was consistent with known positions of *Brassica* candidate genes and collinear regions of the *Arabidopsis* genome. Consequently, this information can be applied to other *Brassica* species for breeding vegetable crops with modified glucosinolate profiles.

Introduction

1.1. Glucosinolates and their occurrence

Glucosinolates are a uniform class of ß-thioglucosides, derived from amino acids and grouped according to their side-chain structure (R). From Figure 1, it can be seen that all glucosinolates share the same aglycone structure and vary in their side-chain representing the precursor amino acids that were involved in the biosynthesis of the individual compound (Schonhof et al., 2004).
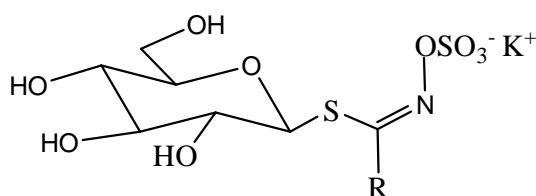
Figure 1 The general structure of glucosinolate (Davidson et al., 2001; Fenwick et al., 1983)

Glucosinolates are plant secondary metabolites that are implicated in decreasing the risk of cancer as a part of a vegetable-rich diet (1.7). They are found in all members of the cruciferous plants, including the *Brassica* crops, Brussels sprouts, broccoli, cauliflower, cabbage, watercress, oilseed rape and mustard. They are responsible for the pungent aroma of these plants (Higdon et al., 2007). Previous studies have focused on glucosinolate natural products from *Brassica,* and a list of structures, names and molecular masses of glucosinolates identified and isolated from different *Brassica oleracea* (*B. oleracea*) species by other research groups is included in Appendix A (Cartea and Velasco, 2008).

The first study regarding the properties of glucosinolates and isothiocyanates was at the beginning of the 17th century; this was followed by the isolation of sinigrin and

1

sinalbin in the 1830s. In 1956, the correct structure and the first chemical glucosinolate synthesis were proposed. Since the 1970s, many studies have focused on the beneficial biological effects of glucosinolates, and their breakdown products, on human and animal nutrition (Cartea and Velasco, 2008).

Plant glucosinolate content is typically a quantitative trait, under polygenetic control and the influence of environmental conditions (Kliebenstein, 2009). The value of identifying the Quantitative Trait Locus (QTL) responsible for the metabolite content in plants as a valuable contribution to plant breeding strategies and the biotechnology industry has been highlighted (Keurentjes et al., 2006). QTL analysis in segregating plant populations has been used to detect the presence of loci affecting metabolite profiles associated with particular synthesis and modification pathways (Gao et al., 2007; Kliebenstein, 2009; Lou et al., 2008). Further investigations of the loci affecting the glucosinolate profile in *B. oleracea* are one of the objectives of this presented work. For example, identification of QTLs controlling seed fatty acid synthesis and the modification pathway in the *Brassica* C genome may have an influence on developing economically viable oilseed crops with modified fatty acid profiles and maximizing the energy efficient yield of oils within crop species (Barker et al., 2007).

1.2. **History, diversity and evolution of *Brassica oleracea***

The diploid species of the genus *Brassica* in the Brassicinae family, have elicited attention because of their complex genetic relationships and the utilization of many of the species as vegetable, oil, and fodder crops (Lanner et al., 1997). The genus *Brassica*, contains a number of important agricultural species, such as rape seed (*B. napus*), cabbage (*B. oleracea*), turnip rape (*B.rapa*) and mustard (*B. nigra, B. juncea*

and *B. carinata*). The genomic relationship between these species is referred to as the triangle of U (U, 1935) (Diagram 1), where the basic diploid species has been classified cytogenetically as (*B. rapa* AA; 2n=20)*, (*B. nigra* BB; 2n=16) and (*B. oleracea* CC; 2n=18) interbreed interspecifically with one another to form three new allotetraploid species (*B. juncea* AABB; 2n=36), (*B. carinata* BBCC; 2n=34) and (*B. napus* AACC; 2n= 38).
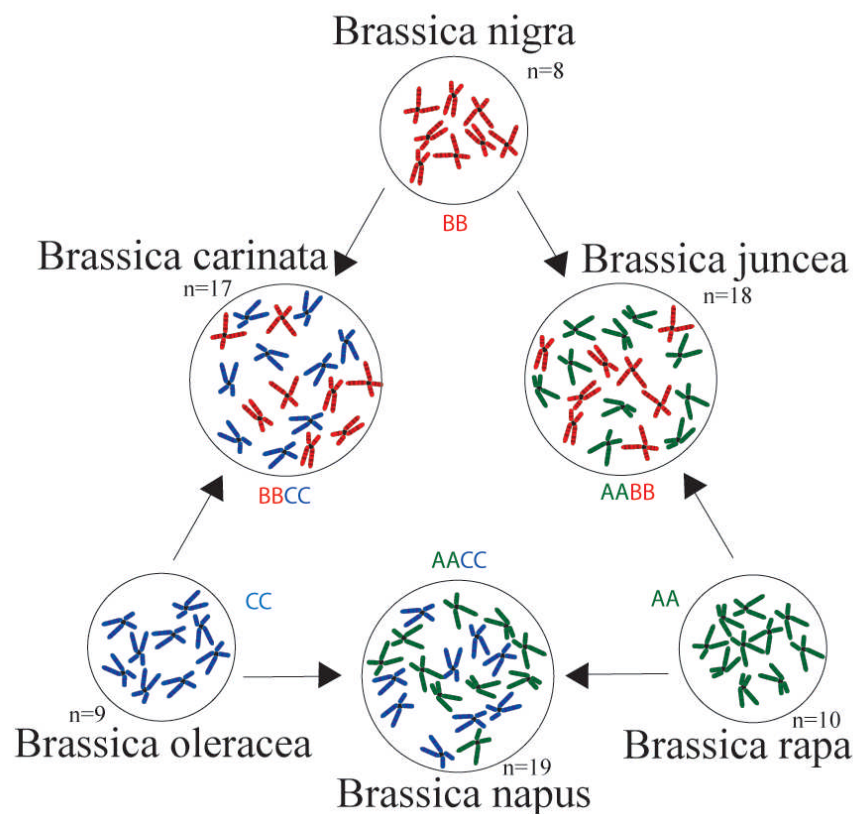
Diagram 1 The "Triangle of U" diagram, showing the genetic relationships between the six species of the genus *Brassica*. Chromosomes from each of the genomes A, B and C are represented by different colours (U, 1935).

In particular, the species *B. oleracea* displays an interesting genetic diversity, represented by 14 cultivated type (Dias, 1995). It has almost twice the chromosome number and four times the genome size of *Arabidopsis* (Suzuki et al., 2003). Although they have diverged from the same ancestor 14.5-20.4 million years ago

(Gao et al., 2004), it has been hypothesized that *B. oleracea* diverged from *A. thaliana* by genome rearrangement as a result of polyploidization. *A. thaliana* has been widely used as a model plant to study the genetics of *Brassica* species, as this plant genomic sequence is well established and known; it also has the advantage of small genome (Qiu et al., 2009). This is in agreement with a study in which a "conservative algorithm" was developed to identify co-linear loci between the genomes. The algorithm identified 34 significant *A. thaliana* regions that are co-linear with >28% of the *B. oleracea* genetic map (Lukens et al., 2003). Several comparative genetic analyses showed some gene homologues between *A. thaliana* and *Brassica* with similar structure and functions (Babula et al., 2006; Mun et al., 2009; Qiu et al., 2009). Gene mapping and sequencing of these two genera show incomplete conservation for the content of the genes and co-linearity between them which is due to chromosomal rearrangement (Gao et al., 2004).

Several evolutionary scenarios have been suggested to be the main cause of the observed diversity in plant metabolites among and between species (Jones and Firn, 1991; Kliebenstein, 2004). It has been suggested that glucosinolate biosynthesis was a result of a well known detoxification evolutionary mechanism in plants expressing cyanogenic glucosides (Wittstock and Halkier, 2002). *Arabidopsis thaliana* (*A. thaliana*) expresses a wide range of secondary metabolites, which can be classified into four major classes: phenylpropanoids, glucosinolates, terpenoids and phytoalexins. Glucosinolates are the largest group among these secondary metabolites in *Arabidopsis* (Kliebenstein, 2004). The three main classes of glucosinolates are well studied in *Arabidopsis* with very wide diversity in their profile and distribution, making it a suitable model for QTL mapping in other related

4

species in order to identify the genes involved in glucosinolate biosynthesis (Kliebenstein, 2009; Wittstock and Halkier, 2002).

1.3. **Glucosinolate biosynthetic pathways**

Different biosynthetic pathways resulted in variations in glucosinolate content, both between, and within the same species, leading to a hypothesis that this content is subject to both genetic and environmental control (Li and Quiros, 2003; Mithen, 2001; Windsor et al., 2005).

Figure 2 shows the general biosynthetic pathways of glucosinolates. Aliphatic glucosinolates are derived from methionine, indole glucosinolates are derived from tryptophan and aromatic glucosinolates are derived from tyrosine or phenylalanine (Schonhof et al., 2004). In addition, a substrate-enzyme dependent route evolved in the core structure formation of the three major classes of glucosinolates, were classified into two main groups depending on the stage of the biosynthesis they control. The first group; cytochromes P450 belonging to the *CYP79* family are responsible for catalyzing the conversion of amino acids to aldoximes, which will be converted into the corresponding *aci*-nitro compound by the aid of the second group of enzymes *CYP83* (Graser et al., 2001; Halkier and Gershenzon, 2006; Mewis et al., 2006). This is followed by the formation of a thiohydroximic acid by a *C-S* lyase, after which desulfoglucosinolate formation is catalyzed by *S*-glucosyl transferase, and finally the formation of glucosinolates by sulfotransferase. The last three enzymes involved in the core structure formation are common for all classes of glucosinolates (Windsor et al., 2005; Zang et al., 2009). The last phase in glucosinolate biosynthesis are the side-chain modifications that involve oxidation,

hydroxylation and methoxylation, which are under genetic and environmental control (Mithen, 2001).

In the biosynthetic pathway of aliphatic glucosinolates, methionine can undergo several elongation cycles for the addition of one methylene group at a time before it can enter the pathway for the formation of the glucosinolate core structure (Textor et al., 2007). Fine mapping of *Gls-elong* loci, on chromosome 5 in *A. thaliana* (de Quiros et al., 2000), enabled the identification of *MAM1, MAM2* and *MAM-Like (MAM-L)* genes coding methylthioalkylmalate synthases (MAMS) that belong to the enzyme class (EC 2.3.3.-) (Benderoth et al., 2009). This is involved in the elongation of methionine in aliphatic glucosinolates by catalyzing the condensation reaction of acetyl coenzyme A with ω-methylthio-2-oxoalkanoic acids to give 2-(ω-methylthioalkyl) malate intermediates (Halkier and Gershenzon, 2006). *MAM1* and *MAM2* encode for 4 and 3 carbon side-chain glucosinolates respectively, while *MAM-L* encodes mainly for 5, 6, 7 and 8 carbon side-chain glucosinolates as well as all other chain lengths (Keurentjes et al., 2006; Textor et al., 2007). In addition, *MAM-L* was found to catalyse the production of isopropyl malate (*IPM*) in the leucine biosynthesis pathway (Halkier and Gershenzon, 2006; Keurentjes et al., 2006; Textor et al., 2007). In a study investigating enzymes involved in the elongation step in the secondary metabolism of aliphatic glucosinolates and the analogous primary metabolism of the amino acid leucine (Field et al., 2004), it was revealed that both metabolic pathways require enzymes to catalyze the condensation of acetyl coenzyme A and 2-oxo acids. These are MAMS and isopropyl malate synthase (IPMS; EC 2.3.3.13), encoded by the *MAM* gene family. MAMS showed high homology to IPMS revealing an evolutionary link with glucosinolates (Halkier and Gershenzon, 2006).

1: Side chain elongation

2: Core structure formation

3: Side chain modification
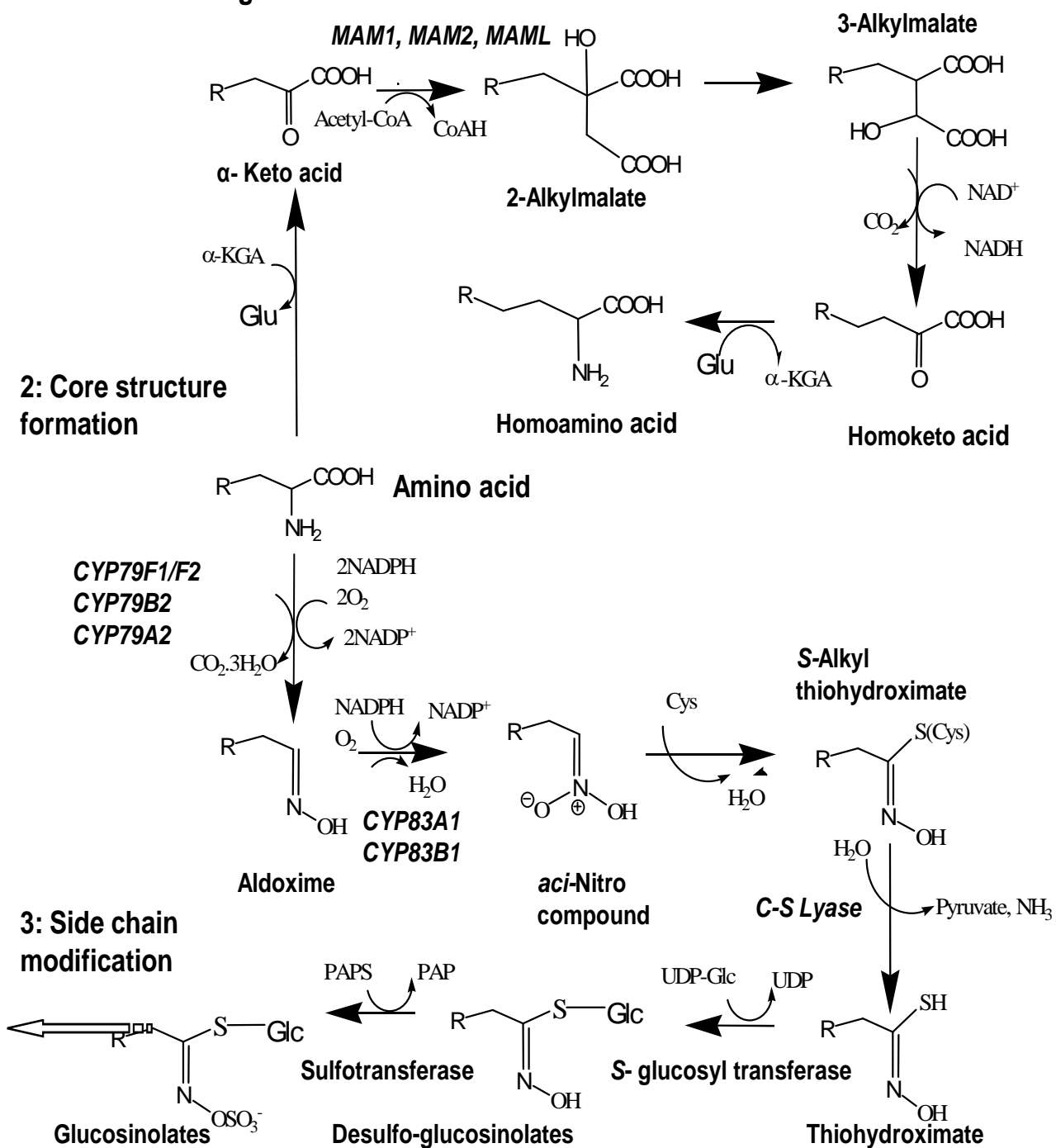
Figure 2 Outline of glucosinolate biosynthesis: 1. Prior to entering this pathway, amino acids can undergo one or more rounds of chain elongation. 2. The conversion of amino acids to the basic glucosinolate skeleton with genes encoding enzymes for each step. 3. Additional side chain modifications can occur following the pathway. R indicates the variable side chain (Halkier and Gershenzon, 2006)

Among the *MAM* genes identified in the model species *Arabidopsis* are *MAML-3* and *MAML-4* found on chromosome 1, *MAML-3* has IPMS activity and showed an impact on the biosynthesis of leucine, while *MAML-4* affected soluble amino acid content (Field et al., 2004).

The long chain aliphatic glucosinolates produced via pathways independent of the short chain glucosinolates, as substrate-enzyme dependent routes were investigated (Wittstock and Halkier, 2002). Elongated methionine was oxidized to aldoximes by *CYP79F1* and *CYP79F2* for the biosynthesis of short and long glucosinolates respectively, aliphatic aldoximes were converted to form the *aci*-nitro compound via *CYP83A1* (Chen et al., 2003).

As in aliphatic glucosinolates, phenylalanine may undergo a chain elongation step before the core structure formation of aromatic glucosinolates. *CYP79A2* and *CYP79B3* control aldoximes derived from phenylalanine and tryptophan for the synthesis of aromatic and indolic glucosinolates, respectively. *CYP83B1* catalyzes the conversion of aromatic and indolic aldoximes into the corresponding glucosinolates (Bak et al., 2001; Halkier and Gershenzon, 2006; Hansen et al., 2001; Hull et al., 2000; Windsor et al., 2005).

Common features in the biosynthetic pathway were found between glucosinolates and the other classes of secondary plant metabolites studied in *Brassica* species; such as cyanogenic glucosides (glycosides); briefly, at the stage of core structure formation, they are both derived from similar amino acids and have oximes as intermediates (Mithen, 2001). Studies revealed that in both pathways, amino acids were converted to aldoximes with the aid of *CYP79* enzymes. The main difference was in the type of enzymes that catalyzed the metabolism of aldoximes; where sequence homology studies showed high similarity between *CYP7IE1* and *CYP83*

that controlled the synthesis of cyanogenic glucosides and glucosinolates respectively (Halkier and Gershenzon, 2006)

1.4. **Glucosinolates and their breakdown products**

Myrosinase (thioglucoside glucohydrolase enzyme; EC 3.2.1.147) has the ability to catalyze the hydrolysis of glycosides, including glucosinolates (Bor et al., 2009). Myrosinase is segregated from glucosinolates in intact plants by cell organelles (Bennett et al., 2006). When this compartmentalization is lost by physical damage to plant tissue during, for example, freezing and thawing, chopping, and chewing, myrosinase catalyzed metabolism of glucosinolates occurs (Song et al., 2005), catalysing their conversion to the corresponding aglycone which then decomposes to isothiocyanates, thiocyanates, or nitriles (Figure 3) (Bennett et al., 2004; Grubb and Abel, 2006). Thus, the presumed health benefits of consuming food containing glucosinolates are achieved by the ability of the alimentary tract to produce different anticancer hydrolysis products in the digestive system from *Brassica* crops using myrosinase enzyme in the digestive tract and in plants. Moreover, it depends on the level of myrosinase activity in the human diet, as this will ensure more complete hydrolysis of glucosinolates (Cieslik et al., 2007). The types of products after the myrosinase hydrolysis are dependent on the substrate type, pH, and availability of ferrous ions and epithiospecific proteins (ESP). In the presence of glucosinoltes with terminal double bond in thir side, chain along with ESP may result in the production of epithioisothiocyanates. Isothiocyanates are the major product at physiological pH (Bones and Rossiter, 2006).
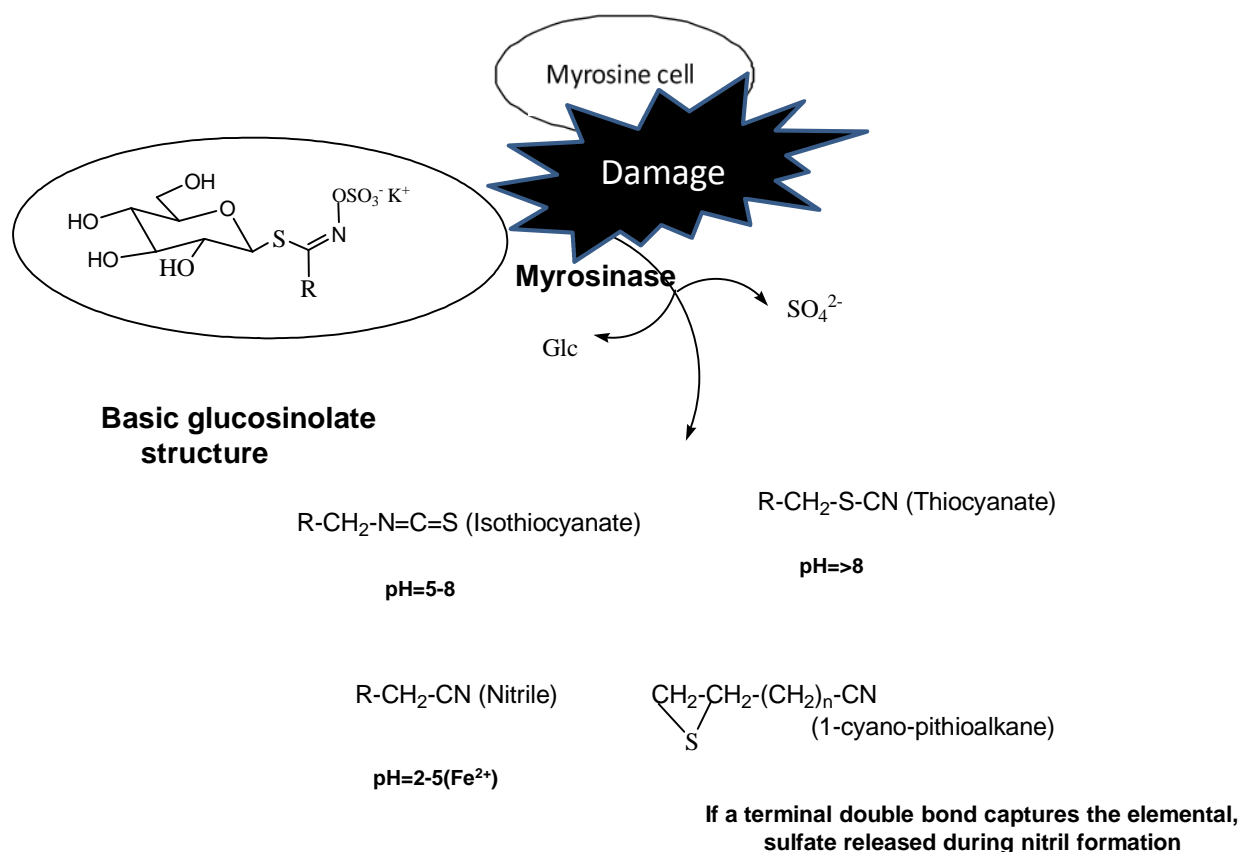
Figure 3 Basic glucosinolate structure and common myrosinase hydrolysis products, shown at different pH values. Based on (Bennett et al., 2004; Grubb and Abel, 2006)

This myrosinase-glucosinolate system may also be considered a part of the plant defence system against herbivores and pathogens, with at least six different types of these enzymes found in *A. thaliana.* By comparing their catalytic properties and functions towards sinigrin hydrolysis and responses to different ascorbic acid concentrations and other conditions, up to 22-fold differences in activity have been found (Andersson et al., 2009). In addition, myrosinase activity was very sensitive to ascorbic acid concentration (with maximum activity in the range of 0.7-1 mM), pH (4-7), temperature, and high salt concentrations. This indicates different functions of

similar enzymes *in planta*, in order to ensure plant fitness in different environments. In turn, this helps our understanding of the evolutionary system among plants and may be useful for plant breeders.

Paying more attention to those factors affecting myrosinase activity, such as climate, that have a pronounced effect on myrosinase activity may bring benefits (Charron et al., 2005). The effect of temperature and photosynthetic photon flux (PPF) on myrosinase activity has been studied; the results show a positive and a negative linear relationship respectively. The expected effects of both temperature and PPF on plant nutrient uptake, which affects myrosinase activity, explained this.

1.5. **Natural variation in glucosinolate profile in *Brassica*, under genetic and enviromental control**

Different *Brassica* crops express glucosinolate profiles with wide qualitative and quantitative variations, where *B. oleracea* has the most diverse profile compared to *B. rapa* and *B. napus*, (Antonious et al., 2009; Bellostas et al., 2007b; Brown et al., 2003; Matthaus and Luftmann, 2000). These differences are mainly due to genetic polymorphisms at the loci determining side chain modifications (Traka and Mithen, 2009).

Significant genotype and environmental effects on the profile of glucosinolates in broccoli florets have previously been observed (Farnham et al., 2004). The interaction between both genotype and environment was pronounced in all cases except for glucoraphanin, for which the genotype effect extended more influence than environment. Previous studies referred to the genetic effect to control the biosynthesis of aliphatic glucosinolates, while variation in levels of indolic glucosinolates was primarily controlled by environmental effect and by interaction of

11

environment and genotype effect. Also it was clearly shown that aromatic and aliphatic glucosinolates are less sensitive than indolic glucosinolates to storage and processing conditions (Cartea and Velasco, 2008).

With regard to environmental effects, many factors interact to contribute to glucosinolate level; plant age, plant parts, temperature, day length, water stress, soil type and pest attack (Velasco et al., 2007). This is in agreement with earlier studies (Schonhof et al., 1999), wherein the results show that the daily mean sum of temperature and light affect the biosynthesis of enzymes contributing to glucosinolate biosynthesis; the sulfur and nitrogen supply and the density of plant spacing all affect glucosinolate biosynthesis in addition to mechanical stress during plant growth.

Table 1 summarizes the relative effects of different environmental factors on metabolite production; including glucosinolates, in the *Brassica* plants commonly used for human consumption (Jahangir et al., 2009). Each compound's production is affected differently from the others by a stress factor; also, the order and the timing of the subsequent stress factors may control the final profile of these metabolites.

Table 1 Summary of the changes in production of different metabolites, increased (+) or decreased (-) as a result of environmental stress factors that affecting the biosynthetic pathway and/ or their regulators in *Brassica* species. N.D: not detected (Jahangir et al., 2009)

Stress induced metabolite production	Herbivore/ pathogen	Jasmonate/ Salicylate/ Ethylene	Temperature/ Light/ UV	Metal/ Fertilizer
Aromatic glucosinolates	+/-	+	+/-	+/-
Aliphatic glucosinolates	+	+	+/-	+/-
Indole glucosinolates	+	+	+/-	+/-
Phenylpropanoids	+	+	+/-	+
Flavonoids	+	+	+/-	+/-
Steroids	+	+	+/-	+
Amino acids	+/-	N.D	+/-	+
Jasmonic acid	+	-	N.D	N.D
Salicylic acid	+	-	N.D	N.D
Sucrose, glucose	N.D	-	N.D	-
Carotenoids	N.D	N.D	+	N.D
Organic acids	N.D	N.D	N.D	+

1.6. **The role of glucosinolates in human diet and promoting health**

In general, plants as foods are well known to have a wide variety of nutrients that have an impact on human health. Some of them are essential as our body cannot synthesise them, for example, vitamin C. Others are specific factors which are sometimes lacking, for example, amino acids and some fatty acids.

Knowing the health benefits of phytochemicals as anticancer, antioxidant, antifungal agents, or as protective factors against other chronic diseases (e.g. diabetes and cardiovascular disorders) led to the development of the "*Five-a-day*" campaign in the UK (http://www.nhs.uk/livewell/5aday/pages/5adayhome.aspx/) to encourage people to eat more healthily. Five portions a day of fruits and vegetable are recommended. Unfortunately, the sometimes perceived unpleasant taste minimizes people's appetite

for certain vegetables and suggests the necessity of supplements or fortified food to meet the need of these essential compounds (Lai et al., 2008), although the benefit from vitamin supplements is much less than a fruit and vegetable rich diet in preventing diseases, such as cancer. A study by (Thompson et al., 2010) supported the role for antioxidant nutrients from vegetables, and some fruits, as protective factors against the risk of non-Hodgkin's lymphoma. The study has not shown an association with supplemental intake of antioxidant nutrients, suggesting that any association is likely to be mediated through foods. This has mechanistic implications (potential synergies between antioxidants and other anti-carcinogenic compounds in these foods) and suggests that prevention approaches will likely need to be targeted toward foods or food groups, and not individual nutrients, particularly taken as supplements.

Evidence from the literature (Abercrombie et al., 2005); has shown the promising healthy effects of glucosinolates are not only for prevention and treatment of cancer, but also for protection from heart disease(s) as well as other neurodegenerative and chronic diseases (Jin et al., 2009). Thus, the need to measure the number and size of dietary servings per week in order to compare efficiency of different crucifers is important.

Numerous studies have shown the health-promoting effects of *Brassicaceae*. Broccoli was found to be especially rich in vitamin C and secondary plant compounds, such as glucosinolates (Verhoeven et al., 1997). Other nutrients and phytochemicals have been found in cruciferous plants; including folate, fiber, carotenoids and chlorophyll all of which have a chemoprotective effect (Higdon et al., 2007). Results from a large study (Michaud et al., 1999) reported a significant correlation between cruciferous vegetable consumption and a reduction in bladder

cancer incidence. Other studies provide evidence that cruciferous vegetable consumption reduces the risk of cancers of the colon/rectum, prostate, breast, and lung, as well as non-Hodgkin's lymphoma. In a recent study carried by (Kusznierewicz et al., 2008), the results showed a strong statistical correlation between total glucosinolate level and antioxidant activity in white cabbage. These findings were as a result of the hight content of polyphenols and flavonoids which are implicated in the protective effects against reactive oxygen species (postulated to be the main cause of many ageing processes and other diseases). This was found to be interrelated with the effect of growing conditions, climate and geographic origin on alteration of bioactive compounds in these plants.

Wet thermal treatments of cultivars expressing health beneficial phytochemicals, were studied in different *Brassica* varieties, especially in broccoli and cauliflowers. The effect of boiling, blanching and steaming on the levels of glucosinolates, phenols, anthocyanins and antioxidant parameters in cauliflowers, cultivated under the same environment conditions and processed immediately after harvest, have been studied (Volden et al., 2009). It was observed that the least influence on these phytochemicals was with steaming and the biggest influence when boiled. The loss of a large fraction of glucosinolates and antioxidants resulted in total glucosinolate levels being reduced by 55 and 42% for boiling and blanching respectively, compared to only 19% by steaming. Other compounds were similarly affected, but to a lesser effect than glucosinolates. Thus, in order to maximize the availability of glucosinolates in our diet, it has been highly recommended that the cooking procedure of these kinds of vegetables be modified, by replacing the boiling step, with steaming, microwaving or stir-fry, to minimize leaching of glucosinolates into the cooking water (Song and Thornalley, 2007). Early studies concentrated on the

toxic and anti-nutritional effects of glucosinolates; *e.g.* oxazolidine-2-thione derived from progoitrin, showed goitrogenic and growth retardation activity on animals, but no similar effect on humans (Cartea and Velasco, 2008). Thus, the need for improving plant metabolomics research through different analytical and genetic tools, is necessary; to ensure relative composition of healthy metabolites in relation to toxic or sensorial unacceptable compounds are optimal and under control (Hounsome et al., 2008).

1.7. **Potential of glucosinolates in prevention and therapy of cancer**

In recent years, several epidemiological studies have suggested that isothiocyanates resulting from the hydrolysis of alkyl glucosinolates found in cruciferous vegetables may play a chemoprotective role in the human diet by reducing the risk of cancer (Hecht, 2000).

A review of anticancer effects of glucosinolates highlights several points where studies can disagree, in particular, the spectrum of anticancer activities. Epidemiological evidences suggested "possible" health benefits rather than a particular activity, which is limited to lung and gastric cancer and not proven to be for all types of cancers by the uptake of a specific subtype of vegetables with critical type and amount of phytochemicals. In addition, they emphasized the role in humans of polymorphism for the genes responsible for glutathione *S*-transferase (GST) enzymes that are responsible for detoxification process for carcinogenic factors, which in turn reduces the risk of cancer and ageing progresses. Different individuals express these genes differently, thus altering their susceptibility for cancer by affecting their responses to cancer "chemoprotective" food (Kim and Park, 2009).

Investigations showed that the anti-cancer effect is not the same with all glucosinolates and their catabolic products, but is dependent on the nature of the side chain of the parent glucosinolate (Schonhof et al., 2004). Studies to compare the health promoting effects of individual products derived from glucosinolate myrosinase hydrolysis; the most important of which are sulforaphane (SF) and indole-3-carbinol (I3C) (Borowski et al., 2008), versus the effect of a whole plant extract containing their precursor glucosinolates, glucoraphanin and glucobrassicin respectively, are the focus of clinical research aiming to find potential cancer preventing and treatment compounds. The pathways triggered by SF and I3C are different, and thus studying the interaction between SF and I3C in an extract, on their effect on any cancer type is still an area of interest for researchers, as a possible synergetic activity for their anti-inflammatory effect is expected (Jeffery and Araya, 2009). Different mechanisms of action were proposed for the chemoprotective catabolic product I3C present in broccoli, cabbage, Brussels sprouts and cauliflowers, which has been proved to be successful against respiratory papilloma, and both breast and cervical cancer through *inhibition of transcription of estrogen-responsive genes* stimulated by 17ß-estradiol, as well as by its condensation product, 3,3'-diindolymethane (DIM), which is produced under acidic pH. These were found to have not only a preventive effect, but also therapeutic treatment for prostate cancer. In an *in vivo* study, mice inoculated with prostate cancer cells, and injected intraperitoneally with I3C (20 mg/ kg body weight), three times a week for 14 days before and after transplantation of prostate cancer cells, showed inhibition of cell growth (78% decrease in tumour volume) by *induction of apoptosis* and *inhibition of cell proliferation* (Souli et al., 2008). The ability of I3C to *induce cytochrome enzymes responsible for Phase II* detoxification, and unfortunately, also as inducer

17

for *Phase I* enzymes, which are known to be responsible for activation of carcinogenesis were also studied (Bellostas et al., 2007a).

SF the isothiocyanate catabolic product derived from sulfur containing glucosinolates such as glucoraphanin present in broccoli sprouts, broccoli, cabbage and Brussels sprouts, was found to be a powerful *inducer of phase II* and *an inhibitor of phase I* cellular enzymes. In addition, it has *antioxidant activity* and the ability to *induce cell apoptosis* as well as its *anti-inflammatory* activity, and *antibacterial* activity against *Helicobacter pylori*. Other hydrolysis products derived from different aliphatic glucosinolates, including iberin from glucoiberin, erucin from glucoerucin, crambene from epi-progoitrin and allyl isothiocyanate from sinigrin, were all found to have anticancer activity similar to SF, and strongly correlated to their R side-chain structure (Cartea and Velasco, 2008; Higdon et al., 2007).

Isothiocyanates derived from aromatic and aliphatic glucosinolates were found to be different in their anti-cancer activity, which is related to their side-chain structure. Secondary metabolites derived from aromatic amino acids such as gluconasturtiin, the precursor of phenethyl isothiocyanate present in watercress, has demonstrated activity against lung, leukemia, colon, prostate, liver and esophageal cancer (Hecht, 2000; Kim and Park, 2009). Benzyl isothiocyanate derived from glucotropaeolin found in cabbage, garden cress and Indian cress, has been studied for its chemoprotective activity. They showed the ability to *induce cancer cell apoptosis* as well as their ability to *induce phase II cellular enzymes* or to *inhibit phase I cellular* enzymes (Cartea and Velasco, 2008; Higdon et al., 2007; Pappa et al., 2006).

The effects of cooked and autolysed Brussels sprouts extracts, which are rich in sinigrin were investigated and the results revealed an inhibition activity for DNA strand cleavage in human lymphocytes exposed to 100 µM H_2O_2 for 5 min, possibly

through *scavenging oxygen radicals*, and to a lesser extent through *induction of phase II* enzymes. The results showed that both cooked and autolysed Brussels sprouts extracts had almost the same inhibition activity (38%) at concentrations of 10 µg/ mL and 5 µg/ mL respectively. The hydrolysis product, the isothiocyanate derived from sinigrin was at higher concentrations in the autolysed plant extracts (Zhu and Loft, 2001).

Comparison of the *cell growth inhibition* activity of SF, phenethyl isothiocyanate, I3C and DIM on human colon cancer cell lines has been studied (Pappa et al., 2006). The half-maximal inhibitory concentration (IC_{50}) in cell lines were: 15 mµM for SF and 10 mµM for phenethyl-isothiocyanate, I3C and DIM after 24 and 48h. The study found isothiocyanate to be cytotoxic, whereas indoles acted in a cytostatic manner.

1.8. Postulated defence role of glucosinolates as a secondary metabolite: mediating the interaction between plants and herbivore

Brassicaceae defence systems contain several chemical compounds, including protease inhibitors, saponins, anthocyanins and glucosinolates. The latter group is the most well known and studied of these compounds. Bearing in mind the postulated defence role of glucosinolates as secondary metabolites, a review (Grubb and Abel, 2006), discusses the variations in systemic distribution of glucosinolate in plants, finding the highest level in reproductive tissue (seeds, siliques, flowers, and developing inflorescence), followed by leaves, roots, and fully expanded leaves.

It is known that some cabbage aphids feed on cruciferous plants and are able to store glucosinolates in their intact form which they can later use them for defence against any attack which cause tissue damage by releasing their own myrosinase enzyme

which results in the production of hydrolysis products such as isothiocyanates (Pratt et al., 2008).

Several studies were performed in an attempt to correlate the plant's glucosinolate profile, with the damage caused by herbivores (Hopkins et al., 2009). Three wild *Brassica* populations, which grow naturally along the Atlantic coastline of the UK and France and one cultivar of *B. oleracea*, grown in a tempreture controlled greenhouse and has lower levels of glucosinolates in leaf tissue than plants of the wild populations, were compared regarding the performance of a specialist *Pieris rapae* (*P.rapae*) and a generalist *Mamestra brassicae* (*M. brassicae*) insect herbivore, and their endoparasitoids (Gols et al., 2008). The results showed that the development of specialist insects correlated to high levels of indolic glucosinolates (neoglucobrassicin and glucobrassicin) which cause a prolonged development time in *P. rapae*, where the level of aliphatic glucosinolates was not changed. Populations expressing higher levels of aliphatic glucosinolates (gluconapin and sinigrin) caused reduced survival against the generalist insects.

Although the wild population in this study was grown over a very short distance (within 15 km), they showed variations in glucosinolate profile caused by "local biotic and abiotic traits". In turn, this caused natural selection pressures, which explained the genetic differences between populations grown in close proximity, expressing different secondary metabolite levels.

As different synthetic pathways are involved in the production of aliphatic and indole glucosinolates, this is selected largely by environmental factors. Depending on the local insect population, glucosinolate composition in plants will be largely dependent on different resistance of these insects to glucosinolates, as the evolutionary changes are correlated to the ratio of specific to generalist insects coexisting in the same area.

20

This can affect the selection of not only genes regulating glucosinolate biosynthesis, but also genes regulating the glucosinolate hydrolysis profile (Kliebenstein et al., 2005; Windsor et al., 2005).

The effect of the level of specific biosynthetically related glucosinolates, on the associated herbivore community composition in the specialist *P. Rapae,* and a generalist *M. Brassicae,* has been studied. This was conducted on *B. oleracea* cultivars, in the field and in the laboratory in order to eliminate the possibility of competition between herbivores or differences in predation rates. Investigating the oviposition performance of both of them, it was observed that variations in the short side chain alkenyl glucosinolates affect the composition of the herbivore community to a significant extent, while higher concentrations of long side chain glucosinolates resulted in higher biodiversity, regardless of the degree of specialization of the herbivore. This indicates that insect biodiversity is not affected by plant biodiversity alone, but also by intra-species variations in plant secondary metabolites (Poelman et al., 2009).

This has shown a very big influence on ecosystem biodiversity by qualitative and quantitative phytochemical variations of the plant phenotypes, and by their degradation product profile, a co-ordered system under environmental control (Hopkins et al., 2009; Newton et al., 2009). This indicates that not only the level of individual metabolite is critical but also that the total and relative concentration is of major importance.

1.9. Different analytical methods for intact and desulfated glucosinolate analysis

Glucosinolate profiles in plants from different origins have been studied qualitatively and quantitatively for health benefits, agricultural, economic and ecological purposes. Two main approaches have been used.

Direct measurements of intact glucosinolates, determining the individual and/ or total concentrations of glucosinolates. Intact glucosinolates have been purified from plant material by reverse phase high performance liquid chromatography *(HPLC)* and identified by nuclear magnetic resonance (*NMR)* and by mass spectrometry *(MS*) methods. These methods reported a high yield (>90%) for individual glucosinolates quantified using the purification method described in a study by (Song et al., 2006).

The indirect measurement of glucosinolate derivatives produced by enzymatic hydrolysis (myrosinase and sulfatase enzyme) in order to measure the level of the break down products. Myrosinase enzymes are used to hydrolyze glucosinolates and the enzymatically released glucose is used for quantitative determination of total glucosinolate concentration (Antonious et al., 2009). Sulfate hydrolysis of glucosinolates, producing desulfated glucosinolates, has been used widely for qualitative and quantitative analysis of individual glucosinolates (Brown et al., 2003; EEC, 1990; Matthaus and Luftmann, 2000). Gas chromatography (*GC)* (Olsen and Sorensen, 1981)*, HPLC* (Leoni et al., 1998) and near infrared spectroscopy *(NIRS)* (Font et al., 2005), were used for separation of desulfated glucosinolates, followed by comparative analysis using tandem *MS* (*MS/MS), UV* absorbance, *NMR* and retention time (RT) with pure standards for confirmation and identification (Agerbirk et al., 2001; Bellostas et al., 2007a; Kiddle et al., 2001).

Quantification based on peak area of desulfated glucosinolates and comparison to peak area of internal or external standards can be accomplished by applying a relative response factor or calibration curves respectively, and a relative concentration for each individual compound has been obtained (Brown et al., 2003; Kim et al., 2009).

Desulfated glucosinolates have been analysed using *HPLC-APCI/MS-MS* methods (Griffiths et al., 2000). Desulfoglucosinolates were identified by application of fragmentation energy; the expected fragment weights were measured and used for qualitative analysis. The quantification method used was based on *HPLC-APCI/MS* and has the advantage of the ability to measure glucosinolates at low concentrations, which was lower than the minimum quantification levels in the *HPLC-UV* method. The disadvantage of this precise quantification method is the need to establish a calibration curve for each individual desulfated glucosinolate using pure standards, which are not commercially available. These would need to be prepared either by chemical synthesis of their analogues or by preparative chromatography from a crude plant extract using solid phase extraction (*SPE)* and *HPLC* purification and separation techniques (Rochfort et al., 2006).

The use of *GC* for separation and identification of trimethylsilyl and pertrimethylsilyl derivatives of intact and desulfated glucosinolates respectively has been studied (Hrncirik et al., 1998). The need for a derivatization step prior to analysis is not suitable for heat sensitive compounds. Moreover, the poor separation of glucosinolates with a methylsulfinyl side chain, made the *HPLC* method more popular as both methods showed the same accuracy and precision.

A strategy for the identification and quantification of intact glucosinolates from plant extract using combinations of different analytical techniques has been proposed.

23

Typically this comprises paired ion chromatography, ^1H *NMR, MS*, and chemical ionization. Myrosinase hydrolysis is used to confirm the identity of glucosinolates and the resulting isothiocyanates used for quantitative analysis. This strategy has the advantages of being simple, rapid and does not need special equipment or derivatization reactions (Prestera et al., 1996).

The indirect analysis approach used in this presented work utilizing the sulfatase desulfation reaction, has the disadvantage of variability of the analytical results, being subjected to enzymatic reaction conditions such as temperature, pH, time of reaction, but having the advantage of the convenience of the analytical methods used. Therefore, optimization of the laboratory working conditions during sample preparation and analysis were essential.

1. 10 Aims

To genetically map regions regulating the synthesis of individual glucosinolates in *B. Oleracea,* mapping populations with the aim of using such information to optimize the glucosinolates content in vegetable crops.

1.11 Objectives

- To develop experimental protocols suitable for analysis of glucosinolate profiles from large numbers of *Brassica* samples.

- To analyse samples from four parental plant lines

- To identify mapping populations with parent plants that is significantly different in their glucosinolate profiles.

- To use QTL mapping to investigate any linkage between genetic factors and expression of glucosinolates.

Extraction and analysis of glucosinolates

2.1 Introduction

Extraction and analysis of intact glucosinolates was performed using the procedure published by (Song et al., 2006), which was developed to isolate and separate glucosinolates with high purity from vegetable seeds. The first step of the chromatographic extraction process involved anion exchange column chromatography exploiting the affinity of the sulfate moiety of intact glucosinolates with Sephadex. Intact glucosinolates were eluted from the column in fractions according to their chemical differences using solvents with different polarities. A structural feature common to all glucosinolates is the presence of a strongly acidic sulfate group. Therefore, paired ion chromatography of glucosinolate mixtures was necessary (Prestera et al., 1996). Separation and identification of intact glucosinolates was performed using reverse phase high performance liquid chromatography *HPLC-MS* analysis. Maximum *UV* absorbance at 224 nm was used for detection of intact glucosinolates.

The standard procedures described in the official Journal of the European communities (EEC, 1990) were used for the extraction and desulfation of glucosinolates with some modifications. The principle of this method is shown in Figure 4.

```
┌─────────────────────────────────────────────────────────┐
│           Addition of an internal standard (IS)          │
│           to plant material at a fixed concentration     │
└─────────────────────────────────────────────────────────┘
                             │
┌─────────────────────────────────────────────────────────┐
│          Extraction of plant material in methanol        │
│          Load extract onto ion exchange column           │
└─────────────────────────────────────────────────────────┘
                             │
┌─────────────────────────────────────────────┐
│  Flushing the ion exchange column with sodium acetate     │
│  buffer at pH 4 to remove the hydrophilic components      │
└─────────────────────────────────────────────┘
                             │
┌─────────────────────────────────────────────────────────┐
│   Desulfation of glucosinolates bound to ion exchange column │
│   overnight at room temperature using sulfatase enzyme   │
└─────────────────────────────────────────────────────────┘
                             │
┌─────────────────────────────────────────────────────────┐
│           Elution of desulfated glucosinolates           │
│           from the ion exchange column with water        │
└─────────────────────────────────────────────────────────┘
                             │
┌─────────────────────────────────────────────────────────┐
│  HPLC separation and identification of desulfated glucosinolates │
│              according to their retention times          │
└─────────────────────────────────────────────────────────┘
                             │
┌─────────────────────────────────────────────────────────┐
│  Quantification of desulfoglucosinolates relative to IS (µmole/g dried │
│  plant material) using absolute peak area at 229 nm UV absorbance │
└─────────────────────────────────────────────────────────┘
```

Figure 4 Flow diagram of the method described in the (EEC, 1990) protocol for the determination of desulfated glucosinolates. Separation of desulfated glucosinolates by *HPLC* and identification by their RT. Quantification of desulfated glucosinolates in the extract was achieved relative to internal standard (IS), by applying a relative response factor to correct for differences in the *UV* absorbance between IS and the detected desulfoglucosinolates.

The standard procedure describes the extraction of glucosinolates using methanol, with sinigrin or glucotropaeolin added at a fixed concentration prior to extraction as an internal standard material (IS) which was used for quantification proposes. Purification of the extract was on an ion exchange column using acetate buffer before the enzymatic desulfation of glucosinolates into their desulfated form utilizing sulfatase enzyme from *Helix pomatia* type H1. The desulfoglucosinolates were eluted from the column with water and then separated using reversed phase *HPLC* with *UV* detection at 229 nm. Identification of eluted compounds was achieved by their order of elution in the chromatography. Quantification of the content of each desulfated glucosinolate, expressed in micromoles/ g of completely dry plant material, was calculated relative to IS. Relative response factors (RRF) were applied to correct for differences in *UV* absorbance between IS and desulfoglucosinolates.

2.1.1 Mechanism of desulfation reaction of glucosinolates

Desulfation of glucosinolates has been achieved using aryl sulfatase (aryl sulfohydrolases, EC 3.1.6.1) Type-1, which catalyses the hydrolysis of sulfate esters of a wide variety of aromatic compounds. The most important source of it is from the intestinal juice of the snail, *Helix pomatia*, which is commercially available and contains large amounts of aryl, steroid, and glucosinolate sulfatase activities, the latter catalyses the hydrolysis of sulfate group from glucosinolates (Leoni et al., 1998).

Figure 5 shows the desulfation reaction at equilibrium when both intact and desulfated glucosinolates are present in the same solution.
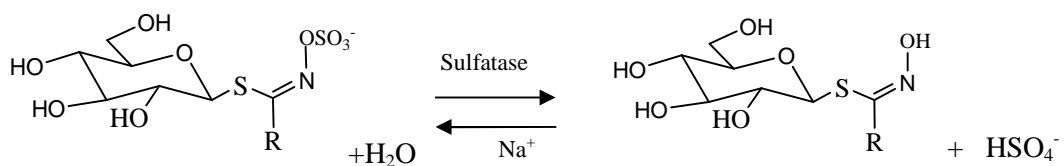
General structure of intact glucosinolates   General structure of desulfated glucosinolates

Figure 5 The desulfation reaction of glucosinolates by sulfatase enzyme from *Helix pomatia*. Desulfated glucosinolates are detected as sodium salts and/ or their protonated equivalent dependent on the pH of the reaction solution**.**

Several studies in the literature have shown the effect of different experimental parameters on the final recovery of desulfated glucosinolates. In a study on the effect of systematic variations in the analytical parameters including incubation time and sulfatase concentration on the desulfation reaction of different glucosinolates (Quinsac and Ribaillier, 1987), it was reported that the ratios between the peak areas of desulfated glucosinolates and the peak area of the IS varied according to the incubation time. Each glucosinolate reached equilibrium after different incubation times according to the nature of the glucosinolate; for example; the rate of desulfation for glucotropaeolin was approximately two fold that for sinigrin using diluted sulfatase solutions (Fiebig, 1991) and, therefore, the calculated amount of total desulfated glucosinolates using sinigrin as an internal standard would be lower than the content calculated via glucotropaeolin. Sinigrin and progoitrin are desulfated less rapidly than gluconapin and gluobrassicanapin (Quinsac and Ribaillier, 1987). In addition, differences in the time necessary to reach the equilibrium step in the desulfation reaction vary linearly with the dilution factor of the sulfatase solution (Quinsac and Ribaillier, 1987). Variations in the desulfation yield were more

28

pronounced between different glucosinolates when diluted sulfatase solutions were used (Fiebig, 1991).

In any study involving desulfated glucosinolates, attention must be paid to experimental design to ensure an efficient, reproducible, robust enzymatic desulfation is achieved.

Depending on the total content of glucosinolates present in the extract, (Wathelet et al., 2004), showed that the desulfation reaction of glucosinolates in a mixture is affected by feedback inhibition of the enzyme causing a slowing down of the reaction resulting in the incomplete desulfation of glucosinolates. In such a case, extending the desulfation time is necessary.

Two approaches have been used for the desulfation of glucosinolates. The reaction can be either performed in solution, or with the intact species bound to a chromatography matrix, as demonstrated by (Fiebig, 1991). They reported that the results obtained from enzymatic hydrolysis of glucosinolates on an ion-exchange column and in solution are different. In addition, the binding of desulfated glucosinolates to the column may vary between different desulfated glucosinolates resulting in varying yields. However, desulfation on an ion exchange resin is preferable to that performed in solution as the latter method lacks the ability to eliminate other hydrophilic materials in the extract, which co-elute with the desulfated glucosinolates and interfere with the analytical method. Using an ion exchange resin, a washing step can be utilised to eliminate unbound compounds prior to the application of the sulfatase enzyme.

2.1.2 **Quantification of desulfated glucosinolates**

Several RRF values for desulfated glucosinolates are available in the literature, calculated relative to particular glucosinolates and at different *UV* wavelengths (Brown et al., 2003; EEC, 1990; Fiebig, 1991; Wathelet et al., 2004). In this study, the values published in the standardised protocol (EEC, 1990) were used. These values were experimentally calculated by different laboratories relative to desulfosinigrin at 229 nm. This list has the advantage that it is the most comprehensive, as it has the RRF for all the glucosinolates previously detected in *Brassica* species. In addition, it has been used in the majority of the published studies.

The previous section described various factors affecting the desulfation reaction rate of glucosinolates in a plant extract, including the incubation time, sulfatase solution concentration, nature of glucosinolate content and the medium where the extraction is performed (in solution or in an ion exchange column). These are important parameters to evaluate when choosing a particular glucosinolate as an IS for use in the relative quantitative measurements of the glucosinolate content of plant extracts.

In the (EEC, 1990) method, only one IS was used prior to the extraction for the validation of the extraction, desulfation, separation and quantification methodology.

For this study, only glucotropaeolin and sinigrin were commercially available to use as pure standard reference glucosinolates. Therefore, it would be necessary to check for the presence of any of these available standards in plant extracts. A suitable "IS1" should be chosen based on the lack of endogenous glucosinolate in plant material. It is essential to experimentally determine the optimum concentration of IS1 in the extract in order to obtain quantitative measurements without exceeding the linear

range of the *UV* detector, whilst ensuring complete resolution of the IS1 peak from other components in the samples.

A second internal standard (IS2) would provide additional information that could be used to improve the reproducibility of the quantitative measurements. This would be added to each sample at a fixed concentration prior to sample injection onto the *HPLC* system. The peak area of IS2 can then be used to correct for the variability in the quantitative measurements due to variation in the injection volume. Ideally, this second "IS2" should elute early in the chromatogram thus being well resolved from the other interesting analyte peaks. Similarly, the concentration of IS2 would need to be experimentally determined.

2.1.3 **Characterisation of desulfated glucosinolates by means of *ESI-MS/MS* analysis**

The *HPLC-UV* method was used for separation of desulfated glucosinolates in the plant extract as described in the standard protocol (EEC, 1990). In this method potential desulfated glucosinolates were identified by comparing their RT in the *UV* chromatogram at 229 nm with well known glucosinolates previously detected in *Brassica* species. In fact, according to the chromatographic conditions being used; variations in the mobile phase gradient, composition, flow rate or the type of column, desulfated glucosinolates can elute with different RT and in a different order in the chromatogram (Wathelet et al., 2004).

Alternatively, desulfated glucosinolates have been identified by comparing their *UV* spectra to those of purified standards (Brown et al., 2003). This method used isocratic rather than gradient elution to avoid any possible drift in the *UV* detector base line, which can result in poor peak separation (Wathelet et al., 2004).

Therefore, coupling the chromatographic method to an analytical method which has the ability to discriminate individual desulfated glucosinolates (*e.g. MS, NMR*), was essential to ascertain reliable identification of desulfated glucosinolates in plant extracts (Kiddle et al., 2001).

The methodology used in this study was based on developing a *HPLC-UV/ESI-MS/MS* method for identification of desulfated glucosinolates in the plant extracts. Desulfated glucosinolates are characterised by *m/z* values presenting the protonated and sodium molecular ions, $[M+H]^+$ and $[M+Na]^+$, respectively (Zimmermann et al., 2007). Confirmation of the identity of desulfated glucosinolates was performed using an *MS/MS* fragmentation method, which produces a typical fergment ion for all desulfated glucosinolates with the general formula of $[M+H-C_6H_{10}O_5]^+$, resulting in the observation of fragment ion with 162.1 Da less mass than the precursor ion (Griffiths et al., 2000; Matthaus and Luftmann, 2000; Zimmermann et al., 2007).

2.1.4 **Aims**

In this study, the aim was to characterize the glucosinolate profiles of the double haploid (DH) lines from two *B. oleracea* reference segregating mapping populations (AG, NG) in order to map the Quantitative Trait Loci (QTL) of the glucosinolates within this population.

2.1.5 **Objectives**

- Establish an experimental protocol for the analysis of intact and desulfated glucosinolates from different plant samples

- Establish robust enzymatic desulfation of intact glucosinolates from each sample.

- Develop an *HPLC* method to resolve individual desulfated glucosinolates.

- Identify potential desulfated glucosinolates using mass spectrometry *(MS)* and confirm identities by tandem *MS (MS/MS)*.

- Identify suitable standards for use in the methodology to improve the quantitative measurements.

- Quantification of desulfated glucosinolates from the chromatogram based on *UV* absorbance at 229 nm and relative response factors.

2.2 Materials and methods

2.2.1 General material

Glucotropaeolin (benzyl glucosinolate) from Applichem and sinigrin (2-propenyl glucosinolate) from Sigma Aldrich are currently the only commercially available glucosinolates. Desulfated neoglucobrassicin was obtained from the Agricultural Research Council, (CRA-CIN) Via di Corticella, Italy.

Barium acetate Ba(OAc)$_2$ and lead (II) acetate trihydrate Pb(OAc)$_2$ •3H$_2$O were purchased from Sigma Aldrich, all were of analytical grade.

Trifluoroacetic acid (TFA) was purchased from Fluka. 0.1% Formic acid and 0.1% formic in acetonitrile were of *HPLC-MS* grade purchased from J.T. Baker. Deionised water (Millipore Q-POD ™ Milli-Q) was used in all experiments.

2.2.2 Plant material

Doubled haploid (DH) lines from two *B. oleracea* reference segregating mapping populations have been described previously by (Sebastian et al., 2000). The AG population was represented by 89 DH lines that had been derived by another culture from an F1 produced from a cross between Chinese kale; A12DH (var. *alboglabra*) and broccoli; GD33DH (var. *italica*). The NG population was represented by 69 DH lines derived from an F1 resulting from a cross between cauliflower; CA25 (var. *botrytis*) and Brussels sprouts Gower; AC498 (var. *gemmifera*). The plant materials used in this study were grown in a glasshouse under standardised conditions from April until September of 2008, as descried below.

The seeds were sown into (FP7) pot with M2 compost at a depth of 2 mm into a hole and covered lightly with compost. Two seeds can be sown in each pot to allow for possible non-germination and the extra one thinned out once growth has been established. Each pot was lablled with the plant accession number. These pots were randomly placed in glasshouse number 16 at 18 ˚C with additional light for day length. Watering only was used for the first 3-4 weeks, and then nutrients were added into the water for three times per week regime, by Warwick HRI horticulture services, and then moved into a larger pot when the plantlet is in suaitable size. Plants were caned and tied to ensure healthy plant growth and to maximise leaf growth.

Young fully expanded healthy leaves were collected at the bud initiation stage and flash frozen in liquid nitrogen, to inactivate endogenous myrosinase, and then stored at -80 ˚C until required (Sanyo ultra low temperature freezer model: MDF-792). Frozen leaves were dried using a freeze drier (Edwards super modulyo model: Prirania 501) and processed for 3-4 days until completely lyophilized, then stored at -20 ˚C.

2.3 Experiments

A robust, quantitatively, statistically valid protocol was developed in this study for the extraction and desulfation of glucosinolates from the analysed plant materials. A detailed description for each step involved in this protocol is described in the following sections.

2.3.1 **Extraction of glucosinolates from plant material**

The method described by (Brown et al., 2003) was used for the extraction of glucosinolates from *B. oleracea* leaves. 0.3 g of plant material was extracted using 20 mL of boiling water containing 2 µmol IS1, 30 µmol $Pb(OAc)_2$ and 30 µmol $Ba(OAc)_2$ to precipitate proteins and free sulfate ions. After 10 min boiling using a hotplate with water reflux (Electromantle CAT NO EM 0500/CE) with 30 min agitation and sonication (Kerry), the samples were cooled to 4 ˚C and centrifuged at 3000 *g* for 40 min (Heraes sepatech centrifuge Omnifuge 2.0RS model: SD00400/ YR 2000).

2.3.2 **Preparation of DEAE-Sephadex A-25 column**

DEAE Sephadex A-25 (GE Healthcare) is a weak anion exchanger with a diethylaminoethyl group. The ion exchange step was used to bind intact glucosinolates, remove contaminating hydrophilic and unbound material that could interfere with any downstream quantification and *MS*-based measurements. The Sephadex column was prepared by suspending 2 g of the powder into an excess of sodium acetate buffer pH 5.5. Complete swelling takes 1-2 days at room temperature. The slurry was washed with excess deionised water, poured into the column and washed with 200 mL of deionised water followed by 200 mL of aqueous methanol before application of the plant extract. This process was found to be essential to remove any polymers which can appear as peaks in the mass spectra from the eluate.

2.3.3 **Purification of intact glucosinolates**

Initially the method described by (Song et al., 2006) was used for the collection of intact glucosinolates in fractions according to their affinity to the ion exchange

column. Extracts of plant material were loaded onto the Sephadex column; distilled water (500 mL) and 50% methanol (200 mL) were used for washing the column. Intact glucosinolates were eluted sequentially with 0.1 mol/ L potassium nitrate (200 mL), 0.1 mol/ L potassium sulfate (200 mL) and 0.2 mol/ L potassium sulfate (200 mL). Each eluted fraction of potassium nitrate and potassium sulfate was collected, evaporated to dryness under vacuum at 50 °C and resuspended in 2 mL *HPLC* water and spin filtered prior to *HPLC* analysis.

2.3.4 **Desulfation of intact glucosinolates**

Plant material was extracted and loaded onto the Sephadex column. The column bound intact glucosinolates were flushed with 67% aqueous methanol (200 mL), and then they were capped and treated overnight with a solution of aryl sulfatase as described by (Brown et al., 2003) to convert the glucosinolates to their desulfated derivatives.

2.3.5 **Preparation of aryl sulfatase solution**

Sulfatase was commercially available and was purchased from Sigma Aldrich. The batch had an activity of 22400 Units/ g solid, at the optimum pH (5-7) at 37 °C.

Sulfatase for glucosinolate desulfation was prepared as described by (Graser et al., 2001). Aryl sulfatase 66 mg was dissolved in deionised water (6 mL) and mixed with absolute ethanol (6 mL). After centrifugation at 5000 *g* for 20 min at 4 °C (DuPont Sorval lRC- 5B, centrifuge model: SD: 00611, YR 2000), the supernatant was mixed with additional absolute ethanol (18 mL) and centrifuged at 4000 *g* for 15 min, at 4 °C. The pellet from this second centrifugation was dissolved in deionised water (150

mL) and divided into 1 mL aliquots and stored at −20 °C. This sulfatase stock solution had a concentration of 10 U/ mL.

2.3.6 **Optimisation of the enzymatic desulfation reaction**

It was essential to establish a robust and reproducible methodology for the desulfation of the glucosinolates bound to the ion exchange column. This was achieved using a range of sulfatase enzyme concentrations to try to ensure as complete as possible desulfation of the glucosinolates present across all the samples.

In the method described by (Brown et al., 2003) after washing the column with 67% aqueous methanol (200 mL), sulfatase solution was added at the top of the column and kept overnight at room temperature. In order to improve the reproducibility of the enzymatic desulfation reaction, several steps within the sample preparation protocol were optimized. After the methanol wash, the Sephadex matrix including the bound intact glucosinolates were equilibrated in sodium acetate buffer pH 5.5 (200 mL). The Sephadex material was then transferred into 100 mL glass bottles and treated with varying concentrations of aryl sulfatase solution at a fixed volume of 1 mL. The total volume of the reaction was adjusted to 40 mL with the acetate buffer. The Sephadex, plant extract and buffered enzyme were incubated in a shaking incubator at 37 °C at 80 rpm (New Brunswick ScientificModel: Innova ® 44) to ensure optimum contact between the enzyme and the bound glucosinolates. After 18 hrs, the Sephadex material containing the desulfated glucosinolates was transferred into an empty column and eluted with 60% aqueous methanol (200 mL). The eluent was evaporated to dryness under vacuum at 70 ˚C (BÜCHI Laborortechnik B-480 with rotavapor R-114), resuspended in 2 mL of 0.1% formic acid and spin filtered

using 0.2 µm cellulose acetate micro spin filter tubes (Grace) at 1000 rpm for 30 seconds (Helena Biosciences Centurion Scientific LTD model: K80) prior to *HPLC-MS/MS* analysis.

In this study, experiments were performed in order to show the effect of sulfatase solution concentration on the production of desulfated glucosinolates, in order to establish the optimised enzyme concentration for desulfation of glucosinolates across the plant materials being analysed. The methodology described above was employed using a range of sulfatase concentrations. The absolute amounts of enzyme for optimal desulfation reaction were determined based on a study for the production of desulfated glucosinolates by (Leoni et al., 1998). In the study, 0.5 g plant material and 3 µmol sinigrin (as IS) were extracted with water and treated with 0.25 units (U) of sulfatase enzyme. A batch of homogeneous plant material was generated and used for all of the enzyme concentrations; 2.5 g of homogeneous plant material was extracted without adding IS1 and divided into five equal volumes. Each portion was desulfated using a serial dilution of sulfatase enzyme solution, ultimately containing 0.25, 1.25, 2.5, 3.75 and 5 U in 1mL final volume prepared as described in section (2.3.5). The relative area (compared to IS2) of each detected desulfated glucosinolate peak in the *HPLC-UV* chromatogram at 229 nm was determined, in three (chromatographic) technical replicates.

2.3.7 Optimisation of concentration of IS1 used for quantitative measurements

In the method described by (Brown et al., 2003), 0.2 g plant material and 0.5 µmoles of sinigrin (as IS) were extracted with water and a relative response factor was used

to determine the concentrations of the desulfated glucosinolates in the samples relative to the IS. In our experiments, glucotropaeolin was used (as IS1) for relative quantitative measurements of glucosinolate content in all plant lines. An appropriate level of IS1 addition, needed to be established prior to the analysis of any plant extracts. Experiments were performed to determine the optimal ratio of plant material to IS1 concentration as follows:

- 0.5 g plant material was extracted with 0.5 µmoles of IS1 (1 g : 1 µmoles)

- 0.5 g plant material was extracted with 2 µmoles of IS1 (1 g : 4 µmoles)

- 0.3 g plant material was extracted with 2 µmoles of IS1 (1 g : 6.7 µmoles)

In order to ensure maximum desulfation of all glucosinolates (including the IS1) in the plant extract, these samples were desulfated using sulfatase solution containing 10 U of sulfatase enzymes in 1 mL, and the optimized desulfation reaction was used as described in section (2.3.6). These samples were then subjected to the *HPLC-MS/MS* analysis and analysed in triplicate.

2.3.8 **Optimisation of concentration of IS2 used for quantitative measurements**

Serial dilutions of intact sinigrin in aqueous solution were prepared at a final concentration of 1 mg/ mL, 2 mg/ mL and 6 mg/ mL. An aliquot containing 5 µL of each concentration was diluted 10-fold in water and analysed. A 25 µL of each of the above solutions was injected into the *HPLC* without using *MS* analysis. The optimum concentration of IS2 was determined initially by comparison of the absolute peak area/s of IS2 compared to the absolute peak areas typically obtained for the desulfated glucosinolates in the extracts analysed to date. A concentration was

selected that provided average peak areas similar to those of the analysts of interest. This concentration was further validated by addition to a plant extract and analysed in triplicate.

2.4 Analysis of glucosinolates

In this study, preliminary analyses for the detection of intact and desulfated glucosinolates in four different plant lines were performed on an Agilent 1100 series *HPLC* coupled with Bruker Daltonics *ESI-MS*. This instrument was not available for use later on, and therefore a Thermo Fisher scientific *HPLC-LTQ XL* mass spectrometer was used.

2.4.1 Preliminary analysis of glucosinolates in plant extract

The initial experiments to detect glucosinolates were performed on an Agilent 1100 series *HPLC* comprising of *UV* detector (diode array detector G1315B); binary pump (G1312A), auto sampler (G1313A), degasser (G1379A) and column oven (G1316A). A Lichrospher RP-C18, 250 x 4.60 mm, with a particle size 5 µm *(Phenomenex)* column was used held at a constant 25 °C, for the separation of intact and desulfated glucosinolates.

2.4.1.1 *HPLC* Separation of intact glucosinolates

The separation of intact glucosinolates was based on the method described by (Song et al., 2006) with modifications. Water (A) and methanol (B) gradient solvents containing 0.1% TFA were used to separate intact glucosinolates at a flow rate of 1mL/ min using Lichrospher RP-C18, 250 x 4.60 mm column, with a particle size of 5 µm. A linear gradient was performed to 50% B over 45 min as shown in Table 2. The effluent was monitored at 224 nm by diode array detection. Three technical

replicates each containing 50 µL of each sample were injected into the *HPLC* system.

Table 2 The *HPLC* gradient used to separate intact glucosinolates. Water (A) and methanol (B) gradient solvents containing 0.1% TFA at a flow rate of 1mL/ min using Lichrospher RP-C18

Time (min)	% B
0	0
5	0
35	15
45	50
50	50
55	0

2.4.1.2 *HPLC* separation of desulfated glucosinolates

The methodology used was based on that published by (Brown et al., 2003) with modifications. Water (A) and acetonitrile (B) gradient solvents containing 0.1% TFA were used to separate desulfoglucosinolates at a flow rate of 1mL/ min using Lichrospher RP-C18, 250 x 4.60 mm column, with a particle size of 5 µm. A linear gradient was performed to 93% B over 25 min as shown in Table 3, within this gradient desulfated glucosinolates eluted approximately between 12 to 32 min. The effluent was monitored at 229 nm by diode array detection. Three technical replicates using 50 µl of sample were performed for each extract.

Table 3 The *HPLC* gradient used to separate desulfated glucosinolates. Water (A) and acetonitrile (B) gradient solvents containing 0.1% TFA at a flow rate of 1mL/ min using Lichrospher RP-C18, 250 x 4.60 mm, with a particle size 5 µm column at Agilent 1100 series

Time (min)	% B
0	0
5	0
6	1.5
11	5
13	7
23	21
28	29
35	43
36	93
40	93
41	1.5
48	1.5

2.4.1.3 **Identification of glucosinolates by *ESI-MS***

Confirmation of the eluting glucosinolates was achieved using an online Bruker Daltonics *HCT* plus mass spectrometer. The *ESI-MS* (G2431A/G2431-60001) instrumental conditions were as follows:

- Capillary exit 121 V, HV capillary 4000 V.

- Drying gas nitrogen temperature 300 ℃, and drying gas flow 10 L/ min, nebulizer= 40 psi.

- Scan range of 50-2000 *m/z*, scan mode was ultra scan in a range of 50-3000 *m/z* at a speed of 26000 (*m/z) /* second. During the tuning procedures, the instrument was operated at a scan range of 50-3000 *m/z* and the following single charged positive ions *m/z* 118.2, 322.2, 822.0, 921.9, 1321.8, 1521.7 and 2121.7 were observed.

- Smart parameter settings were as follows: compound stability is 100% and trap drive is 100%. ICC smart target is 100000 with a maximum accumulation time at 20 msec.

The tuning mixture for the *ESI-MS* was purchased from Agilent Technology; its composition is described in Appendix B.

The outlet of the *diode array* detector on the *HPLC* system was connected directly to the *MS* source. Negative ion mode was used for detection of intact glucosinolates, whilst positive ion mode detection of desulfated glucosinolates was used.

2.4.2 **Optimisation of the experimental method used for separation of desulfated glucosinolates using *HPLC***

For the qualitative and quantitative analyses needed for desulfated glucosinolates a Thermo Fisher Scientific *HPLC-LTQ XL* mass spectromer was used comprising of

Accella autosampler, utilizing full loop injection (25 µL), with all samples housed at ambient room temperature. An *Accella LC* quaternary pumping system was used as a binary system and an *Accella* photodiode array detector (PDA) was used to collect spectral data from 200-600 nm using one channel at 229 nm for absorbance maxima for the desulfated glucosinolates.

The analytical column used for all analyses was a *Zorbax* Eclipse XDB-C18 4.6 × 150 mm 5 µm with in-line *Zorbrax* reliance analytical guard column (4 × 80 mm, 5 µm), from Agilent Technologies.

The outlet from the PDA of the *HPLC* was connected to the mass spectrometer via a T- piece union, which splits 30% of the *HPLC* eluate to waste. The first 5 min and the last 16 min of the chromatographic flow were diverted to waste to reduce contamination of the *MS* ion source.

2.4.3 Optimisation of the *mass spectrometer* experimental conditions for characterisation of desulfated glucosinolates

Identification of the eluting desulfated glucosinolates was achieved using an online Thermo Fisher Scientific electrospray ionisation mass spectrometer *w*ith linear ion trap mass analyser and utilizing the Xcalibur *LTQ* program Rev 2.5.0. The experimental conditions were as follows:

- The *ESI* nozzle was held at 5 kV and a temperature of 280 $^{\circ}$C.

- Nitrogen gas used to assist in the nebulisation of the spray and to aid desolvation with sheath and auxiliary gas flows of 10 psi and 5 arbitrary respectively.

- The instrument was operated in positive ion mode for the detection of desulfated glucosinolates using full scan mode (*MS* and *MS/MS* scans with a resolution of 1000).

The tuning solution was prepared by dissolving *N*-vanillylnonanamide (Sigma Aldrich) in 0.1% formic acid in acetonitrile at a concentration of 0.08 mg/ mL. In order to tune the lenses in the system, the automated process was performed and optimised on the ion observed at *m/*z 294, as the mass range of the expected desulfated glucosinolates was *m/*z 200 to 500.

Calibration of the instrumental conditions was performed using the in-built automatic calibration program every three months and was checked every day to ensure that the mass drift was no greater than ±0.2 Da. The *ESI* calibration solution composition is described in Appendix C. The following singly charged positive ions for caffeine *m/z* 195, MRFA *m/z* 524 and Ultramark 1621 *m/z* 1222, 1522, and 1822 were observed and used for the calibration. The calibration conditions were as follows:

- Calibration solution was infused at a flow rate of 5 µL/ min into the *ESI* source by the syringe pump.

- The *ESI* nozzle was held at 4.5 kV and a temperature of 275 °C.

- Sheath and auxiliary gas flows of 10 psi and 0 arbitrary respectively.

- The instrument was operated in full, normal scan mode. *MS* data were collected over *m/*z range 150-2000 with a maximum injection time of 200 msec.

2.4.4 Development of a reproducible *HPLC* method for optimal separation of desulfated glucosinolates

The method used to separate desulfated glucosinolates (described under section 2.4.1.2) using the Lichrospher RP-C18, 250 x 4.60 mm column, resulted in poor peak separation. Therefore, further improvement was based on the method described by (Matthaus and Luftmann, 2000) with modifications, as follows:

- A shorter *Zorbax* Eclipse XDB-C18 4.6 × 150 mm, 5 µm particle size reversed phase column was used for all subsequent analyses.

- The flow rate used was reduced to 0.25 mL/ min.

- The mobile phase was composed of water (solvent A) and acetonitrile (solvent B) linear gradient, the ion-pairing agent was replaced with 0.1% formic acid to increase the ionization efficiency of the compounds in the *ESI/ MS* source to resolve the desulfated glucosinolates. A linear gradient was performed to 95% B over 33 min as shown in Table 4, within this gradient desulfated glucosinolates eluted approximately between 9 and 29 min.

- The gradient length was increased by 8 min to ensure good peak separation for desulfated glucosinolates during the latter stages of the chromatography.

- The washing and equilibrating time for the column was extended for an extra 9 min, the flow rate used at this stage was increased to 0.4 mL/ min. This improved the reproducibility for the chromatographic separation of the next sample by decreasing the time required for each *HPLC* analysis.

- A full loop injection (25 µL) was used for three technical replicates.

Table 4 The *HPLC* gradient used to separate desulfated glucosinolates using water (A) and acetonitrile (B) solvents containing 0.1% formic acid and Zorbax Eclipse XDB-C18 column based on (Matthaus et al., 2000)

Time (min)	Flow rate (mL/ min)	% B
0	0.25	5
2	0.25	5
28	0.25	41
30	0.25	41
35	0.25	95
40	0.40	95
42	0.40	5
56	0.40	5

2.4.5 Development of *MS* and *MS/MS* method to confidently identify desulfated glucosinolates

To confidently identify desulfated glucosinolates in the plant extracts, an *ESI-MS/MS* method was used based on the method described by (Griffiths et al., 2000; Matthaus and Luftmann, 2000; Zimmermann et al., 2007) with modifications. The *MS/MS* fragmentation of the protonated adduct $[M+H]^+$ under collisional activation conditions is typical for all desulfated glucosinolates, and produces the neutral loss of a 162.1 ion corresponding to the sugar group (Matthaus and Luftmann, 2000). This method has the advantages of producing a fragment ion characteristic of all desulfated glucosinolates with other additional structure specific fragments, used for structure determination of glucosinolates with the same molecular weights.

Within the Xcalibur *MS* program Rev 2.0.7, a mass inclusion list was created, spanning the region 280-399, containing the expected *m/z* of 13 protonated desulfated precursors shown in Table **5** (described in Appendix A), including the *m/z* range of the *MS/MS* acquisitions for each precursor.

Table **5** The mass inclusion list for the expected 13 protonated desulfated glucosinolates precursor, with their *MS/MS* (*m/z*) ranges.

Desulfated glucosinolates	$[M+H]^+$ (*m/z*)	*MS/MS* *m/z* range
Sinigrin	280	65-290
Gluconapin	294	70-305
Glucobrassicanipin	308	70-320
Progoitrin	310	75-320
Glucotropaeolin (IS1)	330	80-340
Glucoerucin	342	80-355
Gluconasturtiin	344	80-355
Glucoraphanin	358	85-370
Glucobrassicin	369	90-380
Glucoalyssin	372	90-390
4-Hydroxyglucobrassicin	385	95-395
Glucohesperin	386	95-400
4-Methoxydroxyglucobrassicin/ Neouglucobrassicin	399	95-410

In the mass spectrometry method, three scan events were specified as follows:

1) The first event involved the collection of *MS* data over the *m/z* range 150-500

2) After five minutes, the instrument switched into a data dependent mode, performed a single *MS* scan, determined if one of the ions specified in the mass inclusion list was present and if so, acquired *MS/MS* data on the precursor ion observed for 30 msec using normalised collision energy of 35 with a 2 Da isolation width. If none of the ions were observed the instrument remained in *MS* mode until a specified ion was detected.

3) After 35 min, the instrument returned to *MS* acquisition mode until the run ceased at 56 min.

As the samples were in sodium acetate buffer, this resulted in a build up of salt at the inlet to the source region. This required frequent cleaning of the heated capillary and ion transfer tube to maintain high sensitivity.

49

For example, desulfoglucotropaeolin was characterised by comparing RT and *m/z* data with pure standard material (Figure 6). In the *UV* chromatogram, desulfoglocotropaeolin eluted at 20.83 min (Figure 6, A). In the *MS* spectrum, desulfoglucotropaeolin was observed with *m/z* 352.0 and 330.0, corresponding to the ions [M+Na] [+] and [M+H] [+] respectively (Figure 6, B). Fragmentation of the protonated adduct (*m/z* 330.0) of desulfoglucotropaeolin gave the expected fragment (*m/z* 168.0), which corresponded to the loss of the sugar group (Figure 6, C).
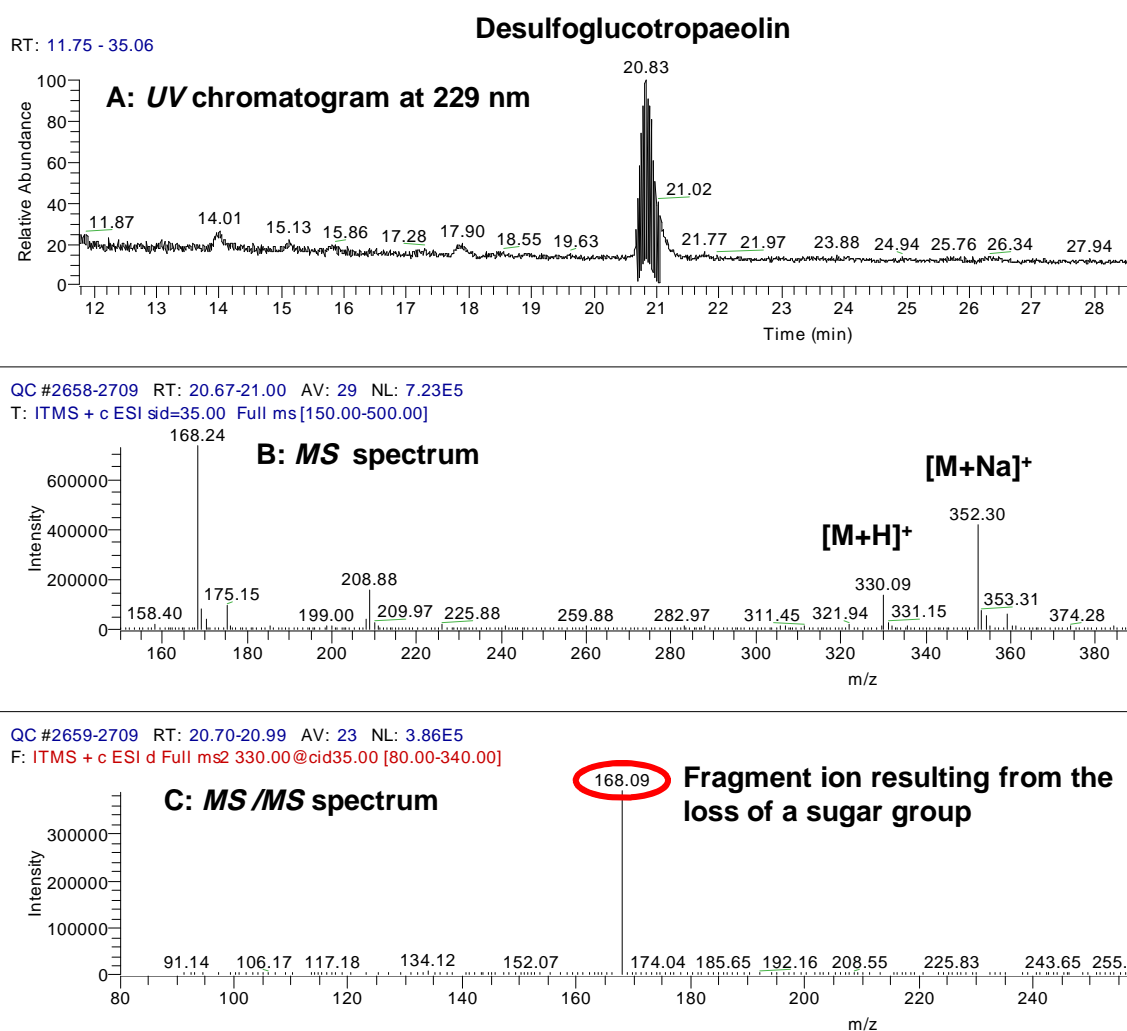
Figure 6 Characterisation of desulfoglucotropaeolin using A: RT at 20.83 min in the *UV* chromatogram, B: the [M+Na] [+] and [M+H] [+] adducts, with *m/z* 352.0 and 330.0 respectively in the *ESI-MS* spectrum, C: the expected fragment with *m/z* 168.0 observed in the *MS/MS* spectrum obtained after selecting the precursor ion with *m/z* 330.0

2.4.6 **Development of statistically valid method for quantifying desulfated glucosinolates in plant extracts relative to IS1**

A statistically valid method for quantifying desulfated glucosinolates from the chromatograms of the AGDH plant extracts and their parental lines relative to IS1 and based on *UV absorbance* at 229 nm utilizing RRF was established.

From the *UV* chromatogram of 0.3 g GDDH33 leaves extracted with 2 µmoles of IS1, desulfated and injected into the *HPLC* as described under section (2.4.4), the *absolute* peak area of desulfated glucosinolates were converted to *relative* peak area (based on IS2). The *relative* peak area value obtained for each peak in the extract, including IS1, was used to correct for the small variations in the injection volumes caused by the auto sampler.

The Avalon peak detection algorithm was used for integrating peak area in the chromatogram of *UV* absorbance of 229 nm. The algorithm was used to integrate the peak areas between the retention time range 8-12 and 16-28 min.

The RRF used in this study, for each desulfated glucosinolate was calculated relative to desulfosinigrin as described in the standardised protocol (EEC, 1990). These RRF were in turn derived relative to desulfoglucotropaeolin (IS1), in order to correct for differences in the *UV* absorbance between different desulfated glucosinolates as shown in Table 6.

Table 6 Relative response factor (RRF) for desulfated glucosinolates determined by *UV* absorbance at 229 nm [a] (EEC, 1990) relative to desulfosinigrin, [b] RRF used in this study derived relative to desulfoglucotropaeolin (IS1), based on (EEC, 1990)

Desulfated glucosinolate	RRF [a]	RRF [b]
Desulfoglucoraphanin	1.07	1.13
Desulfoprogoitrin	1.09	1.15
Desulfosinigrin	1.00	1.05
Desulfogluconapin	1.11	1.17
Desulfoglucobrassicin	0.29	0.31
Desulfo4-methoxyglucobrassicin	0.25	0.26
Desulfoneoglucobrassicin	0.20	0.21
Desulfoglucotropaeolin (IS1)	0.95	1.00

The RRF for each desulfated glucosinolate was applied to the relative area/s to correct for variation in absorbance between IS1 and the desulfated components in the extract. The content of each glucosinolate relative to IS1, expressed in (µmoles/g) of completely dried sample was calculated using Equation 1, as described in the standard protocol (EEC, 1990).

$$= \frac{\text{Relative peak area of desulfoglu cosinolate}}{\text{Relative peak area of IS1}} \times \frac{\mu\text{moles of IS1}}{\text{Dry weight (g) of extracted plant material}}$$
$$\times \text{ RRF of deslfogluc osinolate}$$

Equation 1 Calculation of the content of each glucosinolate relative to IS1, expressed in (µmoles/g) of completely dried sample (EEC, 1990)

A flow diagram of the methodology employed in this study is shown in Figure 7. This methodology will be referred to as the optimized *HPLC-UV/ESI-MS/MS* method and was used to generate all the results detailed in section (2.5.8).
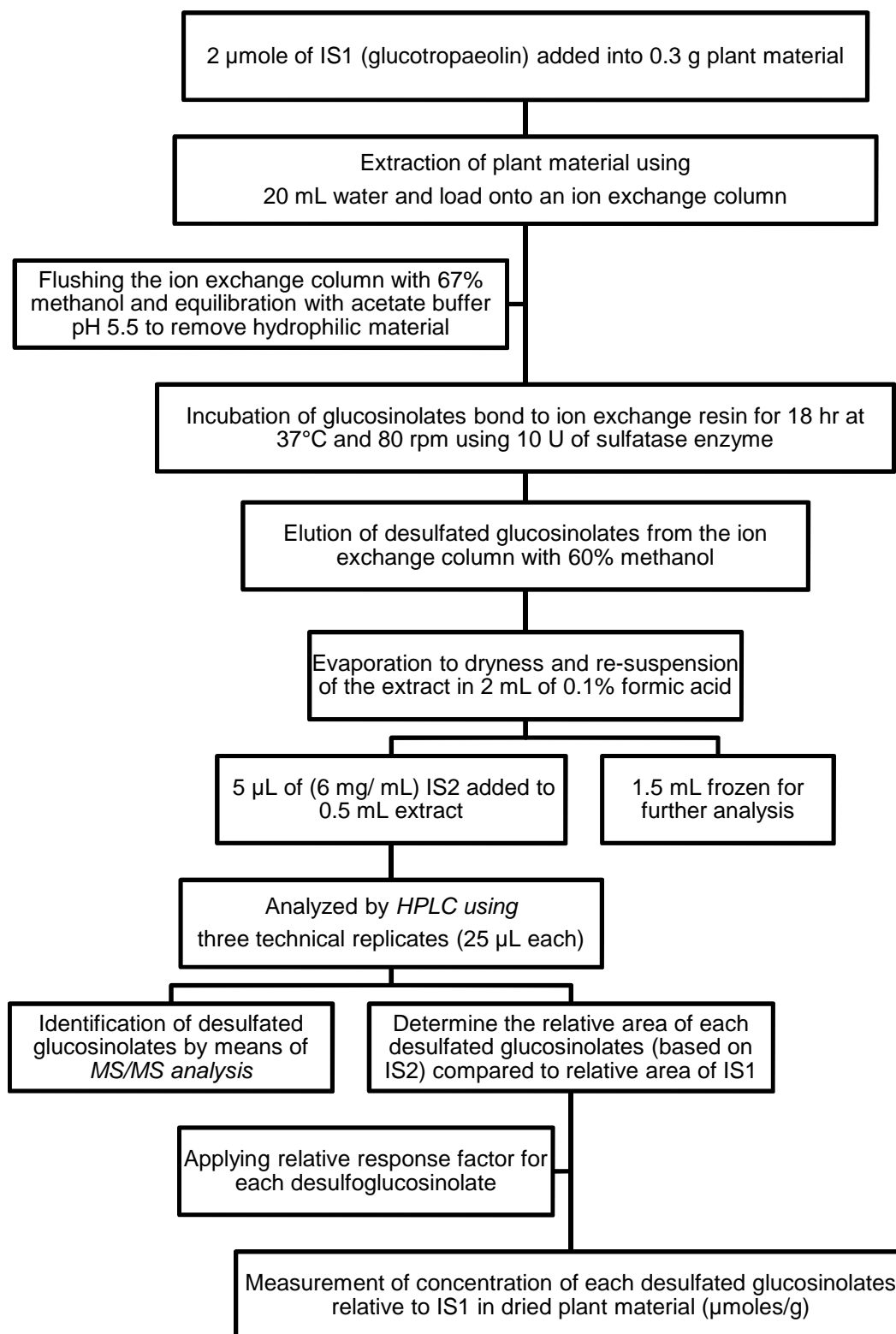
Figure 7 Flow diagram indicating the general protocol developed in this study for the analysis of desulfated glucosinolates from *Brassica* leaves.

2.4.7 Development of a quality control test for the chromatography and *mass spectrometry* performance

It was essential to perform daily quality control (QC) checks to confirm mass accuracy/sensitivity from the *MS* and chromatographic resolution for the *HPLC* prior to any analysis being undertaken. Therefore, it was important to ensure reproducible and stable chromatography prior to any sample analysis, this was achieved by the injection of a sample consisting of 0.1 % formic acid and 5 µL of IS2 at concentration of 6 mg/ mL injected into the *HPLC* using the method described in section 2.4.2) without *MS* analysis. This test was performed several times until the chromatography showed a stable *UV* chromatogram with an IS2 peak appeared at a RT of approximately 8.25 min, with a peak width at half- height of 10 secs.

After that, a QC test solution composed of a mixture of 2 µmoles of sinigrin and 2 µmoles of glucotropaeolin previously desulfated using the protocol described in section (2.3.6) and doped with 5 µL of IS2 at concentration of (6 mg/ mL) was then injected into the *HPLC* using the method described in section (2.4.4). This test was performed at the beginning of the analysis each day and after every 10 samples being analysed, unless any sample during the analysis demonstrated poor chromatography or mass spectrometry. Examples of the chromatography   Figure 8), *MS* (Figure 9) and *MS/MS* (Figure 10) spectra obtained during a QC analysis are shown.
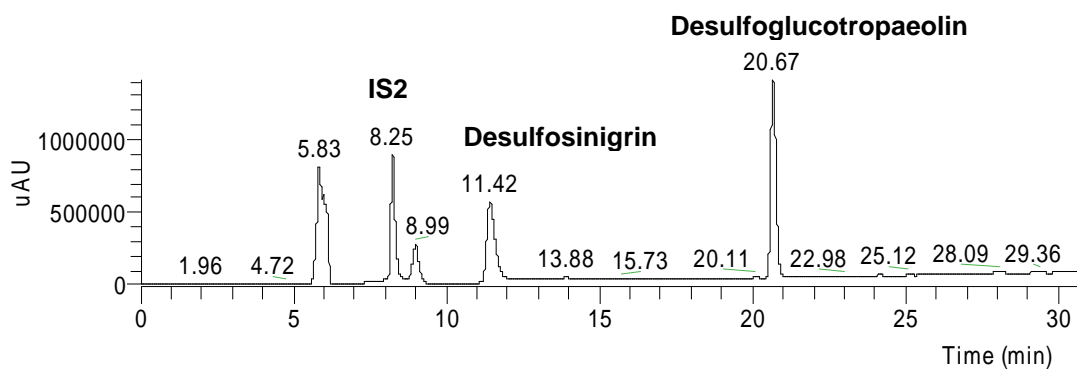
Figure 8 Typical chromatogram obtained at 229 nm of a QC analysis, show peaks of desulfosinigrin, desulfoglucotropaeolin and IS2
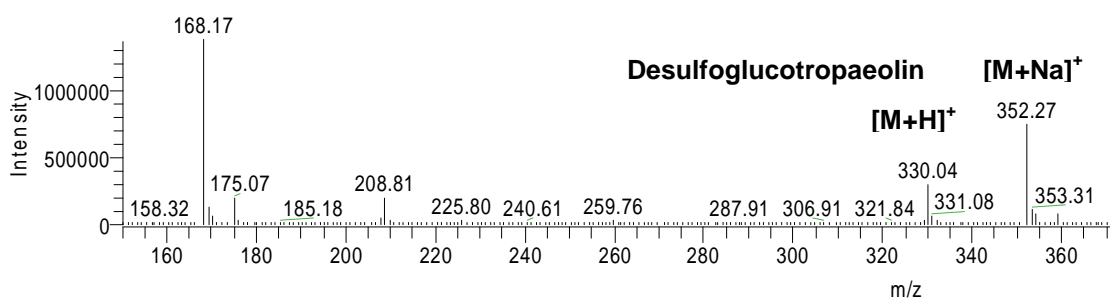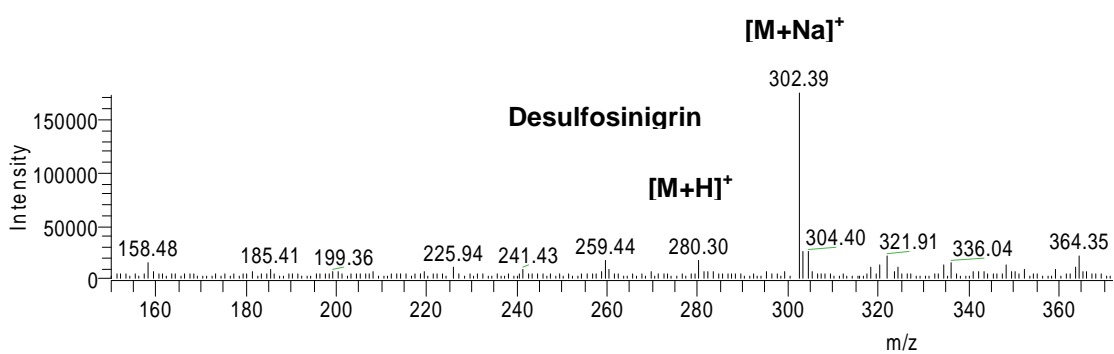
Figure 9 Typical *ESI-MS* obtained from a QC analysis containing 2 µmoles of desulfosinigrin (upper) and 2 µmoles of desulfoglucotropaeolin (lower) respectively
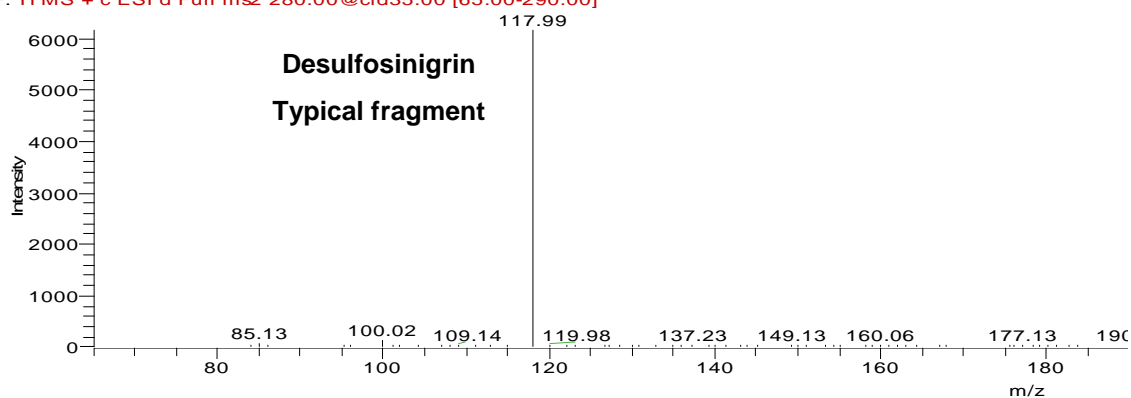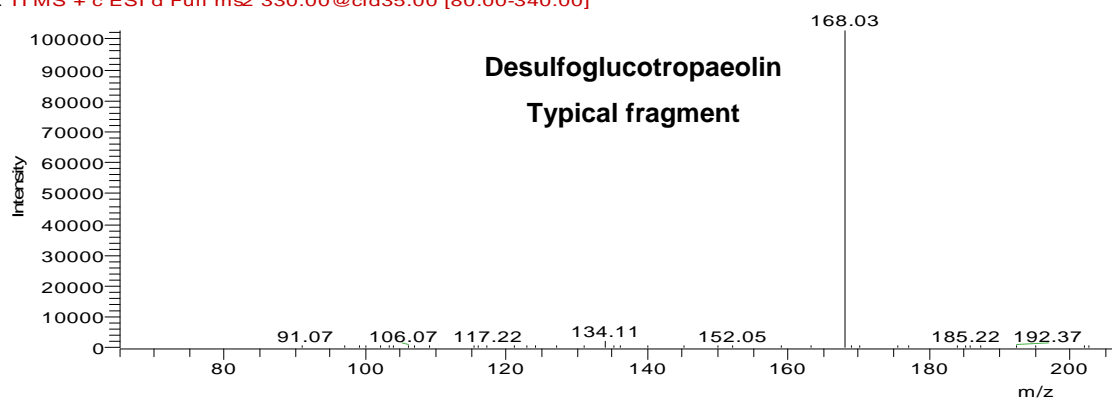
Figure 10 The *ESI-MS/MS* spectra obtained during QC analysis, containing 2 µmoles of desulfosinigrin (upper) and 2 µmoles of desulfoglucotropaeolin (lower), shown are the common fragment ions with *m/z* 117.9 and 168.0, respectively

The typical limits of acceptance for each component in the QC test solution including peak RT, *relative* area, chromatographic resolution calculated as the peak width at half height and the experimentally determined *m/z* for the protonated adduct from the *MS* and *MS/MS* spectra are shown in Table 7.

Table 7 Typical values obtained from a QC analysis including chromatographic RT (min), *relative* peak area (based on IS2), peak width at half height and the experimentally determined *m/z* of the protonated adduct from the *MS* and the characteristic fragment ion observed in the *MS/MS* spectrum

QC solution components	Retention time (min)	*Relative* peak area	Peak width at half height (secs)	[M+H]$^+$ *m/z*	*MS/MS* *m/z*
IS2	8.3	1.0	10	---	---
Desulfosinigrin	11.4	1.14	17	280.3	118.0
Desulfoglucotropaeolin	20.7	1.96	12	330.0	168.0

The accepted shift from the typical values listed in Table 7, were experimentally determined in the range of ±0.3 min, ±0.2 Da and 20% for RT, *m/z* and *relative* peak area, respectively.

The QC test was rejected if any of its components failed to show any of the above listed values within the accepted ranges.

2.5 Results and discussion

2.5.1 Extraction and analysis of intact glucosinolates

Based on the method described by (Song et al., 2006) (and detailed in section 2.4.1.1) using a Lichrospher RP-C18 column, the plant lines A12DH and GD33DH were analysed for the presence of intact glucosinolates. The plant materials were extracted and characterized by *HPLC-MS*. Unfortunately, the chromatography coupled with the mass spectrometry failed to identify any peaks corresponding to individual intact glucosinolates (data not shown), due to the presence of high levels of salts used for eluting intact glucosinolates in fractions from the Sephadex column (as described in section 2.3.3). Therefore, it was decided to desulfate glucosinolates prior to the *HPLC-MS* analysis.

2.5.2 Preliminary qualitative analysis of desulfated glucosinolates

A method for the extraction of the desulfated glucosinolates from *Brassica* leaves was used based on a desulfation enzymatic reaction, utilizing sulfatase enzyme, on an ion-exchange resin, as described by (Brown et al., 2003). Four plant lines; A12DH, GD33DH, AC498, CA25 and a selection of the derived double haploid mapping population; AGDH and NGDH were analysed by the *HPLC-MS* method (as described in section 2.4.1.2) using a Lichrospher RP-C18 column. The chromatography coupled with the mass spectrometry analysis, showed several compounds with *m/z* corresponding to potential desulfated glucosinolates eluting with RT order as expected from the literature.

Different glucosinolate profiles were observed in A12DH and GD33DH, based on

RT and experimentally determined *m/z* values (Table 8). By comparing the two plant

lines profiles, it can be seen that three glucosinolates appear in GD33DH that were

not observed in A12DH; they were desulfoglucoraphanin, desulfogluconasturtiin and

desulfoneoglucobrassicin (Figure 11). Five glucosinolates appear in A12DH that

were not observed in GD33DH; they were: desulfoglucoiberin, desulfoprogoitrin,

desulfosinigrin, desulfoglucoalyssin and desulfoconapin (Figure 12). Only two

compounds appeared to be common to both plant lines, they were

desulfoglucobrassin and desulfo-4-methoxyglucobrassin.

Table 8 Glucosinolates profiles detected in the parental plant lines A12DH and GD33DH in the
preliminary qualitative analysis

Glucosinolates detected in A12DH	Glucosinolates detected in GD33DH	Retention time (min)	*MS (m/z)*
Desulfoglucoiberin		13.0	366.0
Desulfoprogoitrin		14.0	332.0
Desulfosinigrin		17.0	302.0
	Desulfoglucoraphanin	18.0	380.0
Desulfoglucoalyssin		20.0	394.0
Desulfogluconapin		21.0	316.0
Desulfoglucobrassicin	Desulfoglucobrassicin	27.0	391.0
	Desulfogluconasturtiin	28.0	366.0
Desulfo-4-methoxyglucobrassicn	Desulfo-4-methoxyglucobrassicin	29.0	421.0
	Desulfoneoglucobrassicin	31.0	421.0
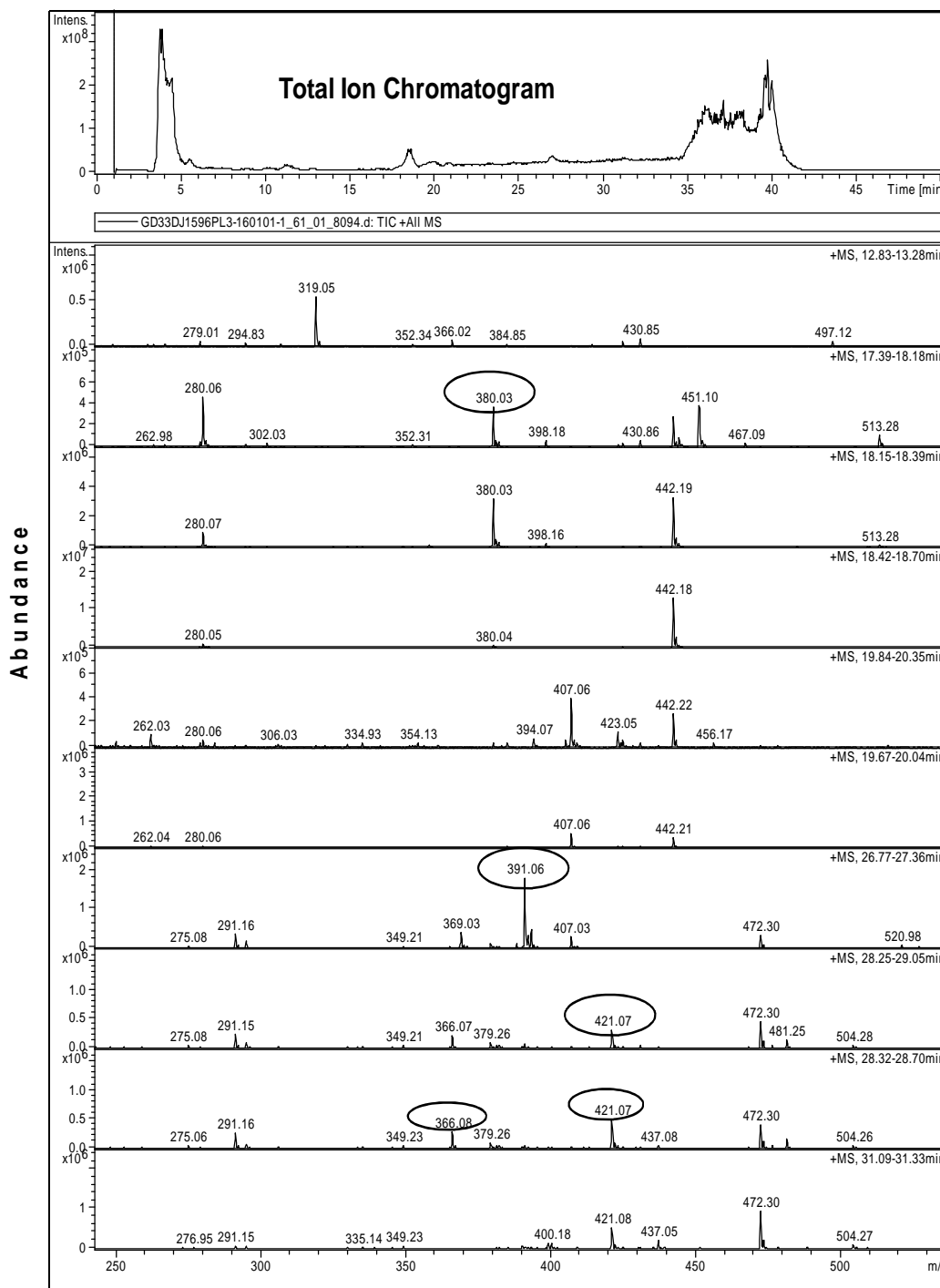
Figure 11 The total ion chromatogram and the mass spectrum of the GD33DH extract in the preliminary qualitative analysis showed five potential desulfated glucosinolates in circles as described in Table 8

60
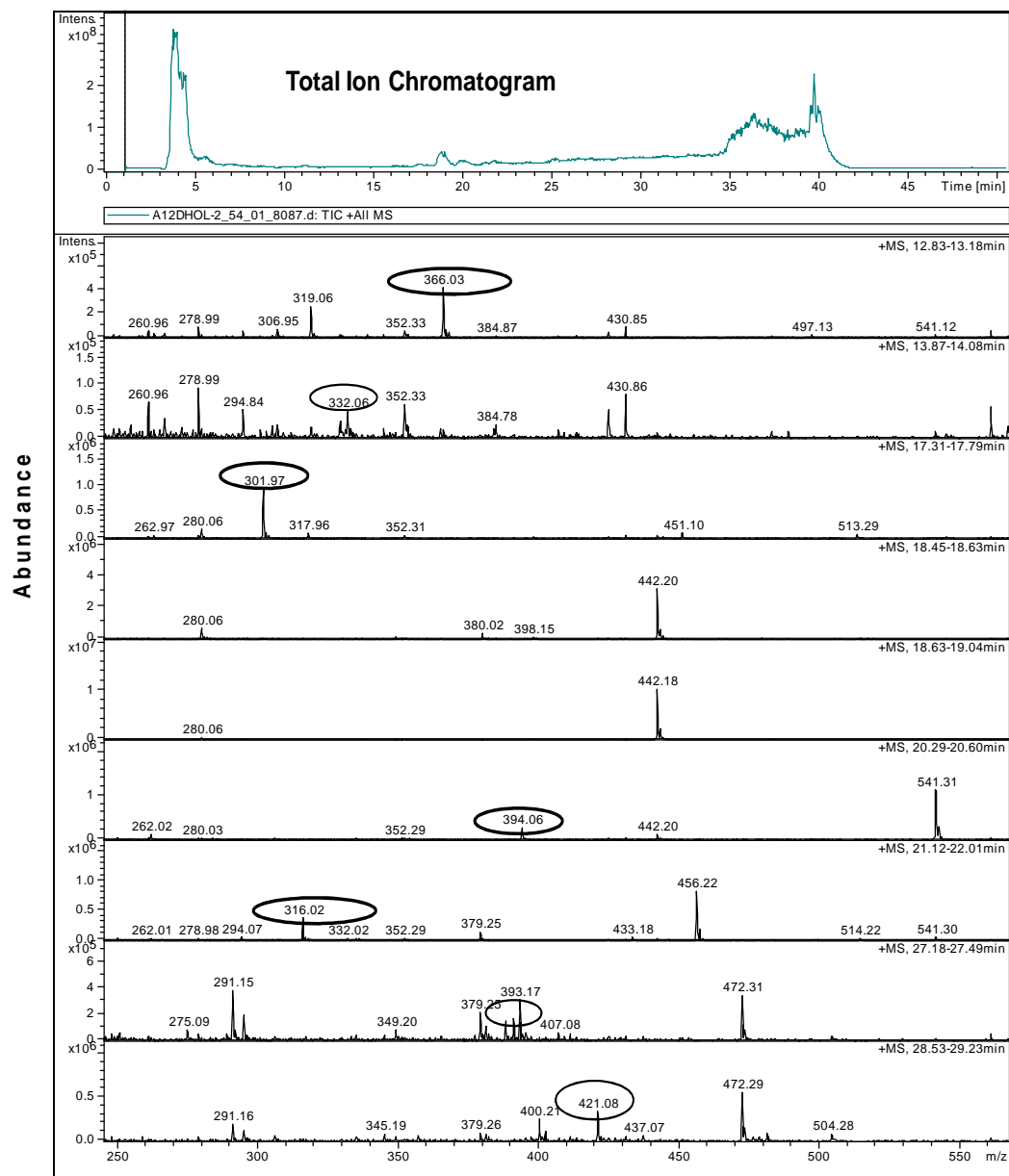
Figure 12 The total ion chromatogram and the mass spectrum of A12DH extract in the preliminary qualitative analysis showed seven potential desulfated glucosinolates in circles as described in Table 8

The glucosinolate profiles observed in AC498 and CA25 lines were compared, as shown in Table 9. Two glucosinolates were only observed in CA25; desulfosinigrin and desulfoneoglucobrassicin, while five compounds were common to both plant lines; desulfoglucoiberin, desulfoglucoraphanin, desulfo4-hydroxyglucobrassin, desulfoglucobrassicin and desulfo-4-methoxyglucobrassin.

Table 9 Glucosinolate profiles detected in the parental plant lines AC498 and CA25 in the preliminary qualitative analysis

Glucosinolates detected in AC498	Glucosinolates detected in CA25	Retention time (min)	*MS* (*m/z*)
Desulfoglucoiberin	Desulfoglucoiberin	13.0	366.0
	Desulfosinigrin	18.0	302.0
Desulfoglucoraphanin	Desulfoglucoraphanin	19.0	380.0
Desulfo4-hydroxyglucobrassicin	Desulfo4-hydroxyglucobrassicin	20.0	407.0
Desulfoglucobrassicin	Desulfoglucobrassicin	28.0	391.0
Desulfo4-methoxyglucobrassicin	Desulfo-4-methoxyglucobrassicin	29.0	421.0
	Desulfoneoglucobrassicin	31.0	421.0

A subset of 36 AGDH douple haploid plant lines derived from cross between A12DH and GD33DH were analysed using the method described by (Brown et al., 2003) and their glucosinolate profiles were found to be varied from their parental lines. Qualitatively, a number of glucosinolates were observed including desulfoglucoiberin, desulfoprogoitrin, desulfosinigrin, desulfoglucoalyssin, desulfoconapin, desulfoglucoraphanin, desulfogluconasturtiin, desulfoneo-glucobrassicin, desulfoglucobrassin, and desulfo-4-methoxyglucobrassin in various combinations (data not shown). At this stage, their were no quantitative mesurments

for glucosinoltes content in these plant samples. The variation in the glucosinolate profiles between the NGDH plants lines were less significant (data not shown).

The measurements obtained were qualitative in nature due to the frequent observation of more than one species eluting in a single *UV* peak (identified by *MS* analysis), which indicated insufficient separation efficiency of the *HPLC* method being used and so further optimisation would be required to obtain appropriate quantitative measurements.

2.5.3 Improvement of the *HPLC* method for optimal separation of desulfated glucosinolates

A sample of 0.5 g of GD33DH was extracted and desulfated as described by (Brown et al., 2003), prior to injection into the *HPLC-MS* using a *Zorbax* Eclipse XDB-C18 column and the mobile phase gradient (as described in Table 3). Poor chromatographic separation for desulfated glucosinolates was observed, especially towards the end of the chromatograph as shown in (Figure 13, A). This mobile phase gradient was improved by increasing the gradient time length with 8 min towards the end of the chromatography and by extending the length of equilibration cycle with 9 min at increased flow rate to improve the reproducibility of the next sample (as described in Table 4). An injection made with a sample of 0.5 g of GD33DH into the *HPLC-MS* using the improved method showed good peak separation for desulfated glucosinolates as shown in (Figure 13, B).
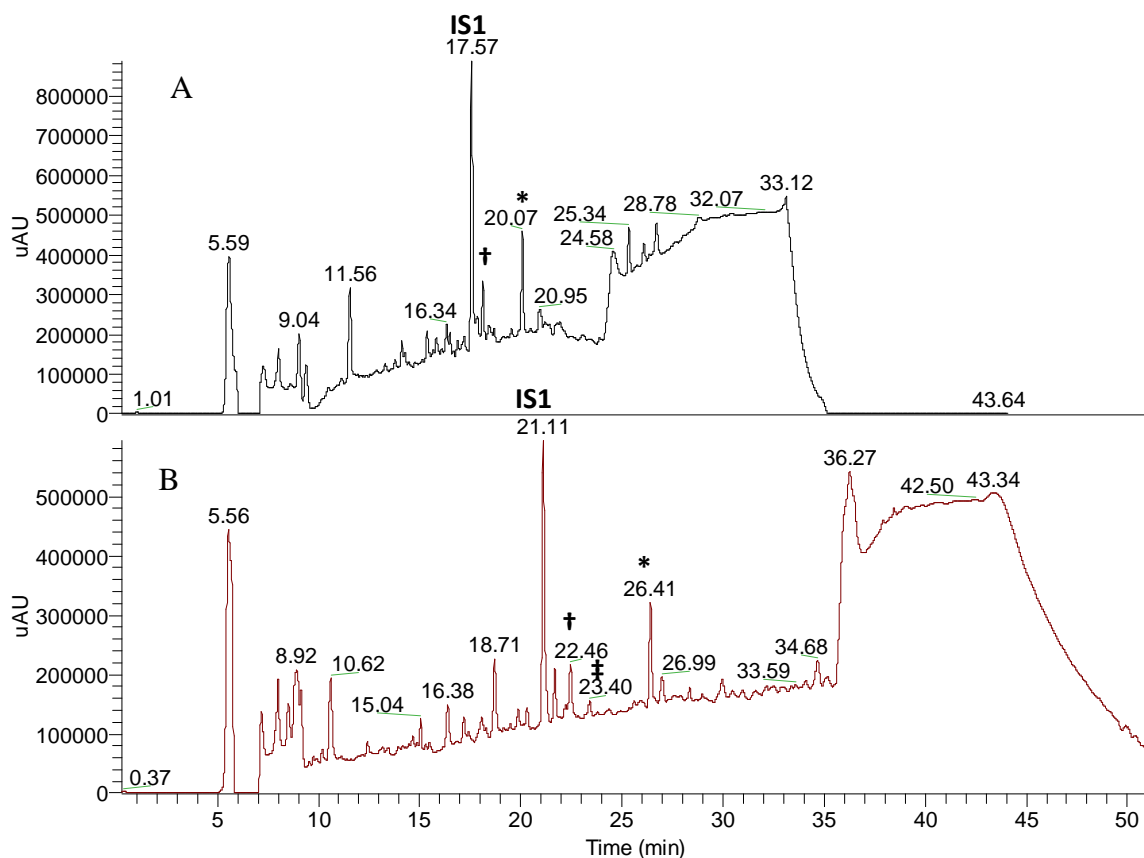
Figure 13 Chromatograms from 0.5 g of GD33DH extracted and desulfated as described by (Brown et al., 2003) prior to injection into the *HPLC-MS* using *Zorbax* Eclipse XDB-C18 column. A: using the mobile phase gradient as described in Table 3; showed poor chromatographic separation for desulfated glucosinolates especially towards the end of the chromatograph. B: improved chromatographic separation was obtained using the mobile phase described in Table 4
IS1: first internal standard (desulfoglucotropaeolin), †: desulfoglucobrassicin, ‡:desulfo4-methoxyglucobrassicin, *: desulfoneoglcobrassicin

A sample of 0.5 g of GD33DH was extracted and desulfated as described by (Brown et al., 2003) and then injected into the *HPLC-MS* as three technical replicates (as described under section 2.4.2) using a *Zorbax* Eclipse XDB-C18 column. With this improved methodology, reproducible chromatography and peak separation with improved resolution was observed between the technical replicates, shown in Figure 14.
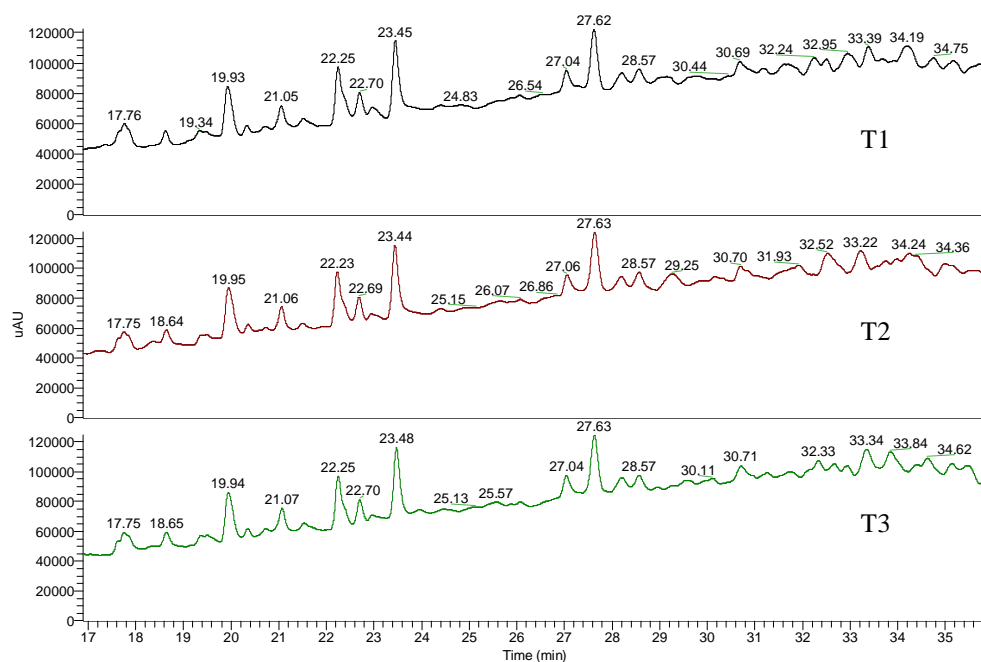
Figure 14 Three chromatograms indicating good resolution of the components from a GD33DH extract and improved reproducibility between technical replicates

2.5.4 **Establishing a robust enzymatic desulfation reaction for intact glucosinolates**

A preparation of GD33DH extract generated from 1 g of plant material was divided into two aliquots and desulfated using the method described by (Brown et al., 2003) and analysed using the *HPLC-MS* method described in section (2.4.2). The chromatograms obtained from the two desulfation reactions showed the presence of desulfated glucosinolates at differing levels in both samples as seen in Figure 15, indicating the lack of desulfation reaction reproducibility from the (Brown et al., 2003) method. Therefore, optimization of the desulfation reaction was required to obtain reproducible quantitative measurements.
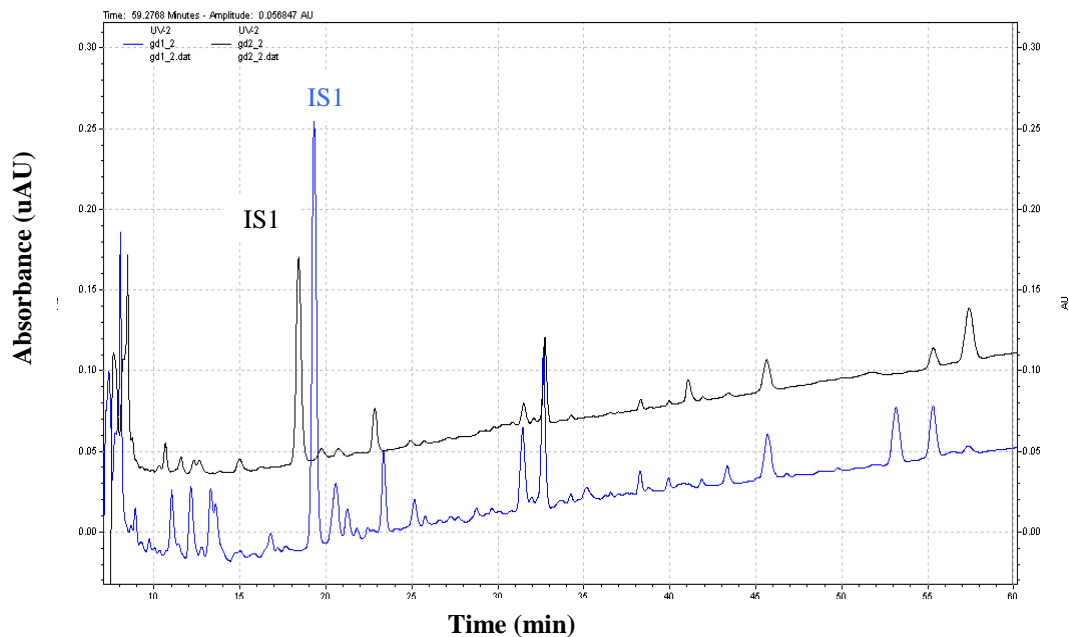
Figure 15 The *UV* chromatograms obtained from two identical samples, processed from the same bulk plant extract of GD33DH plant line, desulfated independently using the non-optimized method*, indicating a lack of reproducibility in the enzymatic desulfation reaction.
* (Brown et al., 2003)

To establish a reproducible and complete desulfation reaction of intact glucosinolates, a batch of homogeneous GD33DH extract was generated from 2.5 g of plant material and divided into five fractions, each fraction was loaded into a Sephadex column. The Sephadex matrix including the bound intact glucosinolates was equilibrated in sodium acetate buffer and treated with different concentrations of sulfatase solution. The Sephadex, plant extract and buffered enzyme were incubated in a shaking incubator as described in section (2.3.6). These samples were then analysed for their desulfated glucosinolate content using the *HPLC-MS/MS* optimized method described in section (2.4.2). The chromatograms obtained from the five concentrations dependent desulfation reactions are shown in Figure 16. Although the desulfated glucosinolates were eluting reproducibly from the *HPLC* column, it was apparent that at the lower enzyme concentration, incomplete desulfation was occurring during the incubation period.
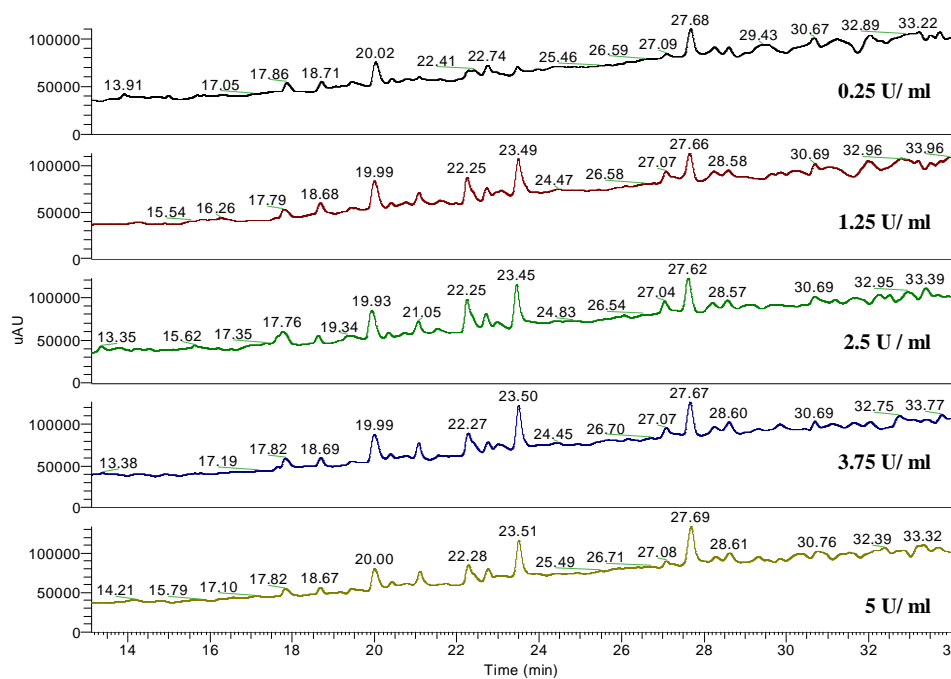
Figure 16 A bulk extract of GD33DH divided into five fractions, each fraction was desulfated using a range of sulfatase solution concentrations; reproducible desulfation reactions of intact glucosinolates were obtained

The average *relative* area (based on IS2) of each detected desulfated glucosinolate peak in the *HPLC-UV* chromatogram at 229 nm was determined in three technical replicates. By increasing the enzyme solution concentration, no obvious differences in the peak areas of the first eluting desulfated glucosinolate (desulfoglucoraphanin) were observed. An increase in the average *relative* area was observed for the last three eluting desulfated glucosinolates (desulfoglucobrassicin, desulfo-4-methoxyglucobrassicin and desulfoneoglucobrassicin), until the measurements plateaued within the range 1.25-3.75 U/ mL, which indicated that the desulfation reaction of these glucosinolates had reached equilibrium. For desulfo-4-methoxyglucobrassicin, the average *relative* area showed a plateau at enzyme solution concentration range of 3.75-5 U/ mL (Figure 17). Within this range,

desulfation of 4-methoxyglucobrassicin (as well as all other glucosinolates in the extract) reached equilibrium.
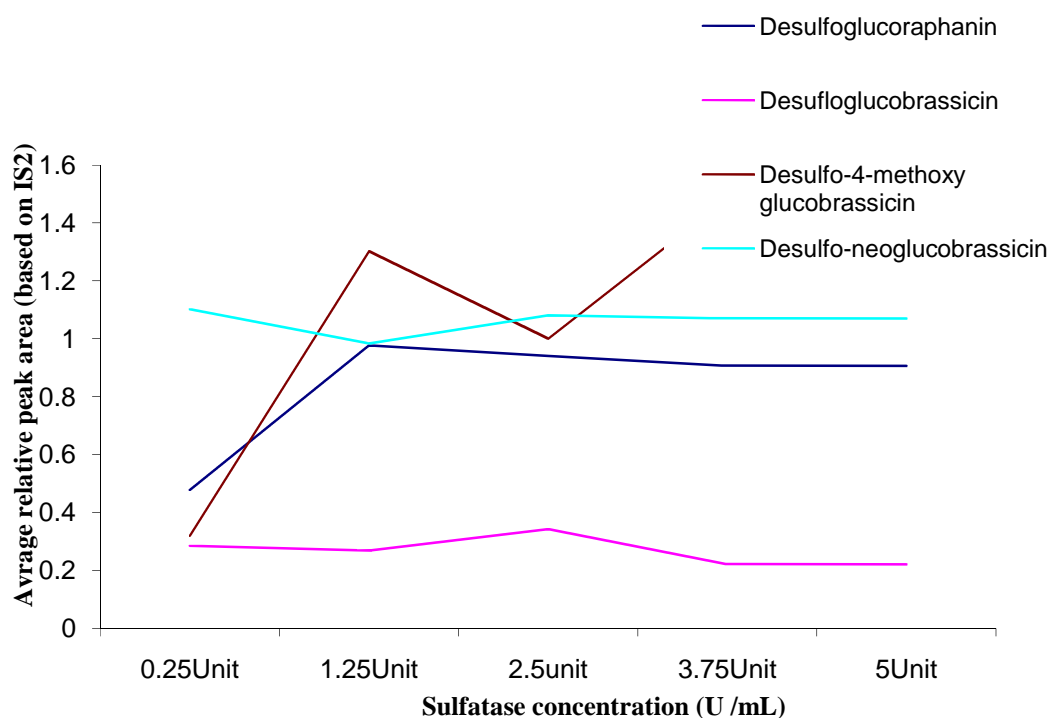
Figure 17 Plot of the *average* relative area (based on IS2) of each desulfated glucosinolate against increasing concentrations of sulfatase solution, indicates the desulfation reaction of all glucosinolates reached equilibrium at the concentration range of (3.75-5 U/ mL)

In order to ensure a robust enzymatic desulfation reaction of glucosinolates across the 89 plant lines which may express glucosinolates at higher content or with different composition, the enzyme solution concentration used in this study was doubled to 10 U/ mL.

2.5.5 Determination of the optimal ratio of IS1 added to the plant material prior to extraction

In order to quantify desulfated glucosinolates, the correct level of IS1 should be experimentally determined prior to extraction. IS1 was added to the plant material

prior to extraction in order to reduce variations in the extraction and desulfation reactions that may occur between different samples. The optimal concentration of IS1 in the extract was determined using different ratios of IS1 to plant material as described under section (2.3.7).

In two independent experiments; 0.5 g of GD33DH material were dopped with 0.5 μmole and 2 μmoles of IS1, equivalent to the ratios of (1g: 1 μmole) and (1g: 4 μmoles); respectively, and then extracted and desulfated prior to the *HPLC-MS/MS* analysis (as described in sections 2.3.6). No IS1 was observed in either experiment (Figure 18, A and B respectively). This suggested the need to increase the ratio of IS1 to plant material extracted in order to increase the intensity of the IS1 peak to a detectable level. Consequently, 0.3 g of GD33DH was extracted with 2 μmole of IS1 (equivalent to the ratio of 1 g: 6.7 μmoles) prior to desulfation and *HPLC-MS/MS* analysis. In the *UV* chromatogram; IS1 was observed at approximately 20.7 min (Figure 18, C), with an absolute peak area was within the range observed from endogenous desulfated glucosinolates in the plant lines.
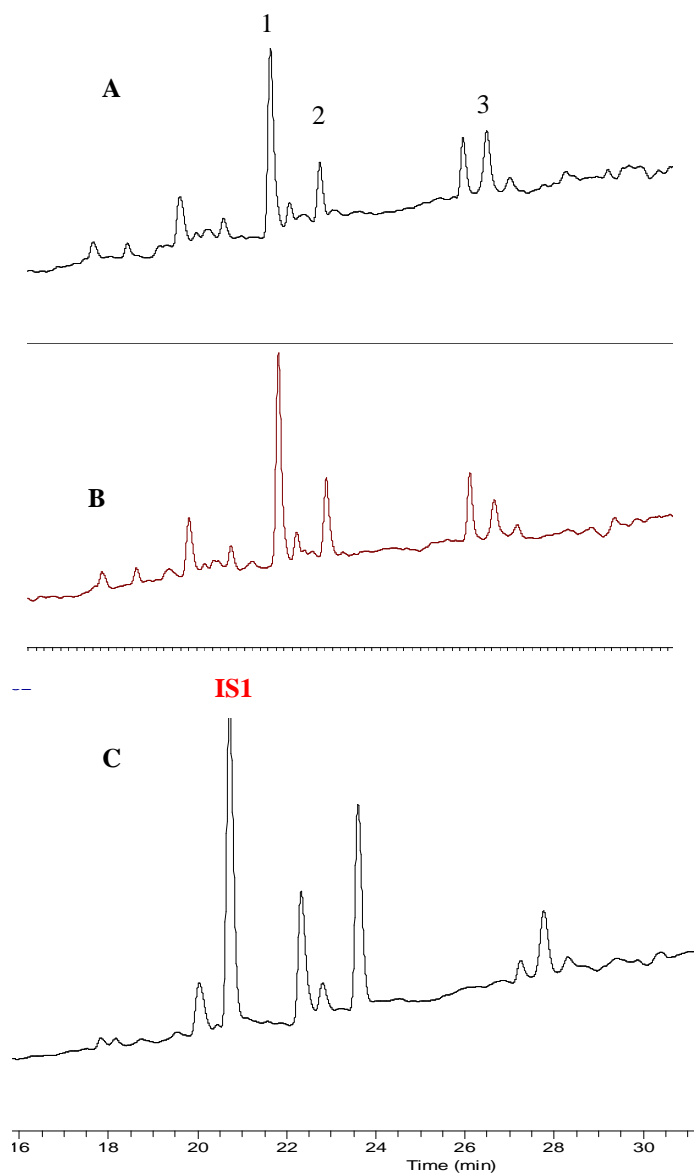
Figure 18 Chromatograms of GD33DH extracted with different ratios of the internal standard glucotropaeolin (IS1). A: (1g: 1 µmole), B: (1g: 4 µmoles), and at the optimal ratio shown in chromatogram C: (1g: 6.7 µmoles), where IS1 peak appeared at 20.7 min.

1: desulfoglucobrassicin, 2: desulfo-4-methoxyglucobrassicin, 3: desulfoneoglucobrassicin

2.5.6 **Determination of the optimal concentration of IS2 used to improve reproducibility of the quantitative measurements**

Initial experiments were performed using only one internal standard (IS1) and the standard deviation of the peak areas obtained for three technical replicates calculated. 2 µmoles of IS1 were added into 0.3 g of GD33DH leaf material and extracted prior to analysis by *HPLC-UV/ESI-MS/MS* using the method described in section (2.4.4). The absolute peak areas of each desulfated glucosinolate in the *UV* chromatogram at 229 nm (obtained from the Avalon peak detection algorithm as described in section 2.4.6) were converted relative to IS1 peak area. The average and standard deviation in three technical replicates, and the % of the ratio of standard deviation to the average peak area are shown in Table 10. The relative areas (from IS1) observed showed variation, which may lead to inaccurate or imprecise quantitative measurements.

Table 10 Integrated peak area measurements for individual desulfated glucosinolate from GD33DH extract, relative to IS1 peak. The average and standard deviation are shown for three technical replicates. The % of the ratio of standard deviation to the average peak area indicating variations in the *relative* peaks area

Compounds	Relative area T1	Relative area T2	Relative area T3	Average replicate	STDEV replicate	% STDEV relative to peak area
IS1	1.0	1.0	1.0	1.0	0.0	----
Desulfoglucoraphanin	0.0682	0.0774	0.0631	0.0696	0.0072	10%
Desulfoglucobrassicin	1.5945	1.9597	1.4983	1.6842	0.2434	14%
Desulfo4-methoxyglucobrassicin	2.3474	2.6953	2.3068	2.4499	0.2135	9%
Desulfoneoglucobrassicin	2.7739	3.0643	2.4677	2.7687	0.2983	10%

In order to improve the accuracy of the quantitative measurements for the plant extract components, IS2 was added to each sample prior to injection into the *HPLC* and the standard deviation of the measurements calculated. The optimum level of IS2 was experimentally determined and therefore, three solutions consisting of 5 µL of IS2 at concentrations of 1 mg/ mL, 2 mg/ mL and 6 mg/ mL were diluted in 500 µL of water and injected into the *HPLC* as described under section (2.3.82.4.4). In the obtained chromatograms of the three samples, IS2 peak eluted at 8.7 min; interestingly only in the injection made with the solution containing (6 mg/ mL) was an IS2 peak detected. The more diluted solutions produced a peak area below the useful level required for its use as a base peak (data not shown).

In order to validate the optimal level of IS2 in the plant extract, a sample consisted of 2 µmoles IS1 added into 0.3 g of GD33DH leaf material were extracted with water and doped with 5 µL of IS2 at a concentration of 6 mg/ mL (1.5 µg on column), was then injected into the *HPLC* in three technical replicates (as described under section 2.4.4). In the *UV* chromatograms, IS2 peak eluted at 8.14 min within the linear range of the *UV* detector at 229 nm (Figure 19).
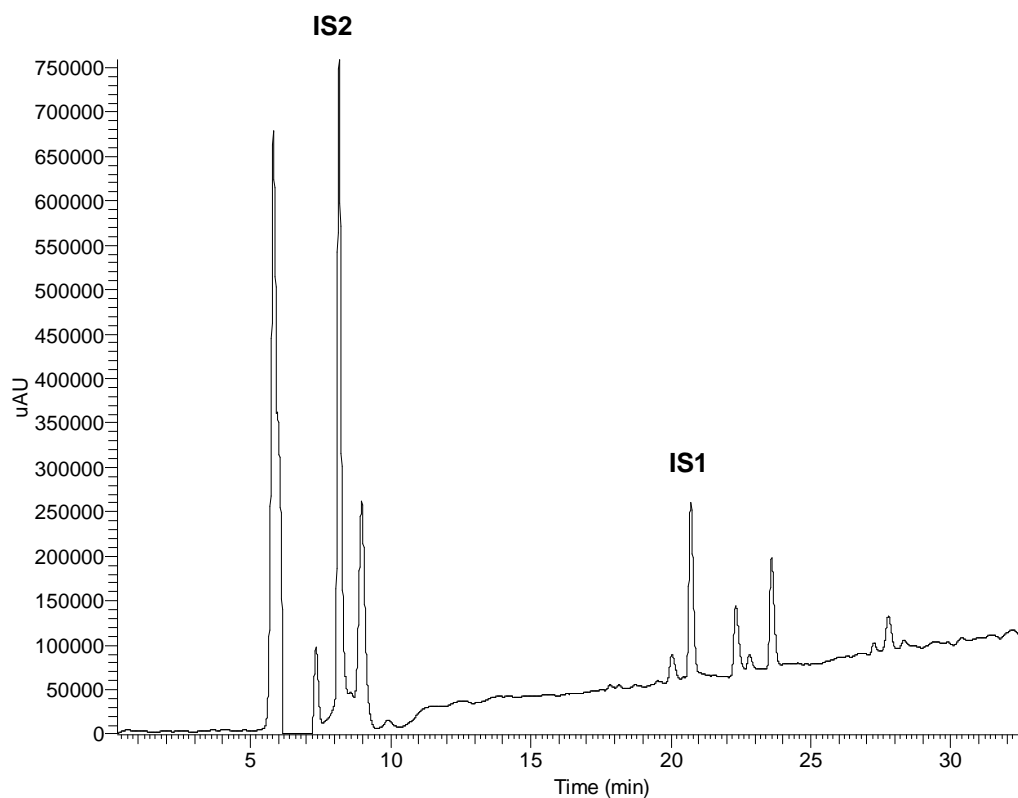
Figure 19 The chromatogram of a sample consisting of 0.3 g GD33DH extracted with 2 µmoles of IS1 and doped with IS2 at concentration of 1.5 µg on column, IS2 peak eluted at approximately 8.17 min was used as a base peak to correct for variations caused by the autosampler observed between the technical repeats

For the quantitative measurements of desulfated glucosinolates, the *relative* peak area for individual compounds was calculated based on the IS2 peak. The standard deviations and the % of the ratio of standard deviation to the average peak area are shown in Table 11. The variations in *relative* areas between the three technical replicates were significantly reduced by applying the peak area of IS2 instead of IS1 as a base peak for all other peaks of desulfated glucosinolates in the extract.

Table 11 Integrated peak area measurements for individual desulfated glucosinolates from GD33DH extract, based on IS2, the average and standard deviation are shown for three technical replicates. The % of the ratio of standard deviation to the average peak area indicating very low variation in the *relative* peaks area obtained from different injections

Compounds	Relative area T1	Relative area T2	Relative area T3	Average replicates	STDEV replicates	% STDEV relative to peak area
IS2	1.0	1.0	1.0	1.0	0.0	----
IS1	0.0788	0.0792	0.0744	0.0775	0.0027	3.5%
Desulfoglucoraphanin	0.0794	0.0805	0.0686	0.0761	0.0066	8.7%
Desulfoglucobrassicin	0.0305	0.0296	0.0349	0.0317	0.0029	9.0%
Desulfo4-methoxyglucobrassicin	0.1167	0.1107	0.1056	0.1111	0.0057	5.0%
Desulfoneoglucobrassicin	0.1382	0.1259	0.1129	0.1257	0.0126	10.0%

2.5.6 Development of an automated *MS* and *MS/MS* method to confidently identify desulfated glucosinolates

In order to identify desulfated glucosinolates in plant extracts, an automated *MS* analysis was developed utilizing the Thermo Fisher Scientific *LTQ XL* electrospray ionization mass spectrometer with linear ion trap mass analyser (as described under section (2.4.5).

Within the 89 AGDH plant population studied in this work, each plant line was dopped with IS1, desulfated and an aliquot of IS2 added prior to injection into the *HPLC-MS/MS* (as described in section 2.4.4). Several compounds were identified as potentially desulfated glucosinolates, by comparing their *m/z* and RT with well known glucosinolates previously detected in *Brassica* species. They were (in order of their RT in the chromatogram); desulfoglucoraphanin, desulfoprogoitrin, desulfosinigrin, desulfogluconapin, desulfoglucobrassicin, desulfo-4-methoxyglucobrassicin and desulfoneoglucobrassicin. These desulfated

glucosinolates were characterised by the *m/z* of protonated and sodiated molecular ions, [M+H]$^+$ and [M+Na]$^+$ respectively, as shown in Table 12. Absolute confirmation of the identity of the desulfated glucosinolates was not possible by RT and *m/z* alone and thus an extra dimension of the analysis was included. A list of the protonated *m/z* of commonly occurring desulfated glucosinolates was generated within the Xcalibur software (as described in section 2.4.5). If any of the desulfated glucosinolates were detected by their *m/z* within a given RT window, the precursor ion was selected for *MS/MS* analysis. The *MS/MS* spectra generated were inspected for the characteristic loss of the sugar group. A combination of RT, *m/z* and the presence of the characteristic fragment ion was used to confirm the detection of a desulfated glucosinolate.

Table 12 Desulfated glucosinolates detected from different AGDH plant lines identified with their RT and *m/z* in the *MS* and confirmation of identify by the characteristic loss of 162.1 Da in the *MS/MS* spectrum. Additional structure-specific fragments were used to distinguish desulfated glucosinolates with identical molecular weight

Desulfated glucosinolates	Retention time (min)	[M+Na]$^+$ *(m/z)*	[M+H]$^+$ *(m/z)*	*MS/MS* fragment ion after loss of sugar group *(m/z)*	*MS/MS* other characteristic fragment ion *(m/z)*
Desulfoglucoraphanin	8.5	380.0	358.0	196.0	
Desulfoprogoitrin	8.7	330.0	310.0	148.0	
Desulfosinigrin	11.2	302.0	280.0	118.0	
Desulfogluconapin	16.3	316.0	294.0	132.0	
Desulfoglucobrassicin	22.2	391.0	369.0	207.0	
Desulfo-4-methoxyglucobrassin	23.5	421.0	399.0	237.0	160.0 [R]$^+$
Desulfoneoglucobrassicin	27.0	421.0	399.0	237.0	205.0 [RCNOH+2H]$^+$ 177.0 [ROH]$^+$ 130.0[R-CH3O+H]$^+$

In the mass spectra of the plant lines expressing desulfo-4-methoxyglucobrassicin and desulfoneoglucobrassicin, which eluted at RT 23.5 and 27.2 min respectively, both glucosinolates have the same *(m/z)* value of 399.0 for the [M+H] [+] molecule (Figure 20, A). As expected they showed the highest intensity for the common typical fragment molecule with *m/z* of 237.0 in the *MS/MS* spectrum (Figure 20, B and C). These two glucosinolates were differentiated by comparing their RT with pure standard desulfoneoglucobrassicin. From this it was possible to confirm the later peak at 27.2 min as desulfoneoglucobrassicin.

Additionally, other structure specific fragments dependent on the R side chain, were used for structural determination of desulfated glucosinolates (Griffiths et al., 2000; Zimmermann et al., 2007). Fragmentation of desulfoneoglucobrassicin produces fragment molecules with *(m/z)* of 205.0, 177.0 and 130.0, corresponding to [RCNOH+2H] [+], [ROH] [+] and [R-CH3O+H] [+] molecules; respectively (Figure 20, C). Fragmentation of desulfo-4-methoxyglucobrassicin produced a structure specific fragment with *(m/z)* of 160.0, which correspond to the [R] [+] molecule (Figure 20, B).
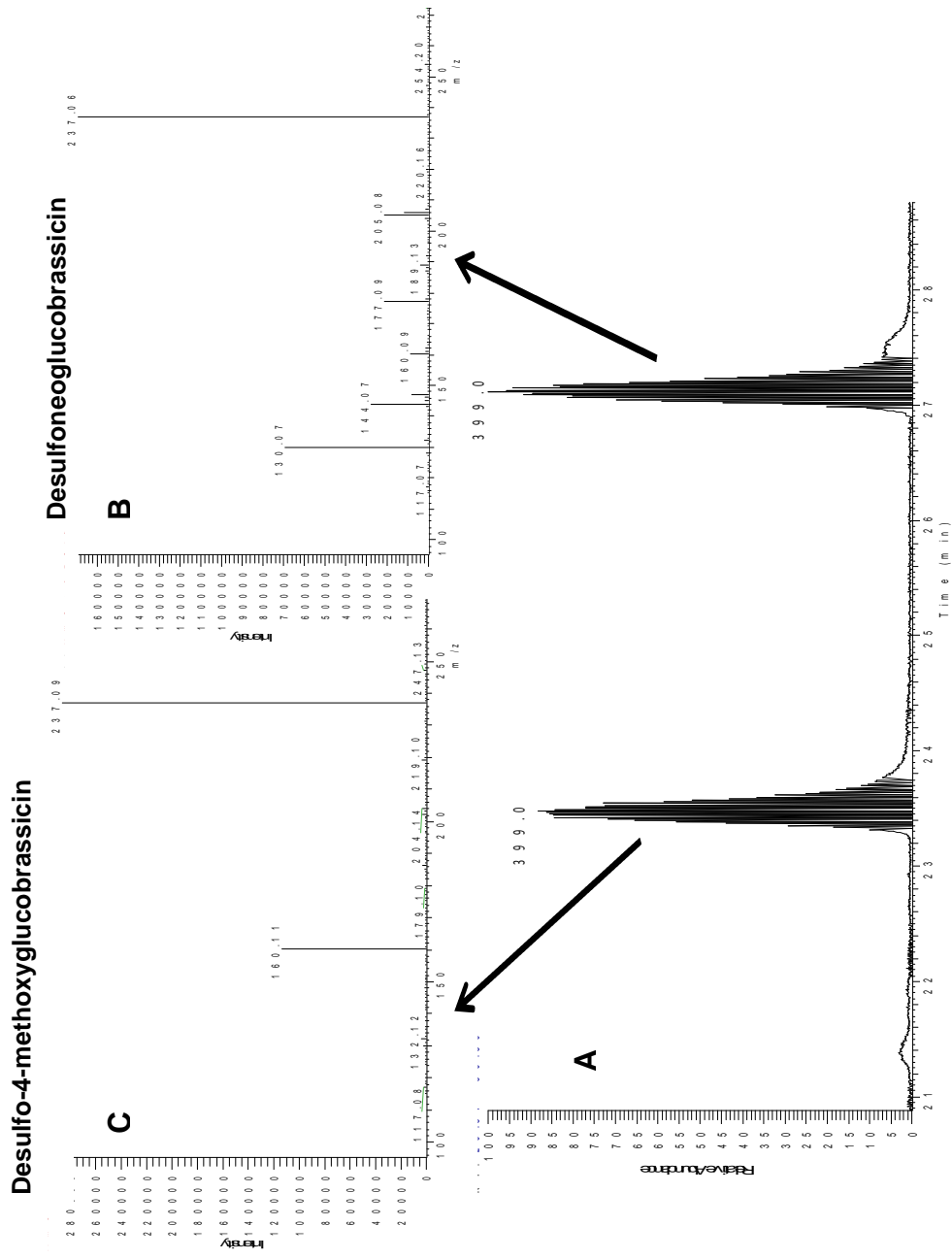
Figure 20 The chromatographic separation and *MS* produced by *HPLC-MS/MS* of desulfated glucosinolates extracted from an AG plant line expressing desulfo-4-methoxyglucobrassicin and desulfoneoglucobrassicin eluted at 23.5 and 27.2 min respectively.(A) In the *MS* full scan mode, both glucosinolates were identified at *m/z* 399.0 for the [M+H]$^+$ ions. The most abundant ion in the *MS/MS* spectrum was *m/z* 237.0. These compounds were distinguished by additional specific fragments in the *MS/MS* spectrum (B) *m/z* of 130.0, 177.0 and 205.0 for desulfoneoglucobrassicin (C) *m/z* 160.0 for desulfo-4-methoxyglucobrassicin

77

In an AGDH plant line extract, a compound eluted at 37 min and with m/z of 294.5. This compound, corresponding to the mass of desulfogluconapin, was subjected to *MS/MS* analysis for further confirmation. No fragment ion at ($m/z)$ of 132.0 was observed for this compound (Figure 21) and thus, it was established that this compound was not desulfogluconapin.
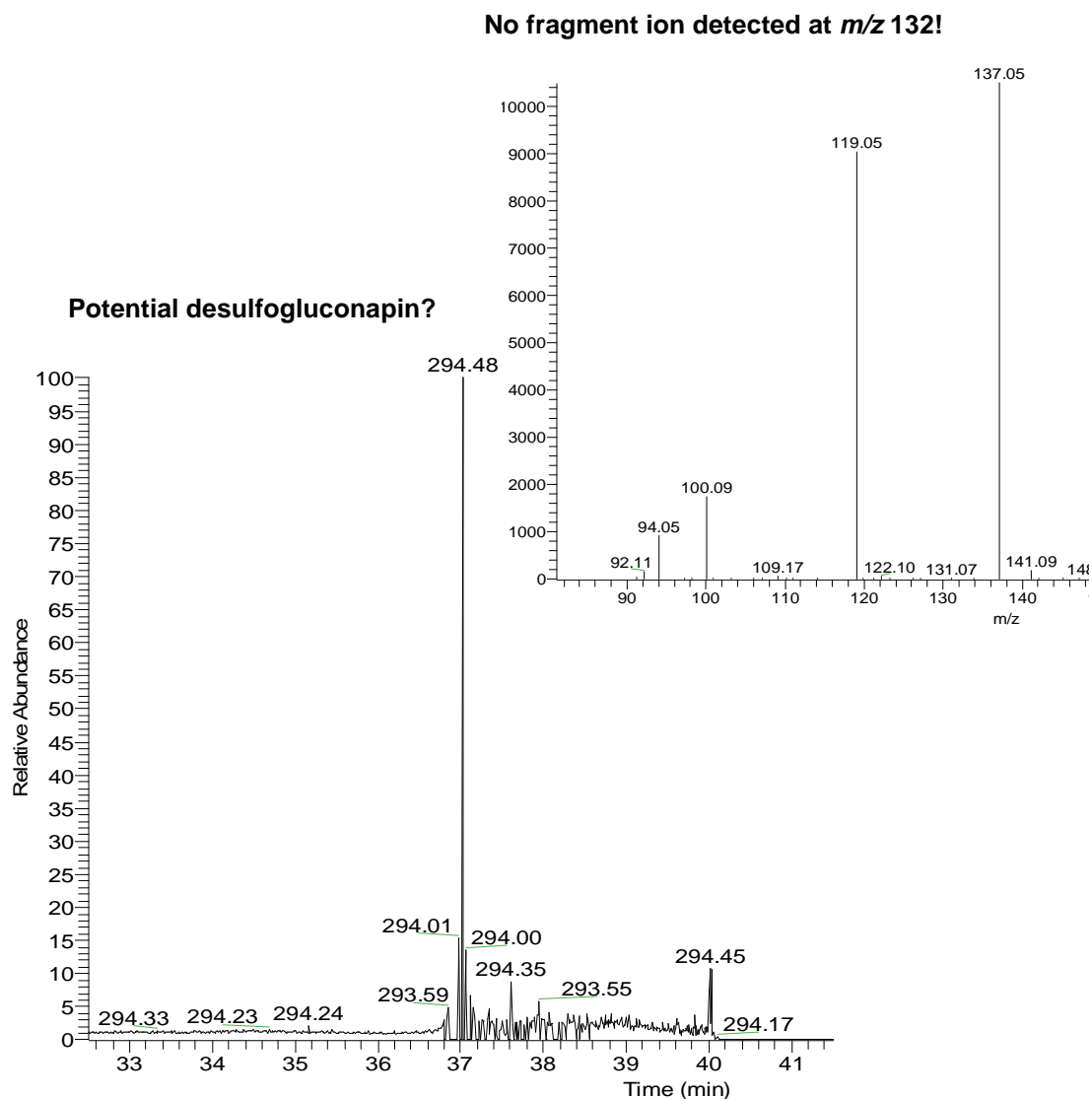
Figure 21 The **c**hromatographic separation and *MS* produced by *HPLC-MS/MS* of an AGDH plant line extract. Peak eluted at 37 min had *m/z* 294.5 corresponding to desulfogluconapin, but failed to show the expected fragment ion at *m/z* 132.0 in the *MS/MS* spectrum

In order to quantify desulfated glucosinolates across all plant lines being analysed in this study, it was essential to determine the lower limit of detection for individual desulfated glucosinolates based on the minimum *relative* concentration at which the characteristic fragment in the *MS/MS* spectrum was observed. Therefore, a generic limit of detection for each desulfated glucosinolates observed in the mass spectrum of plant extracts was calculated and is shown in Table 13

Table 13 The lower limit of detection for individual desulfated glucosinolates, based on the observation of the characteristic fragment in the *MS/MS* spectrum

Desulfated glucosinolates	Lower limit of detection (µmole/ g dry plant material)
Desulfoglucoraphanin	0.30
Desulfoprogoitrin	5.00
Desulfosinigrin	0.90
Desulfogluconapin	0.50
Desulfoglucobrassicin	0.35
Desulfo4-methoxyglucobrassin	0.03
Desulfoneoglucobrassicin	0.05

2.5.7 Effect of Relative Response Factors on quantitative measurements

In this work, the determination of the qualitative and quantitative profiles of glucosinolates in a *B. oleracea* population of 89 AGDH plant lines, produced from a cross between A12DH and GD33DH, was based on the standardised protocol of the (EEC, 1990) method.

Therefore, for each plant line 0.3 g of dried leaf material were dopped with 2 µmoles of IS1, desulfated with 10 U of sulfatase enzymes and injected into the *HPLC-UV/ESI-MS/MS* analysis as described in the flow diagram shown in Figure 7 Flow

diagram indicating the general *protocol* developed in this study for the analysis of desulfated *glucosinolates* from Brassica leaves.

The quantification of individual glucosinolates was based on the peak area for each compound observed in the *UV* chromatogram at 229 nm calculated *relative* to IS2 peak and compared to the *relative* peak area of IS1. The RRF for each desulfated glucosinolate was applied to the relative area/s. The *relative* concentration of each glucosinolate (expressed in µmoles/ g of dried plant material) was obtained (as described in section 2.4.6).

Several response factor values for desulfoglucosinolates are available in the literature and by applying different values of the RRFs, the calculation of the content of individual glucosinolates varied. For example, 0.3 g of AGDH6044 plant line was dopped with 2 µmoles of IS1, desulfated with 10 U of sulfatase enzymes and injected into the *HPLC-UV/ESI-MS/MS* (as described in Figure 7). By comparing the content of individual glucosinolates in the extract calculated using the RRF from the (EEC, 1990) method with their content using RRF from the (Brown et al., 2003) method (Table 14), higher concentrations were obtained from the latter method for all glucosinolates (Figure 22), except for glucoraphanin and gluconapin as they have identical RRF in both methods.

Table 14 Relative response factors (RRF) for desulfated glucosinolates relative to desulfoglucotropaelin (IS1) determined at *UV* absorbance 229 nm in different laboratories. [a] (EEC, 1990) [b](Brown et al., 2003)

Desulfated glucosinolate	RRF [a]	RRF [b]
Desulfoglucoraphanin	1.13	1.13
Desulfoprogoitrin	1.15	Not measured
Desulfosinigrin	1.05	1.25
Desulfogluconapin	1.17	1.25
Desulfoglucobrassicin	0.31	0.38
Desulfo4-methoxyglucobrassicin	0.26	0.38
Desulfoneoglucobrassicin	0.21	0.25
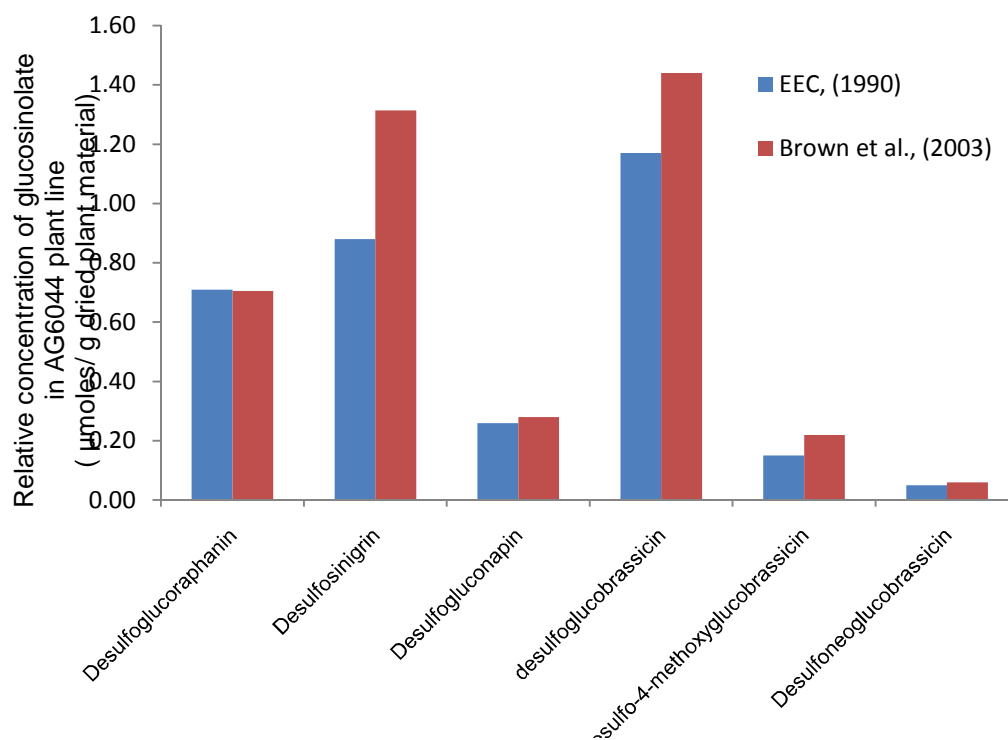
Figure 22 Bar chart representing the relative concentration of individual glucosinolates in the AGDH 6044 plant line, calculated by using different RRF. Higher relative concentrations were obtained using RRF from the (Brown et al., 2003) method as described in Table 14.

2.5.8 Determination of glucosinolate profiles in the AGDH population

Within the 89 AGDH plant lines analysed in this work, seven desulfated glucosinolates were identified using the *HPLC-MS/MS* method developed in this study as described under section (2.4.5). They were found to chemically belong to two different groups, the aliphatic and the indolic glucosinolates. The aliphatic glucosinolate group contains glucoraphanin, progoitrin, sinigrin and gluconapin. The indolic glucosinolate group contains glucobrassicin, 4-methoxyglucobrassicin and neoglucobrassicin.

As expected, the AGDH plant lines contained glucosinolates in different combinations from their parental lines, where the A12DH parental line was found to contain sinigrin, gluconapin, glucobrassicin and 4-methoxyglucobrassicin, the GD33DH parental line was found to express glucoraphanin, glucobrassicin, 4-methoxyglucobrassicin and neoglucobrassicin.

Quantification of the glucosinolate content in the AGDH plant lines was calculated *relative* to IS1 (µmoles/ g of dry plant material) for individual glucosinolates using the RRF published in the (EEC, 1990) protocol, and the quantification method validated in this study (as described in section 2.4.6). These data are presented in Appendix D.

The quantitative analysis of the individual glucosinolate concentration revealed that the AGDH population expressed higher levels of aliphatic than indolic glucosinolates as shown in Figure 23. The box plot for the average relative concentration in three technical replicates for the individual glucosinolates showed variability in the percentage concentration of each glucosinolate expressed in the AGDH population

presented by interquartile range. In addition, the median, maximum and minimum concentrations of individual glucosinolates showed variability between the AGDH plant lines.
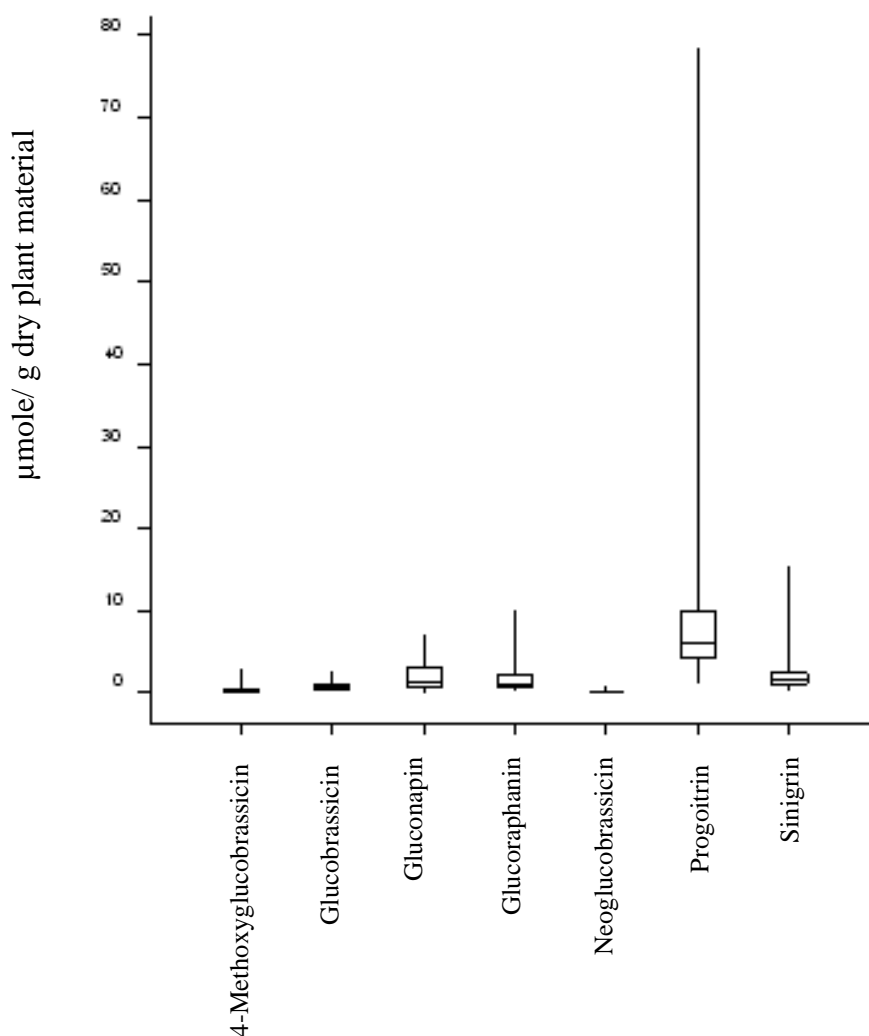
Figure 23 Box plot presenting the variability in the percentage for the average relative concentration of each glucosinolate expressed in the AGDH population. The boxes represented the interquartile range, with median indicated by a line within the box. The whiskers attached to the boxes represent the maximum and minimum concentrations

The predominant glucosinolate in the AGDH plant lines was progoitrin which was present in greater amounts than all the other glucosinolates with average relative concentration of 12.95 µmoles/ g of dry leaves material, followed by sinigrin (2.31 µmoles/ g), gluconapin (1.92 µmoles/ g), glucoraphanin (1.72 µmoles/ g),

glucobrassicin (0.75 µmoles/ g), 4-methoxyglucobrassicin (0.43 µmoles/ g), and neoglucobrassicin (0.22 µmoles/ g).

The ratio between the maximum and minimum concentrations of individual glucosinolate was calculated and Log_{10} transformed. Interestingly, the highest variation was observed between the minimum and maximum levels was with 4-methoxyglucobrassicin exhibiting 2 dynamic fold and the lowest variation was observed for glucobrassicin and neoglucobrassicin with 1.2 dynamic fold only, where all other glucosinolates were with similar range of variation of (1.8-1.6) dynamic fold (Table 15).

Table 15 Variations observed between glucosinolates content in the AGDH population calculated as the Log_{10} of the ratio between the maximum and minimum concentrations of individual glucosinolate

Glucosinolates	Variations between maximum and minimum concentrations (dynamic fold)
4-Methoxyglucobrassicin	2.0
Glucobrassicin	1.8
Progoitrin	1.8
Sinigrin	1.7
Glucoraphanin	1.6
Neoglucobrassicin	1.2
Glucobrassicin	1.2

In this section, the results of the glucosinolates observed in the AGDH plant lines and comparisons with their parental lines, are presented below:

2.5.8.1 **Desulfoglucoraphanin**

Desulfoglucoraphanin is an aliphatic glucosinolate which was the first desulfated glucosinolate eluting in the chromatograms of 36 AGDH plant lines and in the parental line GD33DH, at approximately 8.5 min (Figure 24, A).

Desulfoglucoraphanin was identified in the *MS* spectra with *m/z* value of 380.0 and 358.0, corresponding to the ion $[M+Na]^+$ and $[M+H]^+$, respectively (Figure 24, B). The expected characteristic fragment molecule with *m/z* 196.0 was observed in the *MS/MS* spectrum (Figure 24, C).

The highest *relative* concentration of glucoraphanin was found in the plant line AGDH1058 while the lowest *relative* concentration was found in the plant line AGDH1060 at 10.07 and 0.25, µmoles/ g of dry leaf material, respectively. For the other 34 AGDH plant lines, they contained glucoraphanin at *relative* concentrations in the range of 7.72-0.28 µmoles/ g of dry leaf material (Figure 25). The parental line GD33DH contained glucoraphanin at a *relative* concentration of 2.11 µmoles/ g of dry leaf material, while in the parental line A12DH we were unable to detect a peak corresponding to glucoraphanin.

Figure 24 Desulfoglucoraphanin expressed in AGDH1058 plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 8.5 min, B: in the *MS* spectrum desulfoglucoraphanin was detected with *m/z* of 380.0 and 358.0 corresponding to the ions [M+Na] $^+$ and [M+H] $^+$ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfoglucoraphanin with *m/z* 196.0, was observed.

Figure 25 The average relative concentration of glucoraphanin in 36 AGDH plant lines, red bar presenting the parental line GD33DH, error bars were calculated for three technical replicates

2.5.8.2 **Desulfoprogoitrin**

Desulfoprogoitrin is the second aliphatic glucosinolate which eluted in the chromatograms of 33 AGDH plant lines at approximately 8.7 min (Figure 26, A), but it was not observed in either of the parental lines analysed in this study suggesting that the parental lines may lack different genes within the biosynthetic pathway i.e. there is complementation in some of the progeny. However, this glucosinolate was previously detected in the preliminarly analysis experiments as being synthesised by the parental line A12DH and gluconapin, the precursor of progoitrin biosynthesis, was observed in this study being synthesised by A12DH. Therefore, the absence of progoitrin in the parental line A12DH in this experiment is more likely to be attributed to environmental effects, which may suppress the transaction of the gene involved in converting gluconapin into progoitrin during the side chain modification stage. Apparently, this gene was functional in the AGDH plant lines in the conditions of the preliminary experiment which synthesised progoitrin.

Desulfoprogoitrin was identified in the *MS* spectra with an *m/z* value of 330.0 and 310.0, corresponding to [M+Na]$^+$ and [M+H]$^+$ ions, respectively (Figure 26, B). The expected characteristic fragment ion with *m/z* 148.0 was observed in the *MS/MS* spectrum (Figure 26, C).

The highest *relative* concentration of progoitrin was found in the plant line AGDH4035 while the lowest *relative* concentration was found in the plant line AGDH2221 at 78.58 and 1.25, μmoles/ g of dry leaf material, respectively (Figure 27). For the other 31 AGDH plant lines, they were expressing progoitrin at *relative* concentrations in the range of 65.55-3.32 μmoles/ g of dry leaf material.
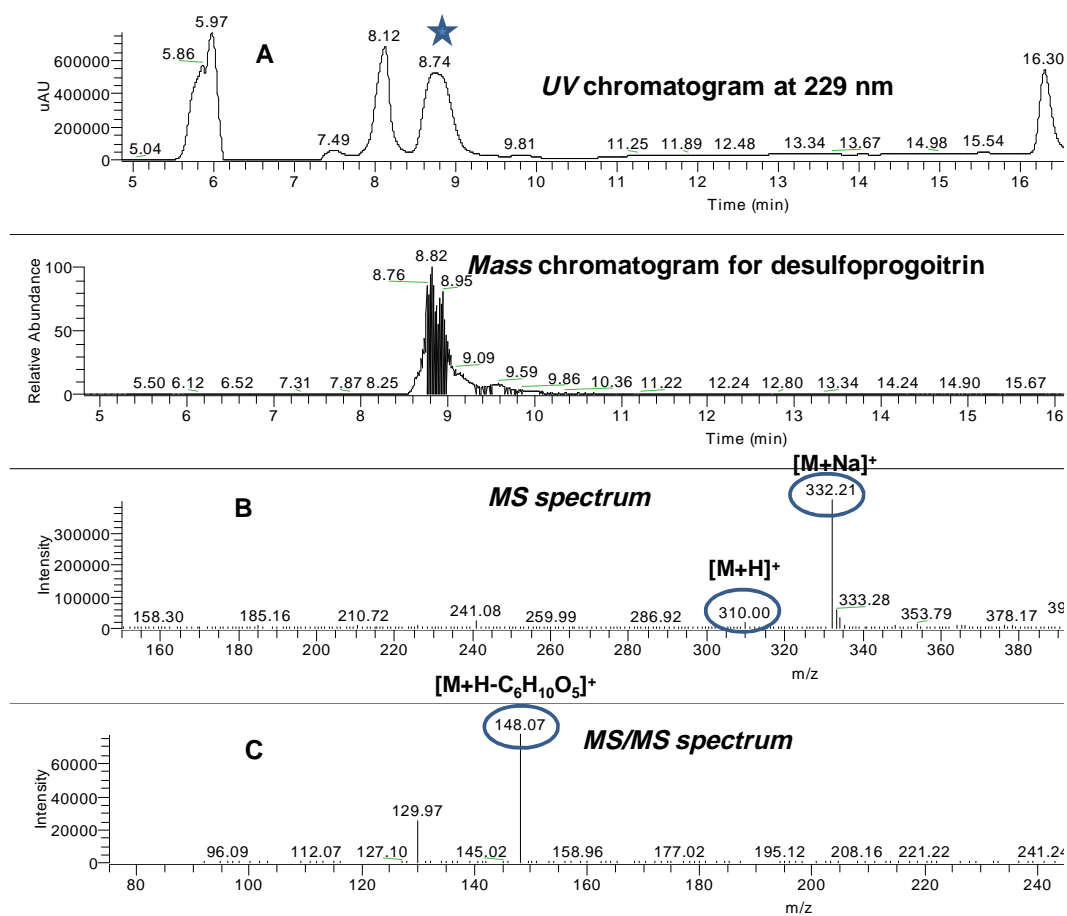
Figure 26 Desulfoprogoitrin expressed in AGDH3088 plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 8.7 min, B: in the *MS* spectrum desulfoprogoitrin was detected with *m/z* of 332.0 and 310.0 corresponding to the ions [M+Na] $^+$ and [M+H] $^+$ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfoprogoitrin with *m/z* 148.0, was observed.
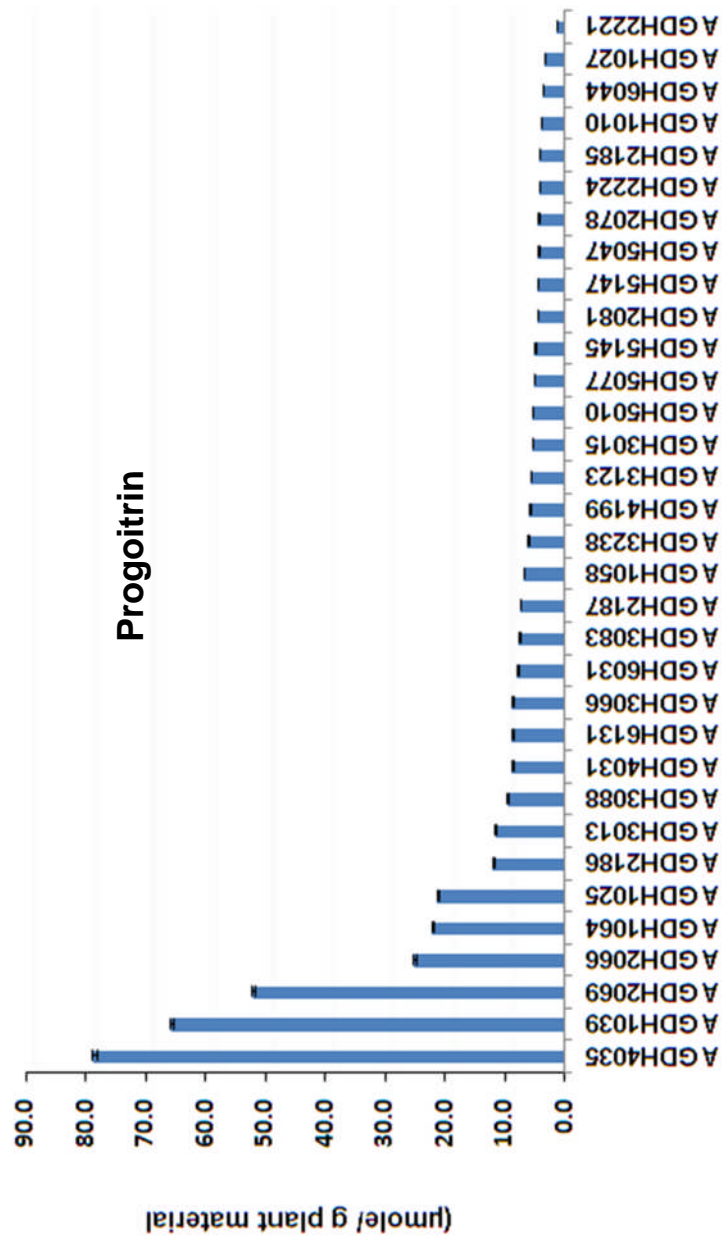
Figure 27 The average relative concentration of progoitrin in 33 AGDH plant lines, error bars were calculated for three technical replicates, note that progoitrin was in the parental lines below the detection level

2.5.8.3 **Desulfosinigrin**

The third desulfoglucosinolate eluting in the chromatograms of the 28 AGDH plant lines and in the parental plant line A12DH, was the aliphatic glucosinolate desulfosinigrin, at approximately 11.2 min (Figure 28, A).

Desulfosinigrin was identified in the *MS* spectra with *m/z* value of 302.0 and 280.0, corresponding to the ions [M+Na]$^{+}$ and [M+H]$^{+}$, respectively (Figure 28, B). The expected characteristic fragment ion with *m/z* 118.0 was observed in the *MS/MS* spectrum (Figure 28, C).

The highest *relative* concentration of sinigrin was found in the plant line AGDH1039 while the lowest *relative* concentration was found in the plant line AGDH2221 at 15.34 and 0.31, µmoles/ g of dry leaf material; respectively. For the other 26 AGDH plant lines, they were expressing sinigrin at *relative* concentrations in the range of 5.15-0.49 µmoles/ g of dry leaf material (Figure 29).

The parental line A12DH was expressing sinigrin at a *relative* concentration of 1.58 µmoles/ g of dry leaf material, while in the parental line GD33DH, it was not possible to detect peak corresponding to sinigrin.
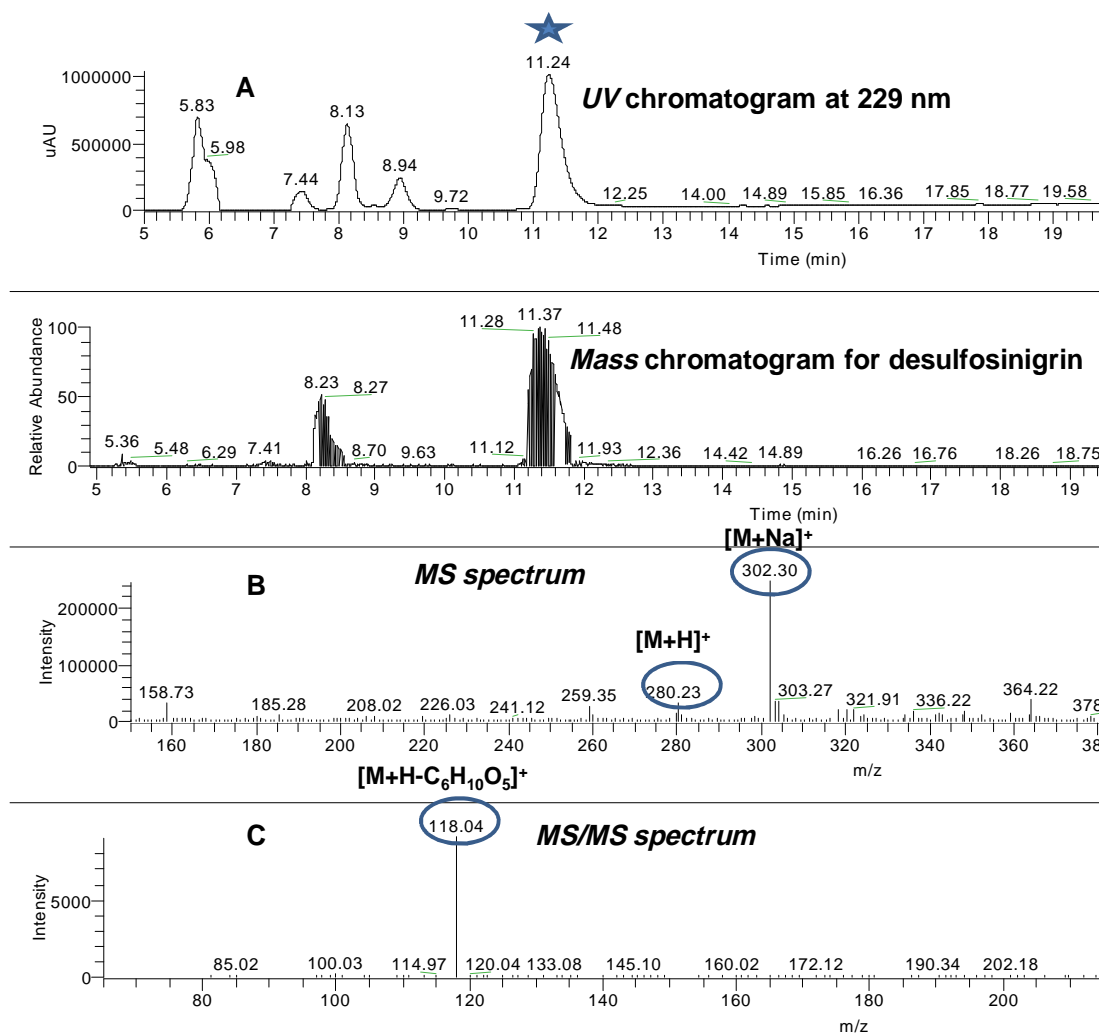
Figure 28 Desulfosinigrin expressed in AGDH1039 plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 11.2 min, B: in the *MS* spectrum desulfosinigrin was detected with *m/z* of 302.0 and 280.0 corresponding to the ions [M+Na] $^+$ and [M+H] $^+$ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfosinigrin with *m/z* 118.0, was observed.
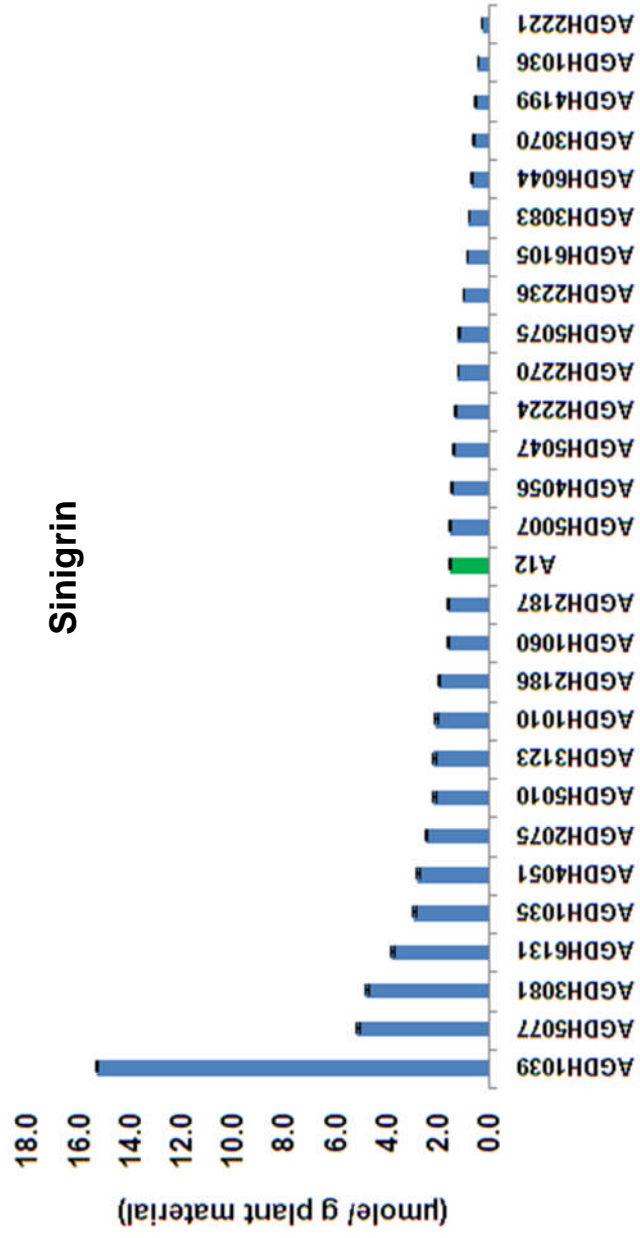
Sinigrin

Figure 29 The average relative concentration of sinigrin in 28 AGDH plant lines, green bar presenting the parental line A12DH, error bars were calculated for three technical replicates

93

2.5.8.4 **Desulfogluconapin**

The fourth aliphatic desulfated glucosinolate eluting in the chromatogram at approximately 16.3 min was desulfogluconapin (Figure 30, A), observed in 45 AGDH plant lines and in the parental line A12DH.

Desulfogluconapin was identified in the *MS* spectra with *m/z* values of 316.0 and 294.0, corresponding to the ions [M+Na]$^+$ and [M+H]$^+$, respectively (Figure 30, B). The expected characteristic fragment ion with *m/z* 132.0 was observed in the *MS/MS* spectrum (Figure 30, C).

The highest *relative* concentration of gluconapin was found in the plant line AGDH1039 while the lowest *relative* concentration was found in the plant line AGDH4199 at 7.17 and 0.11, µmoles/ g of dry leaf material; respectively. For the other 43 AGDH plant lines, they expressed gluconapin at *relative* concentrations in the range of 4.65-0.26 µmoles/ g of dry leaf material (Figure 31). The parental line A12DHd was expressing gluconapin at *relative* concentrations of 2.09 µmoles/ g of dry leaves material, while in the parental line GD33DH, we were unable to detect a peak corresponding to gluconapin.
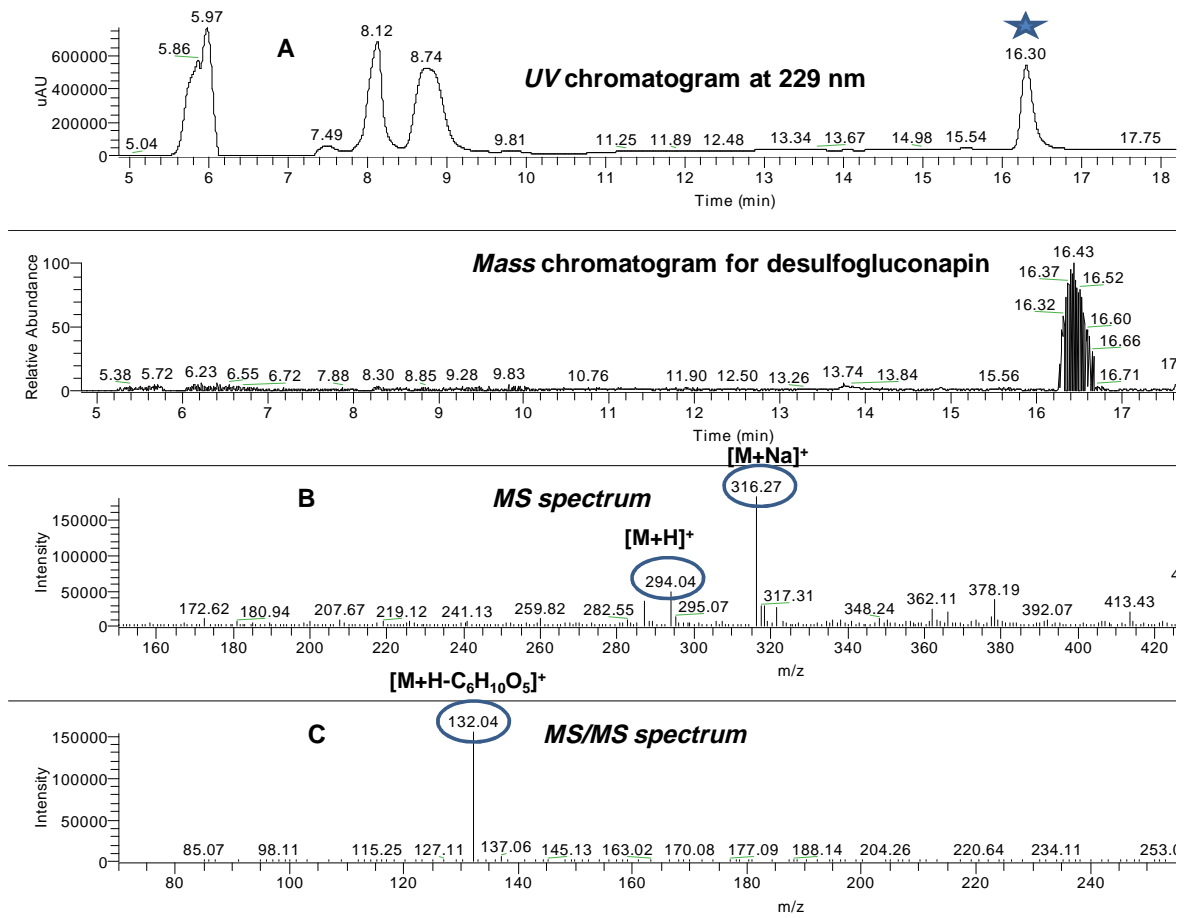
Figure 30 Desulfogluconapin expressed in AGDH3088 plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 16.3 min, B: in the *MS* spectrum desulfogluconapin was detected with *m/z* of 316.0 and 294.0 corresponding to the ions [M+Na] $^{+}$ and [M+H] $^{+}$ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfogluconapin with *m/z* 132.0, was observed.
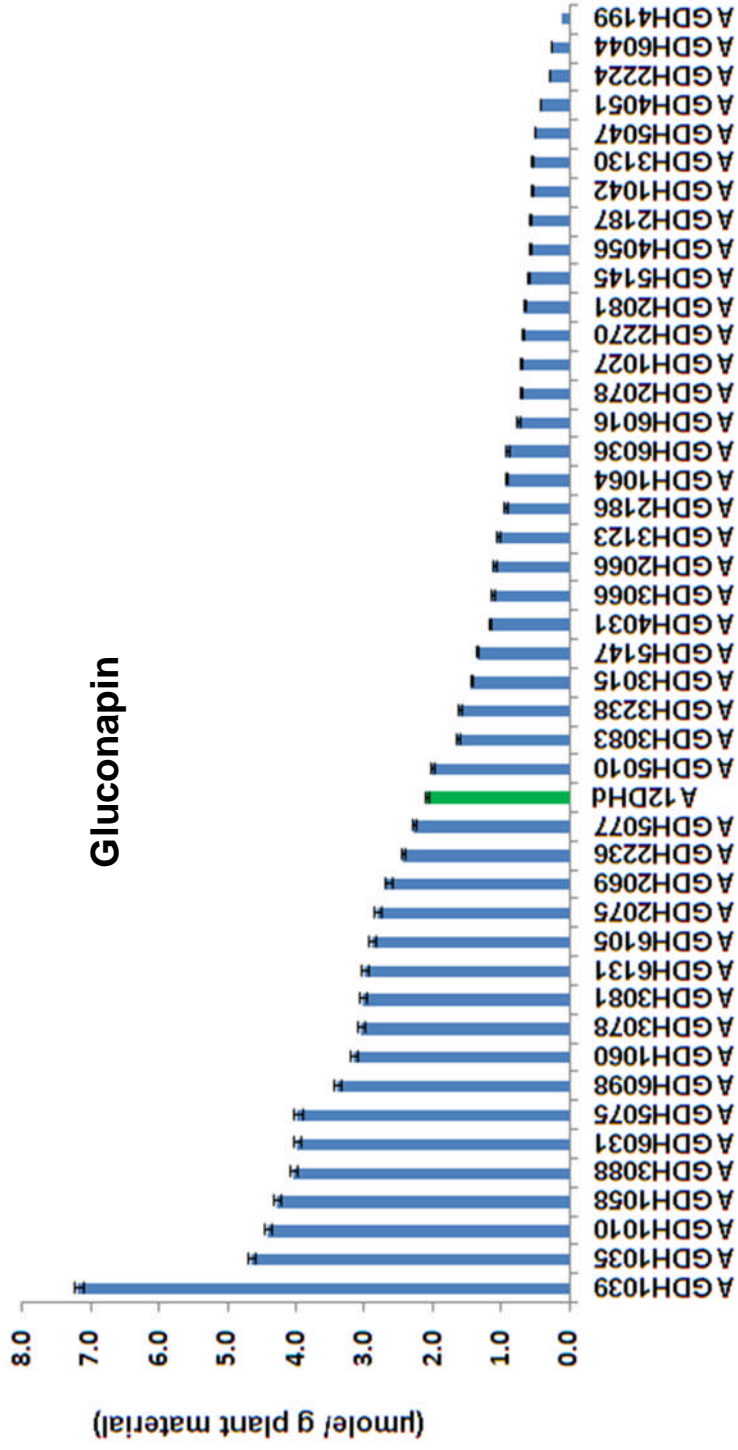
Figure 31 The average relative concentration of gluconapin in 45 AGDH plant lines, green bar presenting the parental line A12DH, error bars were calculated for three technical replicates

2.5.8.5 **Desulfoglucobrassicin**

The fifth desulfated glucosinolate eluting in the chromatogram, at approximately 22.2 min, was the first indolic desulfated glucosinolate; desulfoglucobrassicin (Figure 32, A), which was observed in 85 AGDH plant lines as well as in the parental lines A12DH and GD33DH.

Desulfoglucobrassicin was identified in the *MS* spectrum with *m/z* values of 391.0 and 366.0, corresponding to the ions $[M+Na]^+$ and $[M+H]^+$, respectively (Figure 32, B). The expected characteristic fragment ion with *m/z* 207.0 was observed in the *MS/MS* spectrum (Figure 32, C).

The highest *relative* concentration of glucobrassicin was found in the plant line AGDH3130 at 2.51 µmoles/ g of dry leaf material, while the lowest *relative* concentration was found in the plant lines AGDH5012 and AGDH 3070 at 0.15 µmoles/ g of dry leaves material. For the other 85 AGDH plant lines, they expressed glucobrassicin at a *relative* concentration in the range of 1.99-0.29 µmoles/ g of dry leaf material (Figure 33). The parental lines A12DHd and GDDH33 were expressing glucobrassicin at a *relative* concentration of 0.44 and 0.94 µmoles/ g of dry leaf material, respectively.
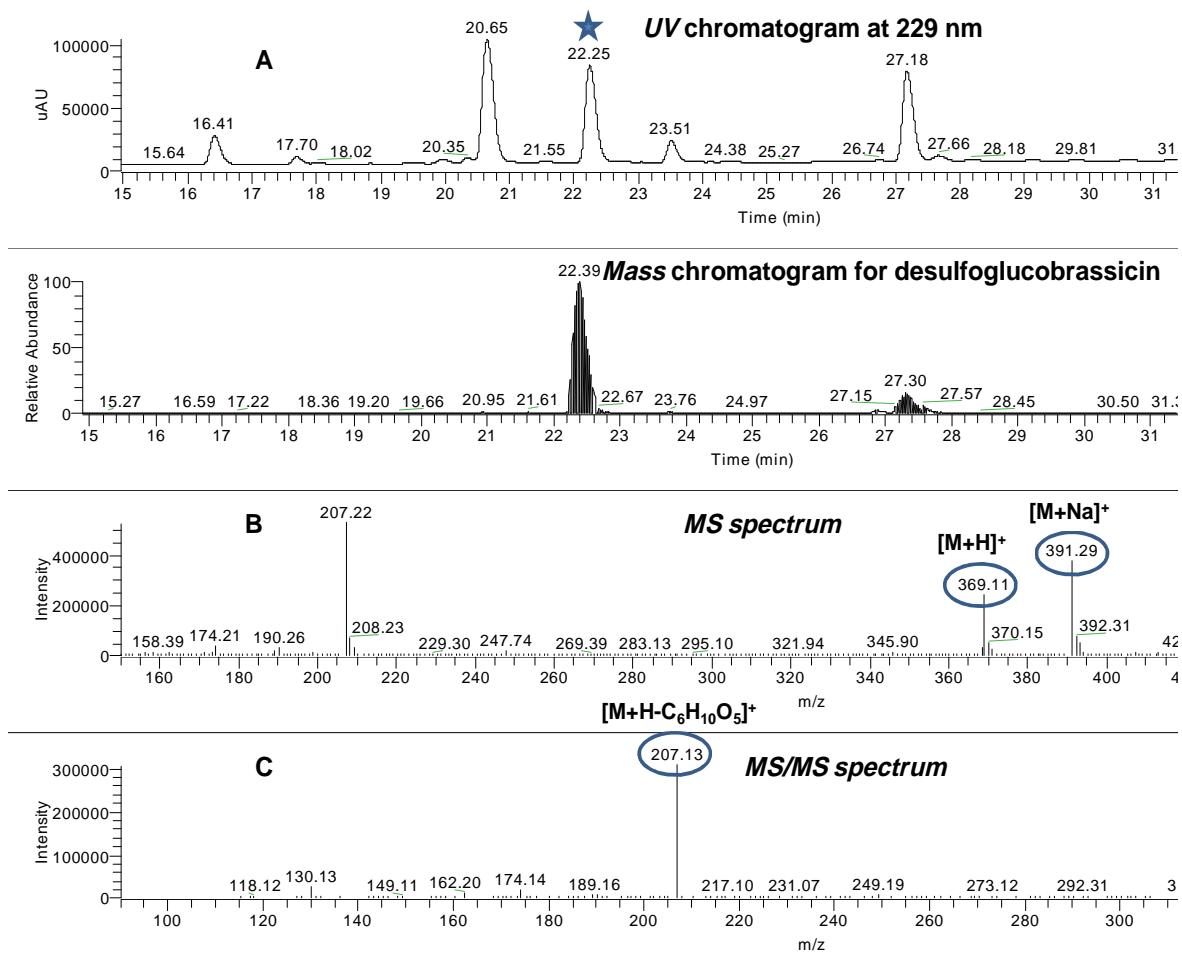
Figure 32 Desulfoglucobrassicin expressed in AGDH5010 plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 22.2 min, B: in the *MS* spectrum desulfoglucobrassicin was detected with *m/z* of 391.0 and 369.0 corresponding to the ions [M+Na] $^+$ and [M+H] $^+$ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfoglucobrassicin with *m/z* 207.0, was observed.
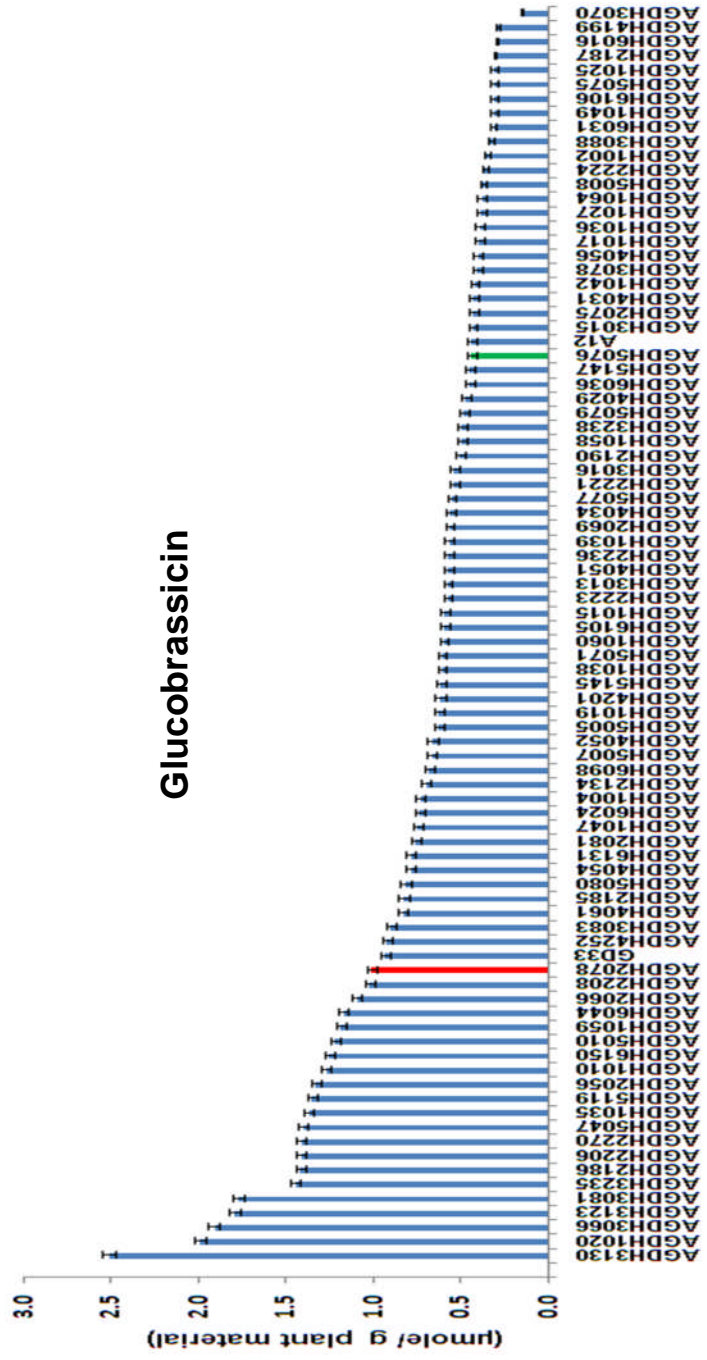
Figure 33 The average relative concentration of glucobrassicin in 87 AGDH plant lines, red and green bars presenting the parental lines GD33DH and A12DH respectively, error bars were calculated for three technical replicates

2.5.8.6 Desulfo-4-methoxyglucobrassicin

The sixth desulfated glucosinolate eluting in the chromatogram at approximately 23.5 min, was the indolic glucosinolate desulfo-4-methoxyglucobrassicin (Figure 34, A); it was observed in all 89 AGDH plant lines as well as in the parental lines A12DH and GD33DH.

Desulfo-4-methoxyglucobrassicin was identified in the *MS* spectra with *m/z* values of 421.0 and 399.0, corresponding to the ions [M+Na]$^+$ and [M+H]$^+$, respectively (Figure 34, B). The expected characteristic fragment ion with *m/z* 237.0 and the structure specific fragment ions with *m/z* 160.0 (corresponding to the fragment ion [R]$^+$), were observed in the *MS/MS* spectrum (Figure 34, C).

The highest *relative* concentration of 4-methoxyglucobrassicin was found in the plant line AGDH5081 at 3.01 µmoles/ g of dry leaf material, while the lowest *relative* concentration was found in the plant lines AGDH4034 at 0.03, µmoles/ g of dry leaf material. For the other 87 AGDH plant lines, they expressed 4-methoxyglucobrassicin at *relative* concentrations in the range of 2.98-0.04 µmoles/ g of dry leaf material (Figure 35). The parental lines A12DHd and GDDH33 were expressing 4-methoxyglucobrassicin at *relative* concentrations of 0.63 and 1.14 µmoles/ g of dry leaf material, respectively.
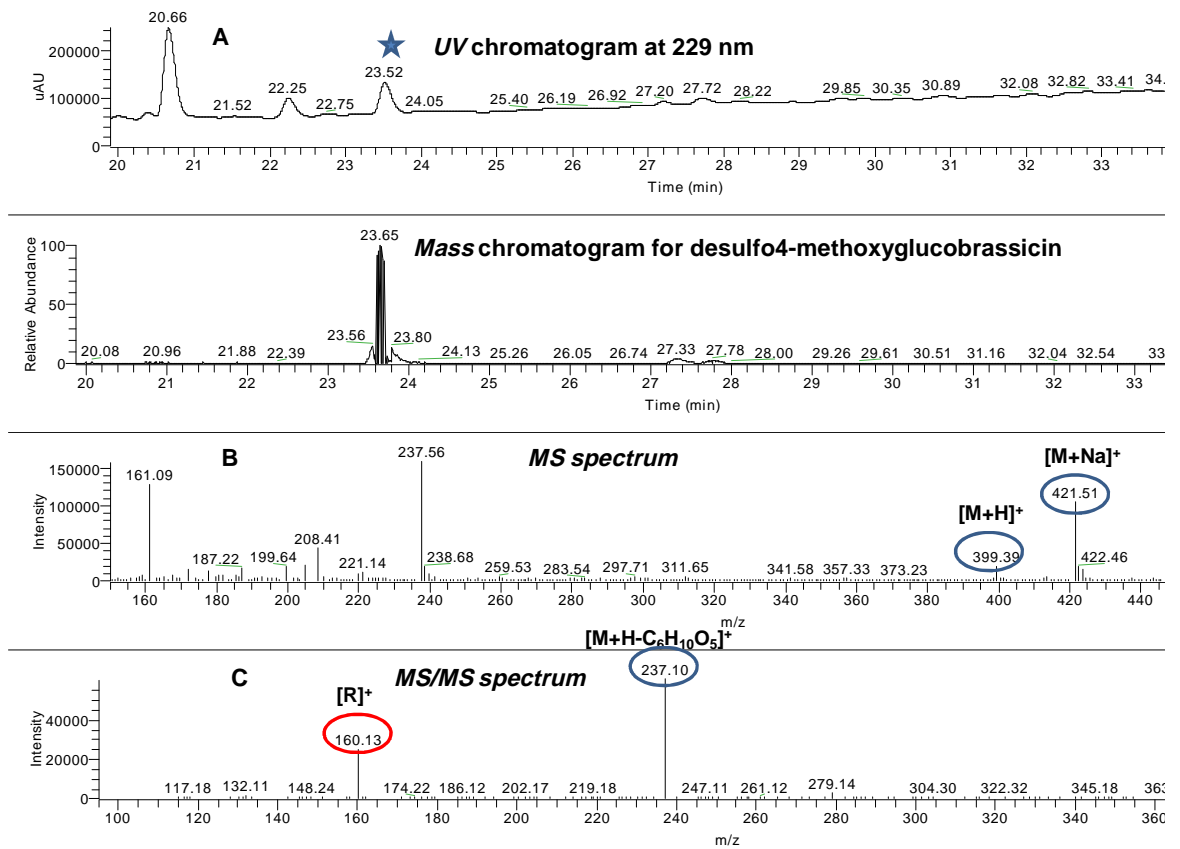
Figure 34 Desulfo-4-methoxyglucobrassicin expressed in A12DHd plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 23.5 min, B: in the *MS* spectrum desulfo-4-methoxyglucobrassicin was detected with *m/z* of 421.0 and 399.0 corresponding to the ions [M+Na]⁺ and [M+H]⁺ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfo-4-methoxyglucobrassicin with *m/z* 207.0, and the structure specific fragment ions with *m/z* 160.0 corresponding to the fragment ion [R]⁺, were observed.
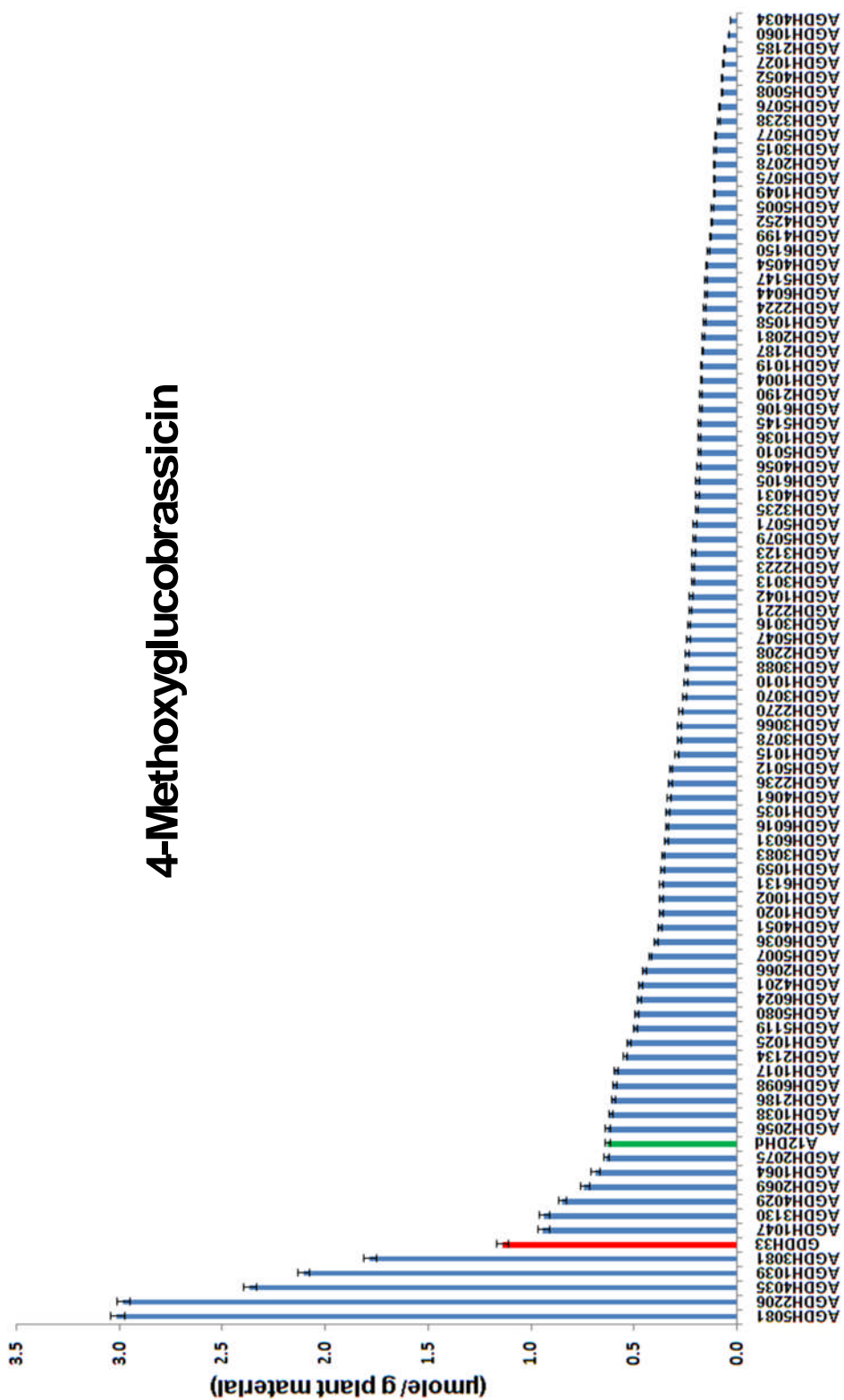
Figure 35 The average relative concentration of 4-methoxyglucobrassicin in 89 AGDH plant lines, red and green bars presenting the parental lines GD33DH and A12DH respectively, error bars were calculated for three technical replicates

2.5.8.7 **Desulfoneoglucobrassicin**

The last desulfated glucosinolate eluting in the chromatogram was the indolic glucosinolate desulfoneoglucobrassicin, at approximately 27.2 min (Figure 36, A), and it was observed in 41 AGDH plant lines as well as in the parental line GD33DH. Desulfoneoglucobrassicin was identified in the *MS* spectra with *m/z* values of 421.0 and 399.0, corresponding to the ions [M+Na] $^+$ and [M+H] $^+$, respectively (Figure 36, B). The expected characteristic fragment ion with *m/z* 237.0 and the structure specific fragments ions with *m/z* 205.0, 177.0 and 130.0 which were corresponding to the fragment ions [RCNOH+2H] $^+$, [ROH] $^+$ and [R-CH3O+H] $^+$ respectively, were observed in the *MS/MS* spectrum (Figure 36, C).

The highest *relative* concentration of neoglucobrassicin was found in the plant line AGDH3123 at 0.90 µmoles/ g of dry leaf material, while the lowest *relative* concentration was found in the plant lines AGDH1042 at 0.05, µmoles/ g of dry leaf material. For the other 39 AGDH plant lines, they expressed neoglucobrassicin at *relative* concentrations in the range of 0.86-0.06 µmoles/ g of dry leaf material (Figure 37). The parental line GD33DH expressed neoglucobrassicin at a *relative* concentration of 0.13 µmoles/ g of dry leaf material, while in the parental line A12DH; we were unable to detect a peak corresponding to neoglucobrassicin.
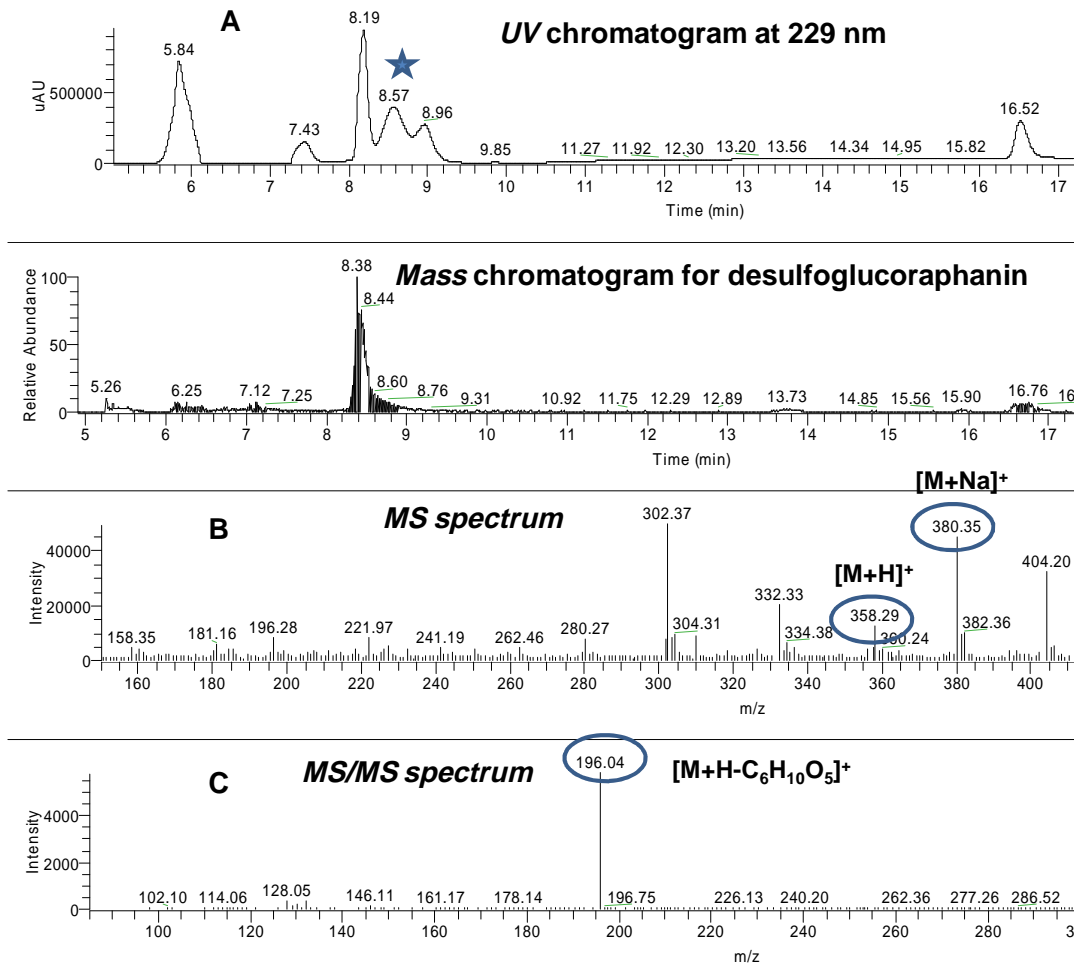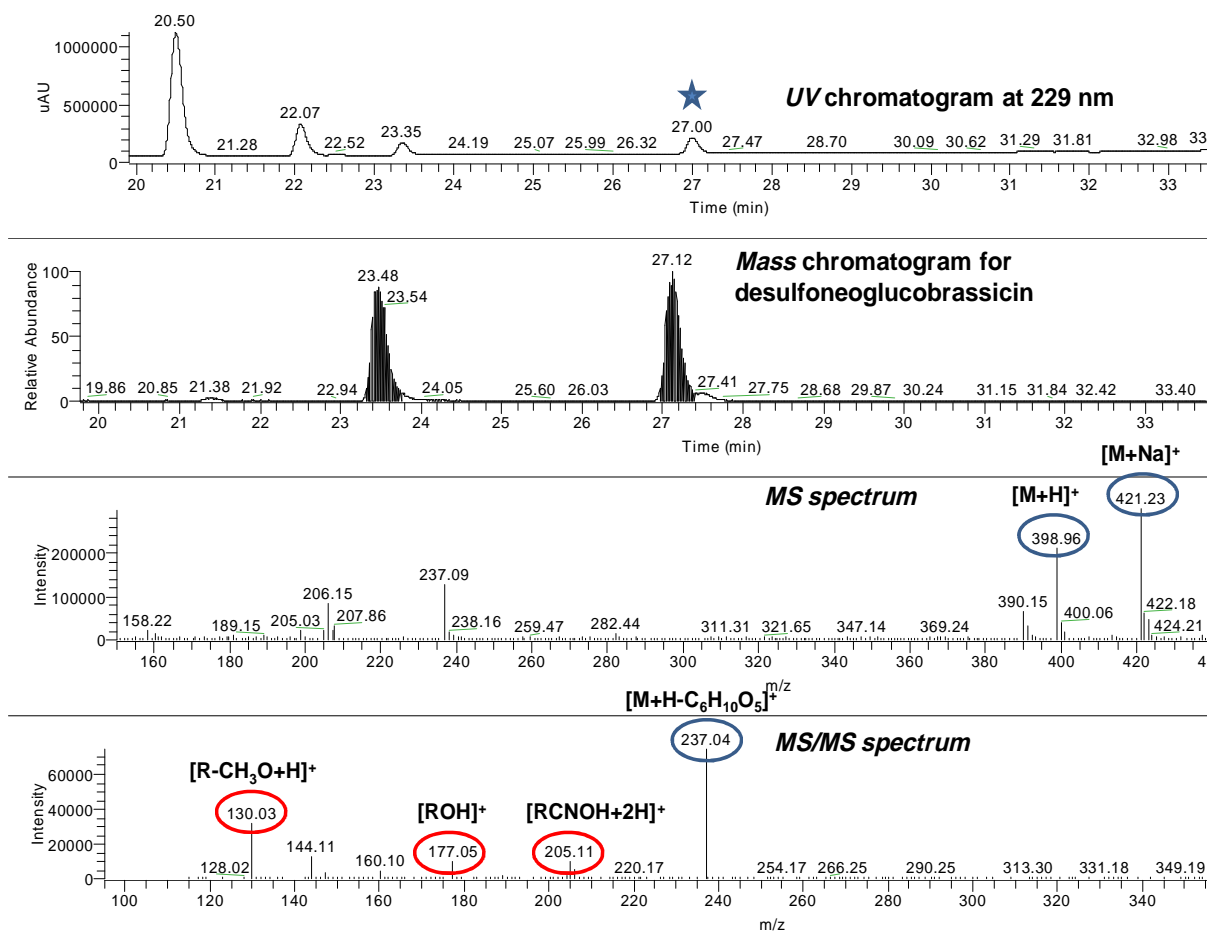
Figure 36 Desulfoneoglucobrassicin expressed in AGDH2190 plant extract, A: eluting in the *UV* and mass chromatogram with RT approximately 27.0 min, B: in the *MS* spectrum desulfoneoglucobrassicin was detected with *m/z* of 421.0 and 399.0 corresponding to the ions [M+Na]$^+$ and [M+H]$^+$ respectively, C: in the *MS/MS* spectrum the typical fragment ion for desulfoneoglucobrassicin with *m/z* 207.0, and the structure specific fragment ions with *m/z* 130, 177 and 205 corresponding to [R-CH$_3$O+H]$^+$, [ROH]$^+$ and [RCNOH+2H]$^+$ respectively, were observed.
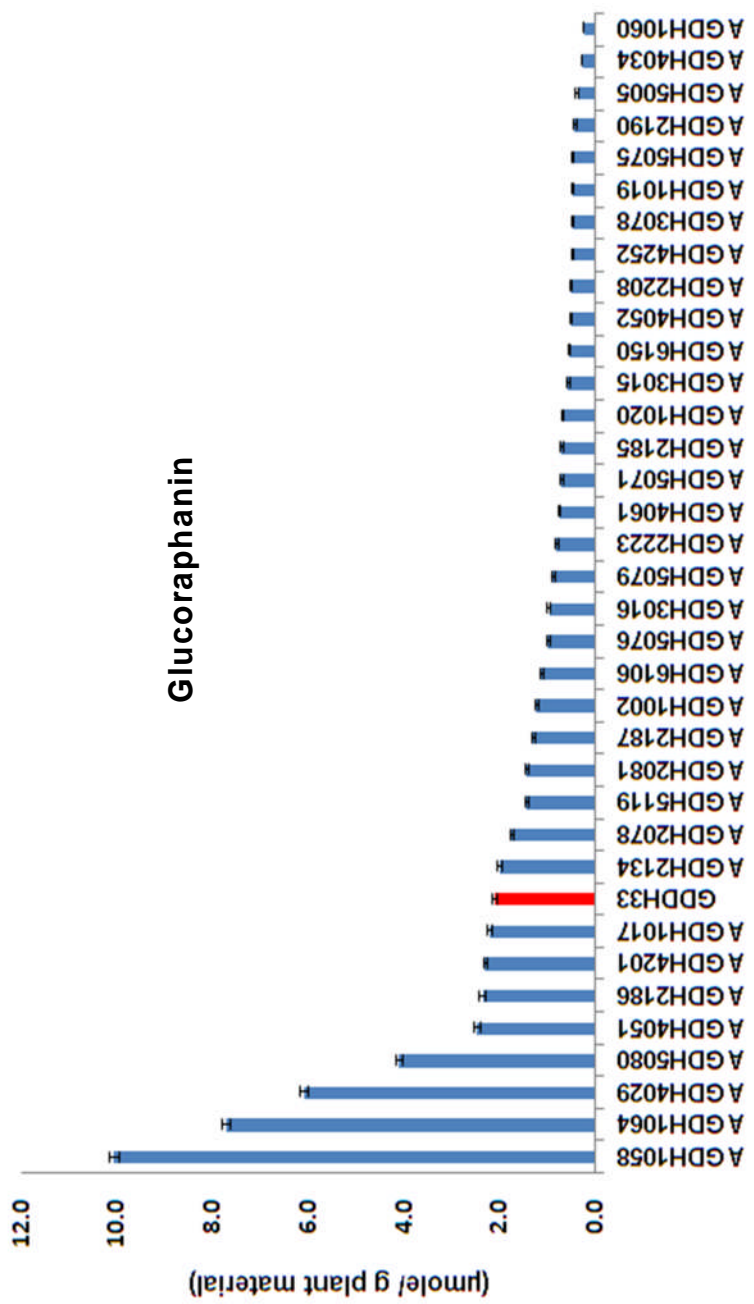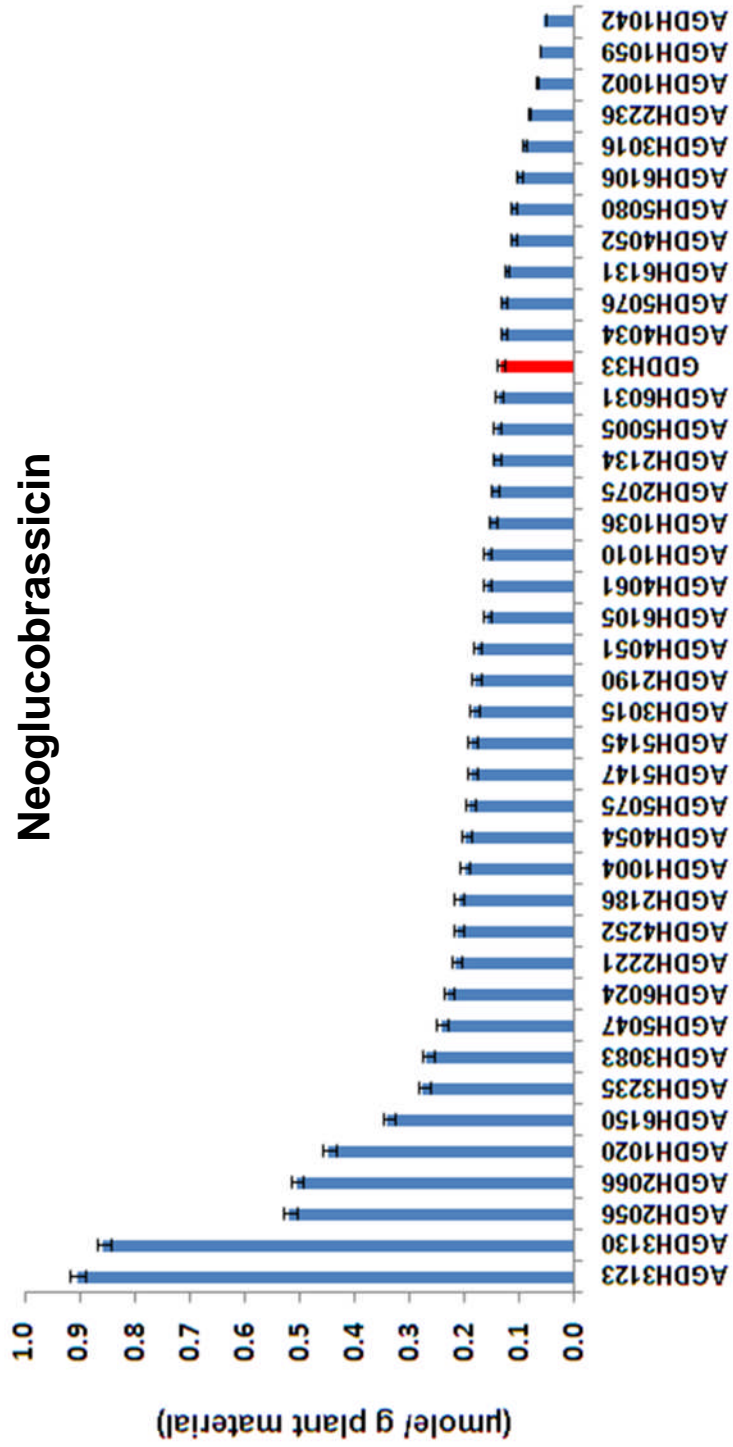
Figure 37 The average relative concentration of neoglucobrassicin in 41 AGDH plant lines, red bar presenting the parental line GD33DH, error bars were calculated for three technical replicates

2.5.9 **Diversification of glucosinolates among species of *Brassicaceae***

The *Brassicaceae* family consists of vegetable crops of biological, economical and agricultural importance attributed to their phytochemical content of which, the active compounds glucosinolates are of great interest in this study. However, their activity is largely determined by their content and, therefore the *Brassicaceae* family has been widely investigated for their glucosinolate composition (Francisco et al., 2009; Meyer and Adam, 2008).

 In the following section, the range and average concentrations for the individual glucosinolates, the sum of aliphatic glucosinolates, and the sum of indolic glucosinolates in the 89 AGDH plant materials, will be discussed. These concentrations, converted to percentage of their contribution to the total glucosinolate content, and to the chemical class to which they belong, are shown in Table 16.

Table 16 Variation in the glucosinolate range and average concentrations expressed in (µmole/ g*) calculated from AGDH plant lines. The % of their contribution to the total glucosinolate content and to the chemical class to which belong is shown. * Dry plant material

Trait	Range (µmole/ g*)	Average (µmole/ g*)	% Total glucosinolates	% Aliphatic glucosinolates	% Indolic glucosinolates
Sum aliphatic glucosinolates	0.28-88.06	8.31	85.0	100	---
Sum indolic Glucosinolates	0.41-4.40	1.26	15.0	---	100
Glucoraphanin	0.28-7.72	1.72	31.5	39.3	---
Progoitrin	3.32-65.55	12.95	73.2	78.6	---
Sinigrin	0.49-5.15	2.31	21.7	24.7	---
Gluconapin	0.26-4.65	1.92	15.9	17.7	---
Glucobrassicin	0.29-1.99	0.75	9.9	---	60.8
4-Methoxyglucobrassicicn	0.04-2.98	0.43	5.2	---	34.2
Neoglucobrassicin	0.06-0.86	0.22	4.0	---	16.1

In all plant lines greater variation in the concentration was observed within the aliphatic glucosinolates, both as a group and as individuals, than within the indolic glucosinolates, as indicated by the wider range of measured concentrations from the plant lines synthesising glucosinolates in the studied plant population (as shown in Table 16). The average aliphatic glucosinolate concentration across all analysed plant lines was 8.31 µmoles/ g dry leaf material; representing 85.0% of the total glucosinolate content. Progoitrin was the most abundant aliphatic glucosinolate and was responsible for 78.6% of the aliphatic glucosinolate content observed in 33 AGDH lines, followed by glucoraphanin (39.3%) by 36 AGDH lines, sinigrin (24.7%) by 28 AGDH lines and gluconapin (17.7%) by 45 AGDH lines.

While the average indolic glucosinolate content was 1.26 µmoles/ g dry leaf material, representing only 15.0% of the total glucosinolates content of the AGDH population. The most abundant indolic glucosinolate was glucobrassicin contributing with 60.8%

observed in 85 AGDH lines, followed by 4-methoxyglucobrassicin (34.2%) by 89 AGDH lines and neoglucobrassicin (16.1%) by 41 AGDH lines of the total indolic glucosinolate content.

A comparison of the relative glucosinolate content determined from the AGDH plant lines in our study with kale, broccoli and cauliflower values in the literature revealed interesting variations in the glucosinolate profiles synthesied by thes vegetable crops as shown in Table 17.

Table 17 Variations in the relative glucosinolate contents in AGDH, kale, broccoli and cauliflower. [a] (Cartea et al., 2008; Velasco et al., 2007), [C] (Schonhof et al., 2004) and [d] (Volden et al., 2009)

Trait	% of total glucosinolate			
	AGDH	Kale	Broccoli [c]	Cauliflower [d]
Total aliphatic	85.0	70.0 [a] 52.3 [b]	>93.4	57.3
Total indolic	15.0	30.0 [a] 46.4 [b]	5.9	42.7
Glucoraphanin	31.5	0.50 [a]	26.5	2.3
Progoitrin	73.2	2.7 [a]	3.7	37.0
Sinigrin	21.7	35.9 [a]	0.9	5.9
Gluconapin	15.9	---	57.7	---
Glucobrassicin	9.9	89.1 [a] 40.8 [b]	3.4	29.1
4-Methoxyglucobrassicin	5.2	---	0.4	5.3
Neoglucobrassicin	4.0	8.8 [a] 5.6 [b]	1.4	2.7

Our results are in good agreement with the quantitative analysis of desulfated glucosinolates concentrations presented by (Cartea et al., 2008; Velasco et al., 2007), who reported the profiles from the edible parts of vegetable kales (leaves and flower buds), showing that aliphatic glucosinolates were dominant. Cartea et al., (2008)

reported 70.0% of the glucosinolate was aliphatic, with the indolic glucosinolates comprising the remaining 30.0%, whilst (Cartea et al., 2008; Velasco et al., 2007) reported 52.3% aliphatic content. High levels of aliphatic glucosinolates were also reported by (Schonhof et al., 2004; Volden et al., 2009), in Chinese broccoli with over 90% of the total glucosinolate content. Cauliflower aliphatic glucosinolate content has been observed at 57.3% (Volden et al., 2009).

In this study, progoitrin was the major glucosinolate observed, representing 73.2% of the total glucosinolate content, followed by glucoraphanin (31.5%), sinigrin (21.7%), and gluconapin (15.9%). This was in agreement with (Volden et al., 2009), who reported the relative content of intact glucosinolates in the florets of five cauliflower cultivars and found that the most abundant aliphatic glucosinolate was progoitrin (37.0% fresh material), followed by sinigrin (5.9%) and glucoraphanin (2.3%). In contrast studies on desulfated glucosinolate content, (Schonhof et al., 2004) found gluconapin to be the most abundant in Chinese broccoli plant material representing (57.7%) of the total glucosinolate content, followed by glucoraphanin (26.5%), progoitrin (3.7%) and sinigrin (0.9%). In contrast, in kale (Cartea et al., 2008; Velasco et al., 2007), desulfated glucosinolate content showed sinigrin present at the highest abundance with (35.9%), followed by progoitrin (2.7%) and glucoraphanin (0.5%).

In general, the levels of indolic glucosinolates in our AGDH population agreed with other published studies (Cartea et al., 2008; Kushad et al., 1999; Schonhof et al., 2004; Song and Thornalley, 2007; Velasco et al., 2007; Volden et al., 2009). In this study, glucobrassicin was present at 9.9%, followed by 4-methoxyglucobrassicin (5.2%) and neoglucobrassicin (4.0%). These results are in agreement with (Schonhof et al., 2004), who studied the desulfated glucosinolates in Chinese broccoli, reporting

the highest proportion within the indolic glucosinolates to be glucobrassicin at 3.4% of the total glucosinolate content, followed by neoglucobrassicin (1.4%) and 4-methoxyglucobrassicin (0.4%).

Higher levels of glucobrassicin were observed in kale varieties by (Cartea et al., 2008) comprising 89.1% of the total glucosinolate content, and neoglucobrassicin (8.8%). The same results were reported by (Velasco et al., 2007), who studied glucosinolate content in kale and found the highest proportion for glucobrassicin with 40.8% of the total, while neoglucobrassicin was only 5.6%, with no observations reported for 4-methoxyglucobrassicin in kale. While investigations of intact glucosinolate content in cauliflower (Volden et al., 2009), showed moderate content of glucobrassicin (29.1%), followed by 4-methoxyglucobrassicin (5.3%) and neoglucobrassicin (2.7%) of the total glucosinolate content.

The reason for these differences might be most probably linked to cultivar differences, growing conditions (temperature, day light, and soil nutrients), water content in the plant material used for analysis, methods of extraction and method of quantification measurements.

2.5.10 Selected AGDH plant lines of biological interest

Investigation of the glucosinolate profiles in all plant lines analysed in this study for their total and individual glucosinolate content, revealed some interesting plant lines that have potential biological applications including medical, agricultural, economic and consumer acceptance.

The parental plant line GDDH33 expressed total glucosinolates at a relative concentration of 4.32 µmole/ g dry plant material, composed of glucoraphanin

(48.8%), glucobrassicin (21.7%), 4-methoxyglucobrassicin (25.4%) and neoglucobrassicin (3.0%) as seen in Table 18.

Studies on the health promoting effects of individual products derived from glucoraphanin hydrolysis, the most important of which is sulforane (SF), are the focus of the clinical research to find potential cancer prevention and/ or treatment compounds. In addition, glucoraphanin may help protect from serious chronic diseases affecting the cardiovascular or the nervous system (Jeffery and Araya, 2009). The benefit of producing crops with the highest content of glucoraphanin as the major aliphatic glucosinolate, is that the hydrolysis product SF is not volatile and so, will not affect the flavour or the aroma of the vegetables, which may increase the customer acceptance for healthy vegetables (Traka and Mithen, 2009).

The parental plant line A12DH, was found to express total glucosinolates at approximately the same relative concentration as GD33DH (4.73 µmole/ g dry plant material), with a composition of sinigrin (33.3%), gluconapin (44.2%), glucobrassicin (9.2%) and 4-methoxyglucobrassicin (13.3%). The disadvantage of vegetable crops synthesising a high content of aliphatic glucosinolates containing alkene bonds can be the bitter taste (Schonhof et al., 2004). On the other hand, a vegetable crop with this glucosinolate profile is considered a safe material for feeding animals due to the natural absence of progoitrin, which is toxic to farm animals.

Table 18 Comparison between selected plant lines; showing the relative content for individual glucosinolate as a percentage of the total glucosinolate content in each plant line, the parental lines are indicated by the red and green columns

	GD33DH	A12DH	AGDH1039	AGDH4034	AGDH2206	AGDH4051
Total glucosinolates (μmole/ g plant material)	4.3	4.7	90.7	1.0	4.4	6.8
% Glucoraphanin	48.8	-----	-----	28.0	-----	36.2
% Progoitrin	-----	-----	72.2	-----	-----	-----
% Sinigrin	-----	33.3	16.9	-----	-----	41.1
% Gluconapin	-----	44.2	7.9	-----	-----	6.27
% Glucobrassicin	21.8	9.2	0.6	56.0	32.3	8.3
% 4-Methoxyglucobrassicin	26.4	13.3	2.3	3.0	67.7	5.5
% Neoglucobrassicin	3.0	-----	-----	13.0	-----	2.6

The *highest content of total glucosinolates* observed among all the analysed plant lines in this population, was for the plant line AGDH1039 which expressed glucosinolates at a total concentration of 90.73 µmole/ g dry plant material, composed of progoitrin (72.2%), sinigrin (16.9%), gluconapin (7.9%), 4-methoxyglucobrassicin (2.3%) and glucobrassicin (0.6%). This plant line had *the highest total content of aliphatic glucosinolates* observed with a concentration of 88.0 µmole/ g dry plant material, composed of progoitrin (74.5%), sinigrin (17.4%) and gluconapin (8.1%). The disadvantages of vegetable crops expressing this glucosinolate profile are not only the bitter taste (Schonhof et al., 2004), but also the toxic effect of the hydrolysis product of progoitrin (Cartea and Velasco, 2008).

Within the 89 AGDH plant lines analysed in this study, 14 lines were found not to express aliphatic glucosinolates at detectable levels, AGDH2206, AGDH5081, AGDH2056, AGDH3235, AGDH1074, AGDH1059, AGDH6024, AGDH1038, AGDH4054, AGDH1004, AGDH1015, AGDH5012, AGDH5008, and AGDH1049.

The *lowest content of aliphatic glucosinolates* was found in the plant line AGDH4034 that expressed glucoraphanin at a concentration of (0.28 µmole/ g dry plant material) with a relative concentration of total glucosinolates at (28.0%). Such a vegetable crop would have higher consumer acceptance because of the good taste, but with minimum health benefits due to its low total glucosinolate content.

The plant line AGDH2206 was found to express *the highest total indolic glucosinolate content* within the AGDH population with a concentration of 4.4 µmole/ g dry plant material, composed of 4-methoxyglucobrassicin (67.7%) and glucobrassicin (32.3%), suggested potential anti-fungal defensive activity for this plant material, due to its high content of 4-methoxyglucobrassicin, which is

important for agricultural applications as an organic bio-fumigant (Bednarek et al., 2009). In addition to the dietary benefit of the most important indolic glucosinolates in *Brassica* vegetables, glucobrassicin is known to decrease the risk for breast cancer (Jeffery and Araya, 2009).

Three of the lines; AGDH5012, AGDH5008 and AGDH1049 expressed trace amounts of indolic glucosinolates in the concentration range 0.41-0.47 µmole/ g dry plant material and appeared to have very limited biological interest.

The plant line AG4051 did not express progoitrin at a detectable level, but was found to express a combination of glucoraphanin and sinigrin at the highest content compared to the AGDH plant lines, when progoitrin was not observed. Sinigrin was produced at concentrations of 2.5 µmole/ g dry plant material corresponding to (36.2%) of the total glucosinolates content. Sinigrin is known as a powerful biofumigant, due to the production of the hydrolysis product allyl isothiocyanate, which reduces the attraction of insects (Bellostas et al., 2007b; Hall et al., 2001). Also it is a precursor for isothiocyanate, known for anticancer activity (Cartea and Velasco, 2008; Higdon et al., 2007). Glucoraphanin is regarded as the most important glucosinolate in *Brassica* vegetables for its benefits to health and was expressed in this plant line at a concentration of 2.82 µmole/ g dry plant material corresponding to (41.1%) of the total glucosinolates content. As progoitrin was not observed at detectable levels in this plant line, there does not appear to be a health risk associated with the consumption of such a crop. Therefore, it is considered an important plant lines for developing a crop for a healthy diet in addition to the other possible agricultural applications.

It was not surprising to find that glucosinolates produced by the same biosynthetic pathway (more detailed discussion will be presented in Chapter 3) were expressed together in most plant lines as shown in (Figure 38).
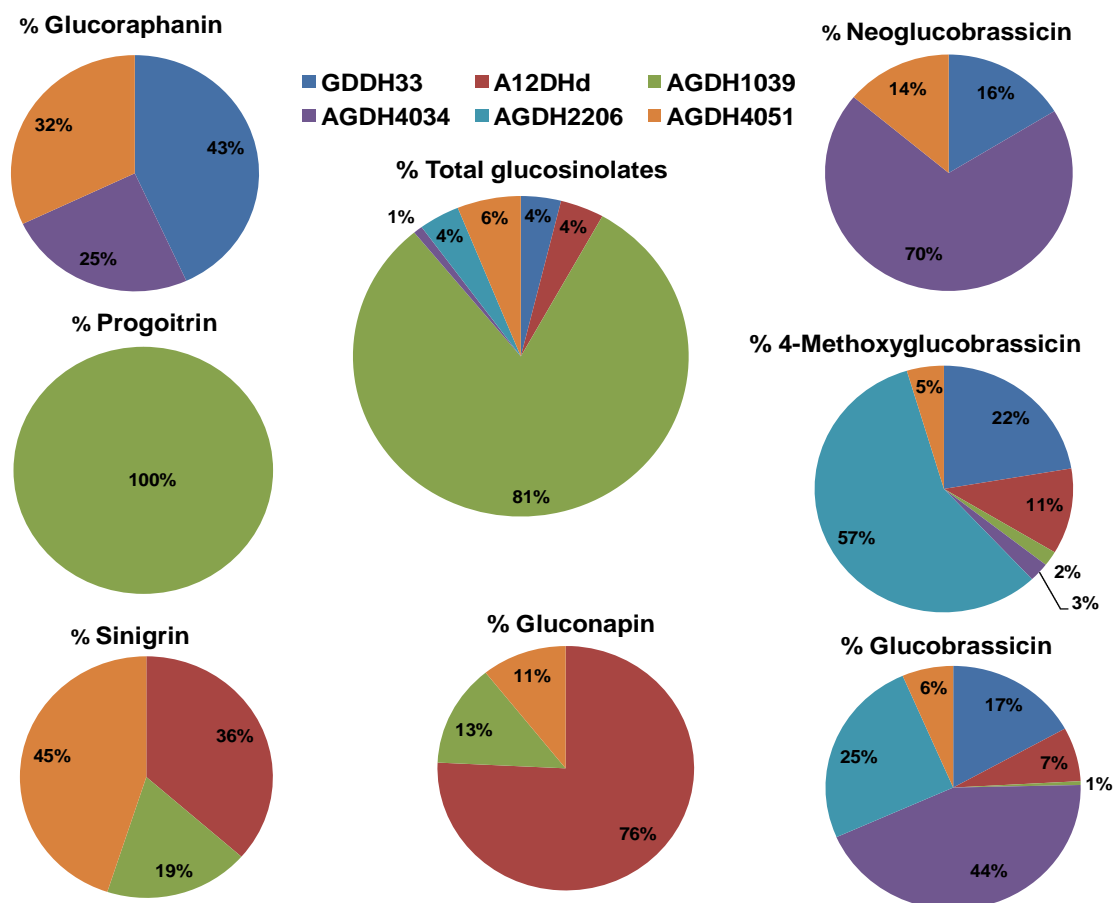
Figure 38 Pie charts showing the total glucosinolates detected across the plant lines; GD33DH, A12DH, AGDH1039, AGDH4034, AGDH2206 and AGDH4051, with the values indicating % of the total glucosinolate and % of the individual glucosinolates observed between the six selected plant lines shown

The aliphatic glucosinolates; sinigrin and gluconapin were expressed by the plant lines; A12DH, AGDH1039 and AGDH4051, while the indolic glucosinolates 4-methoxyglucobrassicin and glucobrassicin were expressed by the plant lines GD33DH, A12DHd, AGDH1039, AGDH4034, AGDH2206 and AGDH4051. In contrast, although glucoraphanin is the precursor of gluconapin biosynthesis; they were observed together in only one plant line AGDH4051, indicating that the plant

115

lines expressing high levels of either of these two glucosinolates expressed the second glucosinolate at undetectable levels. This could be because the number of plant lines analysed in this study was not sufficiently large, or due to the presence of gene(s) controlling their biosynthesis in a qualitative rather than quantitative pattern.

The aliphatic glucosinolate; glucoraphanin and the indolic glucosinolates; neoglucobrassicin, 4-methoxyglucobrassicin and glucobrassicin, were expressed together in GD33DH, AGDH4051 and AGDH4034 plant lines. Although the biosynthetic pathway of aliphatic glucosinolates is independent from the biosynthetic pathway of indolic glucosinolates, they share a common set of enzymes that are involved in the core structure formation of all glucosinolates classes (as described in Chapter 1), which control the total glucosinolates content expressed in plant materials.

The AGDH1039 plant line expressed the highest percentage of total glucosinolates compared to the other plant lines, and was also found to contain high levels of progoitrin compared to the other lines shown in Figure 38. Further discussion of this observation will be presented from genetic point of view in Chapter 3.

In general, the plant lines expressing aliphatic glucosinolates at high levels expressed the indolic glucosinolates at lower levels, and vice versa. This is a known observation of the secondary metabolites in the plant kingdom, where the expression of any of these metabolites can be altered to improve the plant fitness in response to stress (Jones and Firn, 1991; Kliebenstein, 2004).

2.6 Conclusion

- The analysis of intact glucosinolates in this study was unsuccessful presumably due to the high salt content in the plant extract samples. Therefore, the analysis of desulfated glucosinolates was adopted over intact glucosinolate analysis.

- An *HPLC-UV* method was optimized for complete separation of desulfated glucosinolates in the plant extract with high resolution for quantification measurements of glucosinolates.

- *The reproducibility of the desulfation reaction was improved* with the use of an acetate buffer at pH=5.5 for the desulfation step, the use of a shaking incubator during the enzymatic step and the use of the optimal ratio of sulfatase solution concentration to amount of plant material for maximum desulfation reaction.

- An optimized level of internal standard (IS1) was experimentally determined, and used for all subsequent quantitative measurements of glucosinolates.

- I have demonstrated the *first use of a second internal standard (IS2)* to improve the reproducibility of the quantitative measurements.

- *Development of a data dependent MS and MS/MS* based methodology for the identification and characterisation of 13 desulfated glucosinolates.

- The relative concentration of individual desulfated glucosinolates to IS1 was calculated using peak areas in the *UV* chromatogram at 229 nm and the relative response factor (RRF) as described in the standardised procedures (EEC, 1990).

- In total seven glucosinolates were detected in the 89 AGDH plant lines distributed between aliphatic and indolic glucosinolate groups, with different combinations from the parental plants A12DH and GD33DH, and displaying

117

wide qualitative and quantitative variations in their glucosinolate profiles, while no aromatic glucosinolates were detected.

- Aliphatic glucosinolates were predominant over indolic glucosinolates in the 89 AGDH plant lines, whilst progoitrin was found in the highest abundance among the total glucosinolate concentration.

- The observed variations in the glucosinolate profiles of the AGDH plant lines, revealed *the presence of six plant lines expressing glucosinolates with unique qualitative and quantitative contents* of biological importance for medical, agricultural and economical applications.

- The quantitative measurements undertaken can help increase understanding of the biosynthetic pathway of glucosinolates in the studied plant population.

Identifying QTL affecting glucosinolates biosynthesis in

Brassica oleracea

3.1 Introduction

Analysis of quantitative trait loci (QTL) in plant progeny derived from a cross of two parents, which showed significant differences in their trait profiles, involves linkage analysis between a set of markers and phenotypic data, utilising the genetic linkage map of their chromosomes where their molecular markers types and locations are known (Tanksley, 1993; Van Ooijen, 1999). The natural genetic variations observed for the quantitative traits can be exploited using QTL analysis to identify candidate genetic loci or genes that affect metabolite biosynthesis (Kearsey, 1998). Once a significant QTL ($p \leq 0.05$) of a trait has been identified in a population, genes within the QTL confidence interval that affect the trait can be determined by a number of different approaches including comparative genomics at two levels; genetic and physical mapping, and DNA sequencing (Gao et al., 2004). In addition combining gene expression profiles with metabolite profiles may provide more information on genes underlying the QTL (Lou et al., 2008).

The screening of a DH mapping population, which varied in their parental profiles, with its associated molecular map, allows the identification of genetic regions that affect glucosinolate content in the mapping population. Genes affecting glucosinolate biosynthesis have been previously identified in *Arabidopsis* (Kliebenstein, 2009; Mewis et al., 2006; Pfalz et al., 2009) and in other *Brassica* species such as in *B. rapa* (Lou et al., 2008). The conservation of gene order (co-linearity) between *Arabidopsis*, *B. rapa* and *B. oleracea* (Luis Iniguez-Luy et al., 2009) can be used to

identify candidate genes underlying mapped QTLs that affect glucosinolate content. In addition, novel QTLs not previously identified may be discovered. Consequently, information developed in *B. oleracea* can be applied to other *Brassica* species as a result of the close relatedness of the species (Bellostas et al., 2007b; Gao et al., 2004; Kliebenstein et al., 2001a; Lou et al., 2008; Lukens et al., 2003).

Specific glucosinolates have been studied intensively for their bio-fumigation and anti-cancer effects (Bellostas et al., 2007b; Schonhof et al., 2004). The activity of glucosinolates is largely affected by their chemical structure, which is determined by their precursor amino acid and the type of modifications to the carbon side chain group (R) (Halkier and Gershenzon, 2006; Li and Quiros, 2003; Mithen, 2001; Windsor et al., 2005). Major genes underlying QTL affecting these individual glucosinolates biosynthesis ultimately characterized by map based cloning. For breeding purposes, markers tightly linked to the QTL could be adopted for marker assisted breeding strategies (Tanksley, 1993). Such information has potential applications in many different areas of ecological, agronomic, economic and health values.

3.2 Objectives

- To perform the analysis using different mapping methods; interval mapping (IM) and composite interval mapping (CIM), or multiple QTL mapping (MQM) methods, to search for consistency of QTL, utilizing two different QTL mapping programs.

- To determine genes or gene regulators underlying mapped QTL, utilising previously mapped genes in *Arabidopsis* and *B. rapa* for glucosinolate biosynthesis through applying a comparative genomic approach.

- To suggest regions on the genome where novel factors, which are involved in glucosinolate synthesis may be located.

3.3 Materials and methods

3.3.1 The genetic map

The AGDH mapping population is derived from a cross between rapid cycling *B. oleracea* line; A12DH (var. *alboglabra*) as the female parent and the broccoli line GD33DH (var. *italica*) as the male parent (Bohuon et al., 1996). Several versions of the genetic map for this population have been published (Bohuon et al., 1996; Rae et al., 1999; Sebastian et al., 2000). The map has been recently updated with the addition of a number of SSR markers and mapped gene loci (GR Teakle, University of Warwick, unpublished results). In this work, a subset of the AGDH population was available (89) that encompassed recombination breakpoints distributed widely across all the linkage groups. For the QTL analysis a subset of markers distributed at approximately 10 centi-Morgan (cM) intervals, selected based on having the most complete genotype information, was used (Barker et al., 2007).

3.3.2 **Plant material**

The AGDH population were grown under controlled environmental conditions as described in Chapter 2. Each plant line was represented by three genetically identical plants. Young fully expanded healthy leaves were collected at the bud initiation stage from the three plants as a bulk material, and were mixed in order to pool homogeneous plant material.

3.3.3 **Phenotyping**

The initial analysis showed that the glucosinolate profiles varied within the AGDH lines as segregation for these traits was found between the parental lines (as shown in Chapter 2) and therefore, the AGDH plant population was chosen for QTL mapping. In addition, the quantitative analysis of glucosinolate profiles in the AGDH plant lines revealed continuous variation for the individual glucosinolate concentrations, which suggesting they are complex traits controlled by polymorphic genes (Mackay, 1996). Each of these polymorphic genes contributes a small, approximately additive, effect on the phenotype at constant environmental conditions.

3.3.4 **Data analysis for QTL mapping**

In the AGDH plant lines analysed in this study, seven different glucosinolates segregating for content in the parental lines (see Chapter 2) were analysed. They were categorized into two different chemical classes of glucosinolates, the aliphatic glucosinolates including; glucoraphanin, progoitrin, sinigrin and gluconapin, and the indolic glucosinolates including; glucobrassicin, 4-methoxyglucobrassicin and neoglucobrassicin.

For QTL mapping of glucosinolates produced in the AGDH plant population, the data model used was based on the average relative concentration (relative to IS1) for the individual glucosinolates and on the sum of glucosinolates of the same chemical class obtained from three technical replicates (as described in Chapter 2). In addition, the sum of sinigrin and gluconapin was used to map QTL control of the alkene bond formation, where the sum of glucoraphanin and progoitrin was used to map QTL for the biosynthesis of the rest of aliphatic glucosinolates.

When a plant line did not express a glucosinolate at a concentration that could be detected, an arbitary figure of half the concentration of the lower limit of detection was used. The limit of detection for the individual glucosinolates were previously measured as described in Chapter 2 Table 13. This approach was chosen over simply including these points as missing data since biosynthesis for these glucosinolates could occur, producing glucosinolates at undetectable concentrations. These concentration values were $\log_{10}$ transformed to increase the homogeneity of variance between plant lines.

3.3.5 QTL mapping

Mapping QTL to chromosomal regions was performed using Windows QTL Cartographer ver 2.5 (Win QTL Cart) (htt: //stangen.ncsu.edu/qtlCart/WQTLCart. htm). The IM analysis was used to scan the whole genome searching for an interval locating between a pair of linked marker loci with regard to their effects on the quantitative trait, utilizing a molecular linkage map to obtain variance ratios to each significant QTL (Tanksley, 1993). Kosambi mapping function was used to translate from recombination frequency to distance on the chromosome with precision of 2 cM. To get more precise QTLs, CIM analysis was used (Zeng, 1993). CIM can

remove the effect of non-target QTL on the estimation of a target QTL. In the Win QTL Cart, the cofactors were chosen automatically by setting the program parameters using standard model 6, control marker number 5 (number of markers to control for the genetic background), window size of 10 cM (to block out regions of the genome on either side of the markers that are tightly linked to the test site) and the forward regression method.

The results were confirmed by reanalysing the data using Map QTL® ver 4.0 analysis (Van Ooijen, 2002), utilizing the IM analysis and the MQM analysis previously described by (Jansen and Stam, 1994). In the MQM analysis, markers nearby previously identified QTL in the IM analysis were fitted and used as cofactors, in order to absorb the effect of the QTL in their background, and therefore enhancing the power of the search for other segregating QTLs (Van Ooijen, 1999).

The $\log_{10}$ of the likelihood ratio (variance ratio); statistically determined as the Log of Odds (LOD score) (Mackay, 1996) for the presence of a segregating QTL compared to absence of segregating QTL, was calculated at a given position on the genome for individual plant lines in the population from the marker genotypes and the linkage map (Lander and Botstein, 1989). In all analysis methods, the frequency distribution of the maximum LOD score was determined by implementing a 1000 permutation test in order to determine the genome-wide significant threshold at ($p \leq 0.05$). When the QTL is significant, a LOD score peak exceeding the threshold value appears. Map positions for maximum LOD score and the two LOD support intervals were determined relative to the genetic map in cM, and shown for each QTL with the markers allocated at these points where applicable.

At any QTL locus the contribution of an increasing allele from the female parent (A12DH), indicated by a positive additive effect, is equivalent to the contribution of

a decreasing allele from a male parent (GD33DH), indicated by a negative additive effect. The genetic variance explained by QTL was calculated using Equation 2.

$$\frac{\left(\text{Additive effect}\right)^2}{\sum\limits_{n}\left(\chi - \bar{\chi}\right)^2} \times \frac{1}{n}$$

Equation 2 The genetic variance explained by QTL was calculated from the additive effect and the mean in 89 AGDH line, x: trait value, $x\square\square$: the mean of trait values in three technical replicates, n: size of plant population (Griffiths et al., 1996)

MapChart © ver 2.2 (Voorrips, 2002) software was used for the graphical presentation of linkage maps and QTLs.

3.4 Results and discussion

3.4.1 Variations of the glucosinolate content in the AGDH plant lines

The relative concentrations of individual glucosinolates (relative to IS1) were determined for the 89 AGDH population and their parental lines, A12DH and GD33DH in three technical replicates (as described in section 2.4.6). The average and standard deviation for the three technical replicates for each plant line were calculated, and the results are shown in Appendix D.

The initial analysis of the quantitative data, obtained using the average relative concentration of individual glucosinolates, indicates large variations in glucosinolate concentrations in all plant lines. For example, 87 AGDH plant lines were found to express glucobrassicin (as described in section 2.5.8.5), the average relative concentrations of glucobrassicin expressed by these plant lines, varied between 0.15 and 2.51 µmole/ g dry plant material (Figure 33). The standard deviations for the

relative concentrations, calculated from three technical replicates for each plant line indicate that the variance increased at higher concentrations in comparison to lower concentrations, as shown in Figure 39.
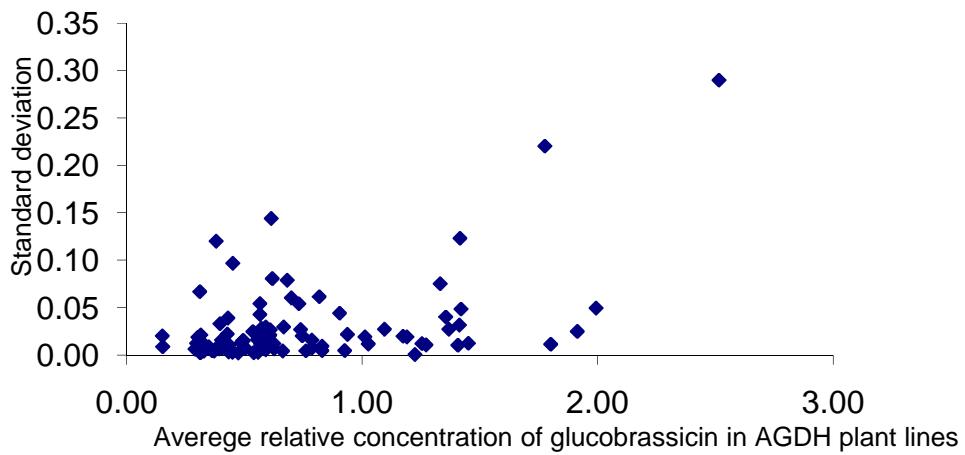
Figure 39 Scatter plot showing the standard deviation in three technical replicates compared to the average relative concentration (to IS1) of glucobrassicin in the 87 AGDH plant lines

When the quantitative data were $\log_{10}$ transformed, using the model of data analysis described in section (3.3.4), a continuous, fairly unimodal distribution of the average relative concentrations of glucobrassicin in the 89 AGDH plant lines was obtained, which approximated to a normal distribution as shown in Figure 40 .
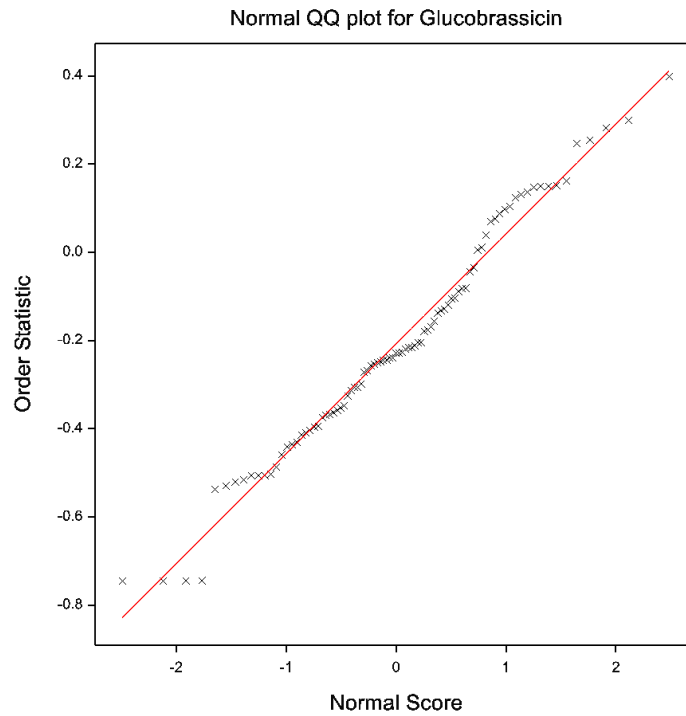
Figure 40 Normal Q.Q plot of the average relative concentration (to IS1) of glucobrassicin in 89 AGDH plant lines

The frequency histogram of the average relative concentrations of glucobrassicin in the 89 AGDH plant lines, obtained using the model of the analysis (as described in section3.3.4), shown in Figure 41, also shows a continuous distribution of glucobrassicin conentrations within the AGDH population. Glucobrassicin relative concentrations in the A12DH and GD33DH plant lines are indicated by arrows, showing that the synthesis of glucobrassicin by the parental plants lines were within the concentration range of its synthesis in the AGDH offsprings.

All other glucosinolates synthesied by the AGDH plant lines, were analysed using the same model as described for glucobrassicin analysis, and showed continuous distributions of their content. This indicates that the synthesis of glucosinolates is a quantitative trait possibly under the control of polygenes. The histograms showing

the continuous distribution of glucosinolate content in the AGDH plant lines and their parental lines are shown in Appendix D.
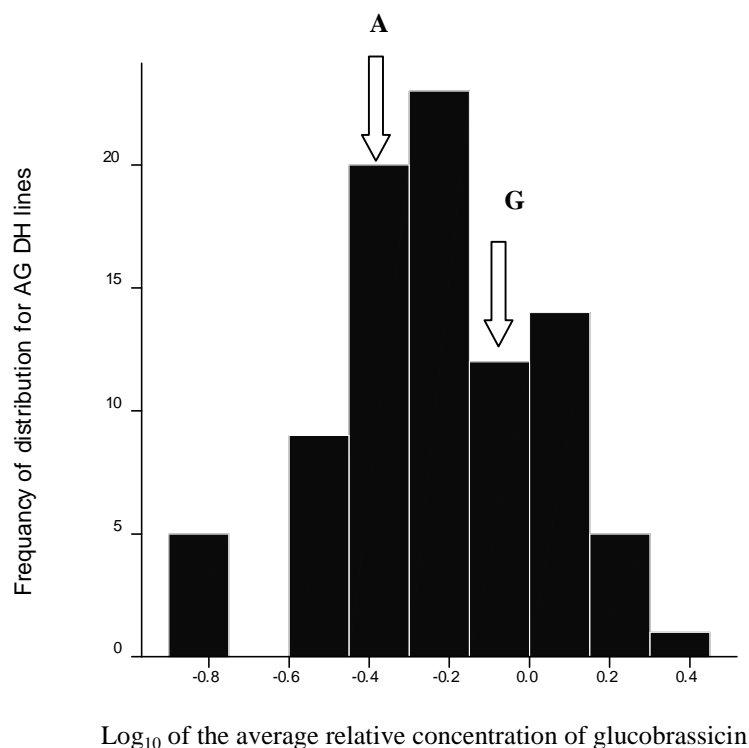
Figure 41 Frequency distribution of glucobrassicin in 89 AGDH plant lines. Parental scores are indicated by A = A12DH and G = GD33DH

3.4.2 Predicting the key points in glucosinolate biosynthesis pathways

Glucosinolates are classified into three major groups, namely aliphatic, indolic and aromatic glucosinolates, based on the amino acids from which they are synthesized via independent metabolic pathways, and share a common set of enzymes involved in the core structure formation of glucosinolates, which is under genetic control (Figure 2) (Halkier and Gershenzon, 2006). The general biosynthesis pathway of glucosinolates involves three main phases as summarised below:

- Prior to entering the biosynthesis pathway for the synthesis of aliphatic glucosinolates, methionine can undergo several elongation cycles for addition

of one methylene group at a time before it can enter the pathway for the formation of the aliphatic glucosinolate core structure.

- Conversion of amino acids to the basic glucosinolate skeleton.

- Additional side chain modifications can occur following the pathway, resulting in the vast diversity of glucosinolate content observed.

Consequently, the pattern of relationships between levels of individual glucosinolates across the AGDH population can provide information on the genetic contribution to the synthesis and modification pathways. Correlation analyses between glucosinolates expressed by the AGDH plant lines were illustrated using a scatter plot matrix shown in Figure 42. These plots allow the interpretation of trends within, and between, different groups of glucosinolates, demonstrating how genetic variations within a single biochemical step can affect subsequent products. By identifying genetic variation for such steps, we were then able to investigate them in more detail through QTL mapping.
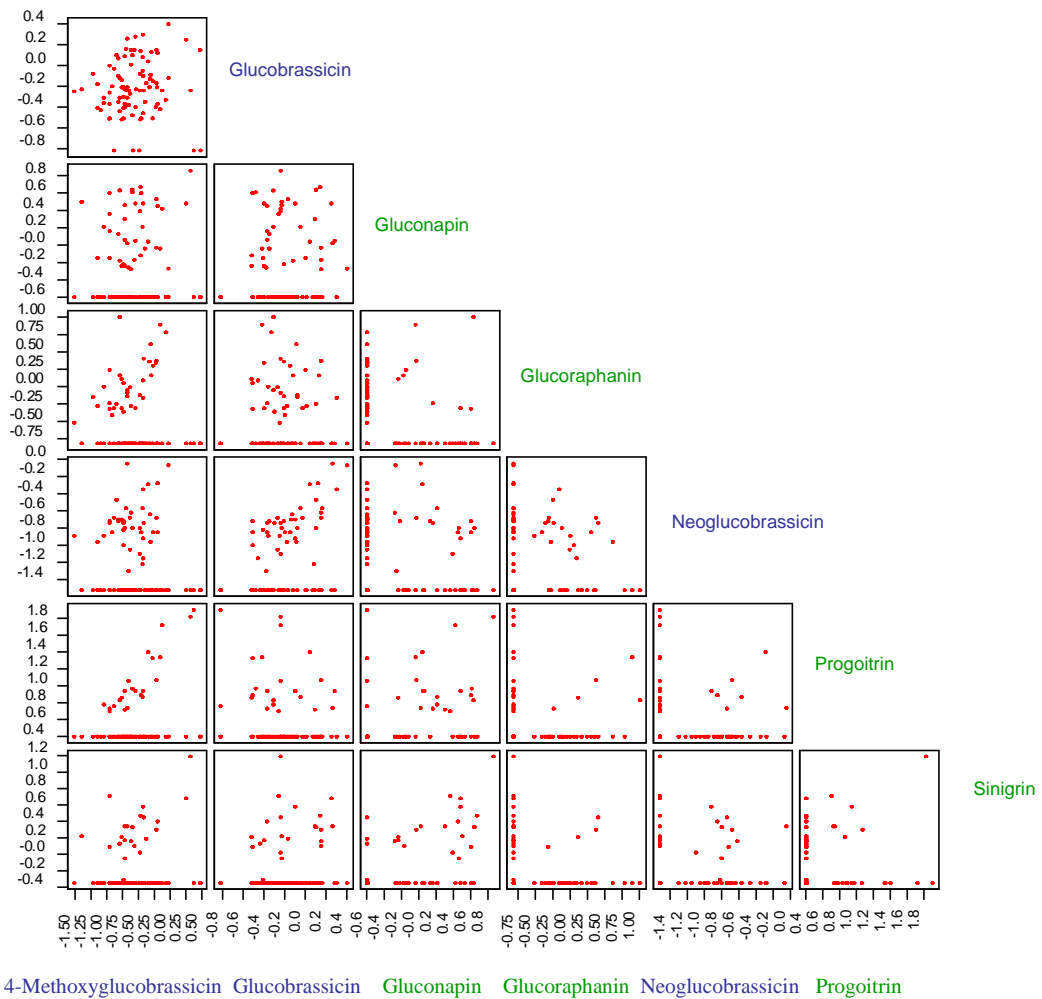
Figure 42 Scatter plots matrix of glucosinolate in 89 AGDH plant lines. Each individual plot represents a pairwise comparison for the average relative concentrations of two glucosinolates. When a plant line did not express glucosinolates at a detectable level, it was scored as half the amount of detection. Blue and green were for indolic and aliphatic glucosinolates, respectively.

The relationship between individual aliphatic glucosinolates; sinigrin with progoitrin, glucoraphanin with gluconapin, and progoitrin with gluconapin; indicate that when a plant line expressed any of them at concentrations lower than the detection level (as shown by the dots line near the edges of the pairwise squares), the other glucosinolates of the pair were expressed at detectable levels. This pattern indicates the presence of genes that control the biosynthesis of specific aliphatic glucosinolates, possibly through regulating the side chain modification phase. As

expected, when singrin and gluconapin were compared, there is no obvious relationship observed between their levels as they undergo independent pathways during the elongation phase prior to core structure formation of glucosinolates that is largely under genetic control.

However, comparing glucoraphanin to sinigrin and glucoraphanin to progoitrin, there are some indications of a positive relationship in their content indicating the presence of genes that control the core structure formation of the analysed aliphatic glucosinolates. Therefore, QTL mapping of individual and total aliphatic glucosinolates may reveal the presence of candidate loci that control aliphatic glucosinolate synthesis at specific positions in their synthesis.

A positive linear relationship was observed between the indolic glucosinolates analysed in the AGDH plant lines, which indicates that there is a shared step in their biosynthesis pathway.

As the biosynthesis of all classes of glucosinolates share a common set of enzymes involved in the core structure formation, the pairwise plot for 4-methoxyglucobrassicin with sinigrin, progoitrin, and glucoraphanin, showed a positive linear relationship between aliphatic and indolic glucosinolates levels.

3.4.3   Analysis of QTL affecting glucosinolates content in *B. oleracea*

All the plant populations used in this experiment were grown in the same glasshouse and under constant environmental conditions in order to minimize the environmental effect (as described in section 2.2.2). These mapping analyses highlighting different QTLs for different glucosinolates can be divided into two main categories as discussed below.

3.4.3.1 **QTL associated with aliphatic glucosinolate biosynthesis**

This section, will be considering the significant QTLs with $p \leq 0.05$ only. The LOD scores at each marker position were calculated for individual aliphatic glucosinolates, total aliphatic glucosinolates, the sum of glucoraphanin and progoitrin and for the sum of sinigrin and gluconapin utilizing the CIM analysis, as shown in Table 19. All other non-significant QTLs detected using the CIM analysis was not detected using other mapping methods. Thus, they were not investigated further.

QTL mapping revealed the presence of 26 total QTLs, of which only 18 QTLs were significantly affecting aliphatic glucosinolate content, distributed on 6 out of 9 linkage groups (LG) over the C genome of *B. oleracea,* as shown on Map 1.

Interestingly on LG3 (Map 1-A), a QTL was found for progoitrin, which co-localized with QTLs for gluconapin, and for the sum of sinigrin and gluconapin were identified within two discrete regions. The QTL identified for the sum of sinigrin and gluconapin was supported by the presence of the same QTL within approximately the same interval (44.7-71.7) cM, with a significant LOD score using Map QTL/ MQM analysis (Table 3 in Appendix E).

Table 19 QTLs detected for aliphatic glucosinolate and sub classes of aliphatic glucosinolates, in 89 AGDH segregating mapping population sorted by trait type using the Win QTL Cartographer program with CIM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) and ** (p≤0.01) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. Italic for QTL confirmed by IM or by Map QTL analysis. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance calculated using Equation 2 for the significant QTLs only

Trait	LG	Position (cM)	Marker	Additive effect	LOD score	LOD Threshold	Two LOD support interval (cM)	% variation explained by QTL
Glucoraphanin	1	30.4	pW239E2	0.1736	2.45	2.7		
	7	0.0	pO87E2	-0.1396	1.56			
	7	9.1	pO131E2	-0.1346	1.65			
	7	62.0	pCeriE3	0.1488	1.79			
	*9**	14.9	pO125E1N	-0.2120	3.46		0.0-23.1	17.14%
Progoitrin	*3**	30.5	pW111J1	-0.1990	3.15	2.5	22.1-40.3	30.94%
	4	23.2	pO171J1	-0.0994	2.20			
	*7**	0.0	pO87E2	-0.1307	2.66		0.0-17.9	13.34%
Sinigrin	*3*	83.6	pN207E1	-0.0810	1.75	2.6		
	*5**	33.6	pW164E1	0.1243	3.74		21.5-42.8	13.27%
	8	58.1	pO143E2	-0.0989	1.91			
	*9**	0.0	pN52E2	0.1250	3.73		0.0-12.3	13.42%
Gluconapin	3	20.4	pN102E1	0.1131	1.67	3.0		
	*3**	56.2	pW143J1	-0.1551	3.12		44.3-71.3	7.99%
	8	47.4	pR97J1	-0.1637	2.18			
	*8**	58.1	pO143E2	-0.1700	3.40		39.8-75.7	9.60%
	*9***	12.9	pO125E1N	0.2601	8.45		2.2-20.9	22.49%
Total aliphatic glucosinolates	*7**	0.0	pO87E2	-0.1355	3.25	2.5	0.0-17.1	16.69%
	*8**	58.1	pO143E2	-0.1189	2.51		39.2-75.7	12.85%
	*9**	4.0	pN52E2	0.1082	2.60		0.0-23.3	10.55%
Sum of glucoraphanin and progoitrin	2	0.0	pW116E1	-0.0415	1.93	2.6		
	*7**	0.0	pO87E2	-0.0638	3.59		0.0-17.4	16.76%
	*8**	58.1	pO143E2	-0.0595	3.04		39.5-75.9	14.58%
	*9**	4.0	pN52E2	0.0603	3.77		0.0-22.9	14.79%
Sum of sinigrin and gluconapin	*3**	66.4	pN213J2	-0.1120	2.95	2.7	44.7-71.7	8.62%
	*8**	56.6	AC-	-0.1405	3.45		51.4-75.7	13.57%
	*9***	10.0	CAAE05	0.1859	5.33		0.0-22.9	23.75%
	*9**	58.4	pN52E2	-0.2035	2.69			28.46%
			pW233J1					

133

C3
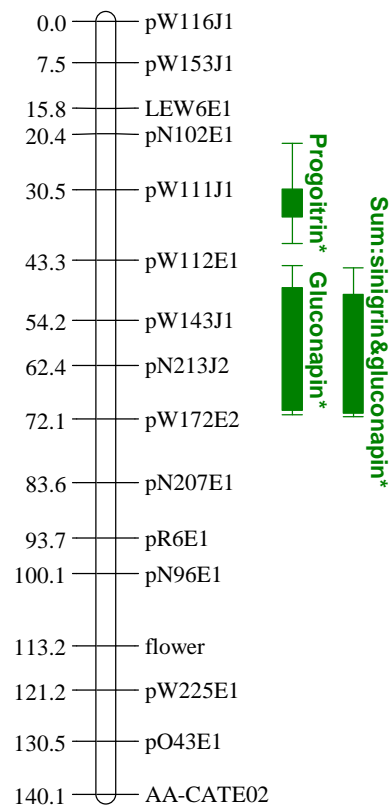
Map 1-A *Brassica oleracea* linkage map based on AGDH population (Unpublished revision of the Sebastian et al (2000) Integrated map by Graham Teakle) with QTLs detected for individual aliphatic glucosinolates, the total aliphatic glucosinolates, the sum of glucoraphanin and progoitrin and the sum of sinigrin and gluconapin using Win QTL Car. CIM analysis. C: chromosomes1-9, significant QTL determined at *(p≤ 0.05) and ** (p≤ 0.001)

On LG5 Continue Map -B) in the interval (21.5-42.8) cM, evidence of a QTL for sinigrin was strengthened by MQM analysis using the MAP QTL program (Table 3 Appendix E), which yielded a significant QTL.

C5

0.0	pN21E2
8.6	pO92J1
21.0	pO105J1
33.6	pW164E1
43.6	pN148E1
53.6	pN91E3
60.7	pO123J1
78.3	AC-CACE02
87.0	pN113E1

Sinigrin*

Continue Map 1-B

On LG7 at (0.0-17.1) cM (Continue Map 1-C), three QTLs for progoitrin, total aliphatic glucosinolates and for the sum of glucoraphanin and progoitrin are co-localized. The QTLs identified for the sum of glucoraphanin and progoitrin were observed at different positions using IM analysis at 50.3 cM (Table 20) and Map QTL/ IM analysis at 50.3 cM (Table 1 Appendix E).

C7
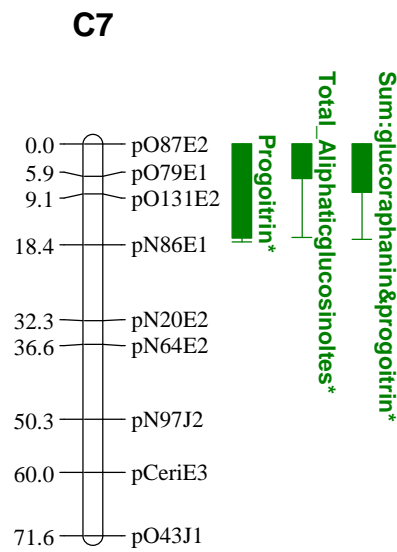
Continue Map 1-C

Table 20 QTLs detected for individual glucosinolate, total aliphatic glucosinolates and sub classes of aliphatic glucosinolates, in 89 AGDH segregating mapping population sorted by trait type using the Win QTL Cartographer program with IM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) ** (p≤0.001) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. Italic for QTL confirmed by CIM or by Map QTL analysis. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance calculated using Equation 2 for the significant QTLs only

Trait	LG	Position (cM)	Marker	Additive effect	LOD score	LOD Threshold	Two LOD support Interval (cM)	%variation explained by QTL
Glucoraphanin	*9*	14.9	pO125E1N	-0.1313	1.26	2.7		
	9	27.4	pW137J1	-0.1236	1.13			
Progoitrin	1	73.3	pN53E2	-0.1419	1.48	2.4		
	3	30.5	pW111J1	-0.1044	1.83			
	4	23.2	pO171J1	-0.0967	1.62			
Sinigrin	1	81.3	pN53E2	-0.1266	1.67	2.5		
	1	99.6	pW216J1	-0.1017	1.78			
	3	83.6	pN207E1	-0.0977	1.78			
	5	33.6	pW164E1	0.0910	1.51			
	9	4.0	pN52E2	0.1113	2.0			
Gluconapin	*3*	54.2	pW143J1	-0.1482	2.15	2.9		
	7	50.3	pN97J2	0.1371	1.84			
	*9***	8.0	pN52E2	0.2905	7.82		0.5-21.6	28.06%
	*9***	25.4	pW137J1	0.2295	5.18		0.0-37.1	17.51%
Total aliphatic glucosinolates	*1*	75.3	pW112E1	-0.1410	1.52	2.5		
	9	8.0	pN52E2	0.1042	1.81			
Sum of glucoraphanin and progoitrin	*4*	23.2	pO171J1	-0.0415	1.57			
	7	50.3	pN97J2	0.0426	1.71	2.6		
	8	58.1	pO143E2	-0.0507	1.70			
	*9**	8.0	pN52E2	0.0593	2.61		0.0-46.1	14.48%
Sum of sinigrin and gluconapin	1	79.3	pN53E2	-0.1516	1.54	2.7		
	1	97.6	pW216J1	-0.1076	1.54			
	3	54.2	pW143J1	-0.1183	2.21			
	3	66.4	pN213J2	-0.1038	1.50			
	8	60.1	pO143E2	-0.1247	1.58			
	*9***	8.0	pN52E2	0.2137	6.64		0.0-23.7	31.38%

On LG8 at the interval (39.2-75.9) cM Continue Map 1-D), QTLs for gluconapin, total aliphatic glucosinolates, the sum of sinigrin and gluconapin, and for the sum of glucoraphanin and progoitrin are co-localized, QTLs for gluconapin and total aliphatic glucosinolates were only detected using CIM analysis. However, QTLs for the sum of sinigrin and gluconapin, and for the sum of glucoraphanin and progoitrin were only significant using CIM analysis, and they were supported by the presence of QTLs within the same intervals near to each other, shown using IM analysis (Table 20) and Map QTL/ IM analysis (Table 1 Appendix E). These QTLs suggest the presence of two co-localized QTLs with different genes affecting the expression of particular glucosinolates, possibly controlling side chain modifications for aliphatic glucosinolates.

C8
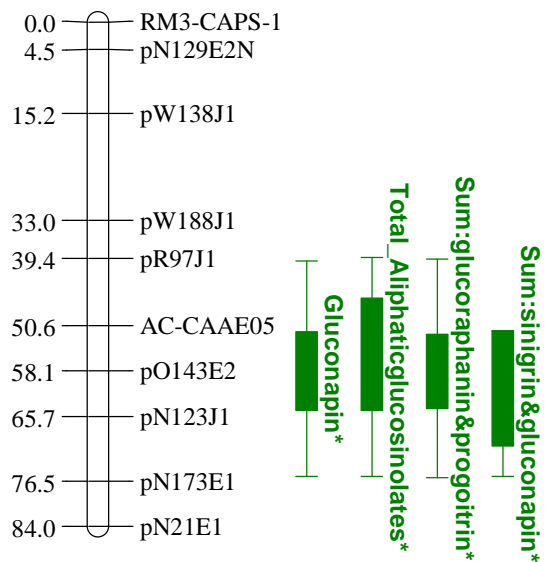
Continue Map 1-D

On LG9 (Map 1-E), six co-localized QTLs were detected at the interval of (0.0-23.3) cM underlying for glucoraphanin, sinigrin, gluconapin, total aliphatic glucosinolates, the sum of glucoraphanin and progoitrin, and for the sum of sinigrin and gluconapin. For sinigrin, QTLs within the same interval were detected using the IM analysis (Table 20) and Map QTL/ IM analysis (Table 1 Appendix E); they were below the significant threshold level. Therefore, there was a weak evidence to support the presence of these QTLs. In addition, a QTL for the sum of sinigrin and gluconapin was detected at 58.4 cM with (p≤0.05).

Continue Map 1-E

139
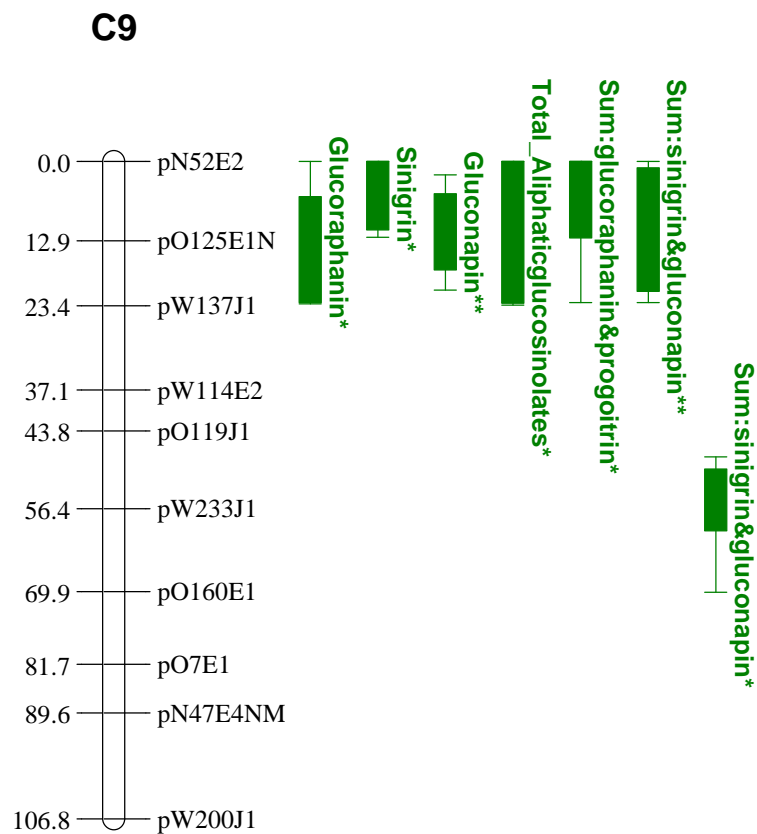
3.4.3.2 **Major gene effect**

The QTLs underlying gluconapin, total aliphatic glucosinolates, the sum of glucoraphanin and progoitrin, and the sum of sinigrin and gluconapin, were all located on LG9 within the same interval (0.0- 12.9) cM using CIM (Table 19), IM (Table 20) and Map QTL/ IM and MQM analysis (Tables1 and 3 in Appendix E) and were all significant ($p<0.05$).

Interestingly, the QTLs for gluconapin and the sum of gluconapin and sinigrin were highly significant ($p<0.001$) using CIM, IM analysis as well as using Map QTL/ IM and MQM analysis, and were located on LG9 near to each other. Approximately half the plant lines produced a detectable level of gluconapin and half did not, while 35% of the plant lines produced a detectable level of sinigrin and 65% did not (Figure 43). These findings raised a strong suggestion for the presence of a major gene effect at this locus. That was investigated by converting the quantitative data into presence or absence scores corresponding to the parent, which formed a set that could be genetically mapped.
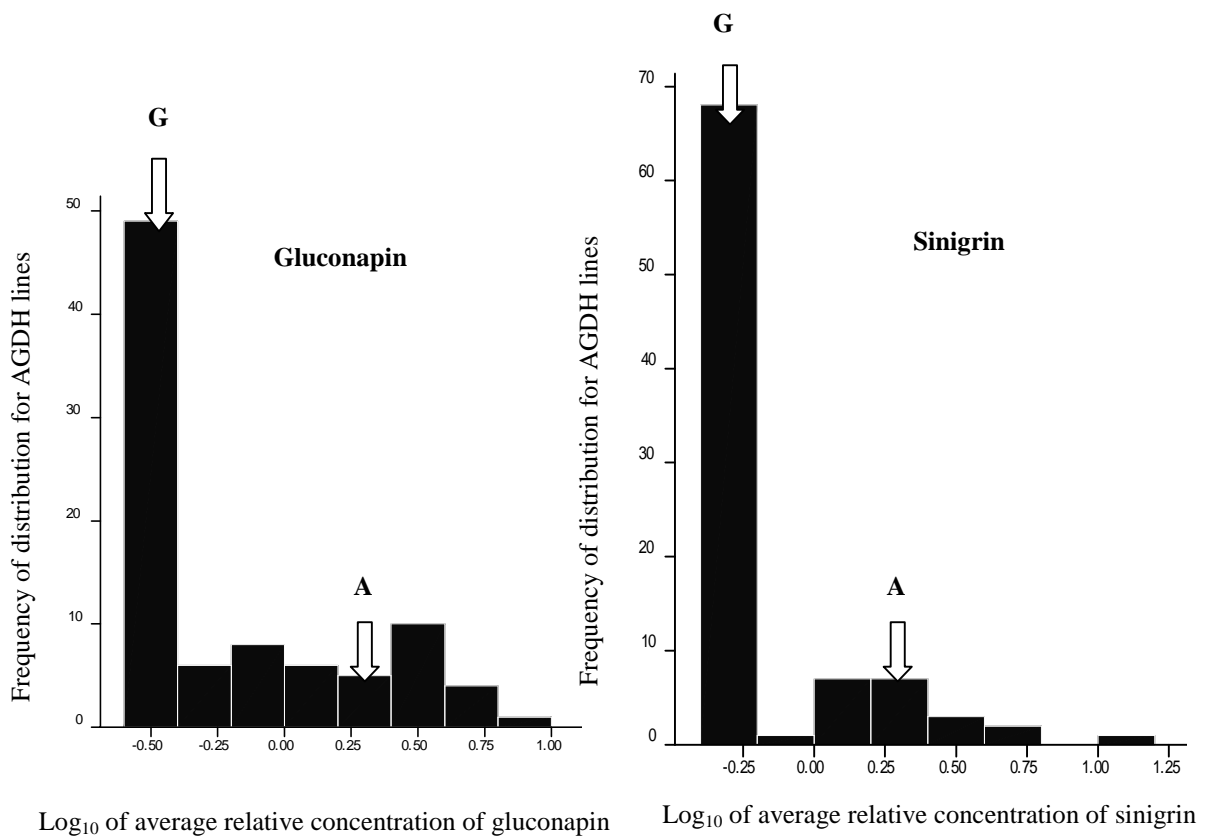
Figure 43 Frequency distribution of gluconapin and sinigrin in 89 AGDH plant lines. Parental scores are indicated by A = A12DH and G = GD33DH

Gluconapin showed linkage to LG9 and mapped convincingly as a single locus at 9 cM, indicating that a major gene controlling the content of gluconapin is associated with this locus. A single dominant Mendelian gene in *B. oleracea* controls the production of alkene side chain glucosinolates, this has been mapped on LG9, at the interval between the markers pW157 and pW137 at 12.3 and 23.4cM, respectively (Hall et al., 2001).

There are 9 double recombinant scores (DRs) for this marker that don't fit the pattern of recombination in the map and usually indicate scoring errors. These were identified for the following plant lines: AGDH1036, AGDH2056, AGDH2185, AGDH2206, AGDH2221, AGDH5012, AGDH6016, AGDH1038 and AGDH3013, their dominant effect removed and were stated as missing values.

141

In order to map the gene with the major effect on the synthesis of gluconapin, the plant lines expressing gluconapin at concentrations lower than the detection level were given missing values to eliminate their gene effect, consequently mapping QTL for the major gene.

Once a major gene effect was identified, in order to remove its influence on traits known to be linked through our knowledge of the metabolic pathway, the QTL data was reanalysed following removal of lines associated with the dominance. As expected, progoitrin and sinigrin were in agreement with presence or absence categories with gluconapin, and 8 of the 9 DR lines were within the group of absence scores. These absence scores were replaced with missing values except for the DR lines and they were given values of half the concentration at the lower limit of detection for the corresponding glucosinolates. This data model was used for mapping QTLs underlying major gene effects controlling gluconapin, sinigrin and progoitrin content (Map 2, Table 21). In another model, the DR lines scores were replaced with missing data.

Table 21 QTLs detected for glucosinolates expected to be under the control of major gene effect in 89 AGDH segregating mapping population, sorted by trait type using the Win QTL Cartographer program with CIM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. A missing value was used when a plant line was expressing the corresponding glucosinolate at concentrations less than the detection level and with the 9 double recombinant plant lines were indicated as (–DR). Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. Italic for QTL confirmed by IM or by Map QTL analysis. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance calculated using Equation 2 for the significant QTLs only

Trait	LG	Position (cM)	Marker	Additive effect	LOD score	LOD Threshold	Two LOD support interval (cM)	%variation explained by QTL
Gluconapin	7*	62.0	pCeriE3	-0.2332	3.73	2.9	48.4-68.0	18.08%
	8	0.0	RM3-CAPS-1	0.1840	2.13			
Progoitrin	3*	99.7	pR6E1	0.1810	2.71	2.5	82.7-101.4	25.60%
	3	113.2	flower	-0.1720	2.28			
	3	125.2	pW225E1	-0.1384	1.59			
	7	50.3	pN97J2	-0.1250	1.98			
	7	70.0	pCeriE3	-0.1143	1.71			
	9*	37.1	pW114E2	-0.1584	3.13		23.2-54.2	19.60%
Sinigrin	2*	104.7	pW141E1	0.1880	3.24	2.7		30.38%
	2*	113.8	pO119E2	0.1979	3.67			33.66%
	5*	12.6	pO92J1	-0.2437	3.72		1.9-20.6	51.05%
	5	33.6	pW164E1	0.1637	2.04			
Progoitrin -DR	3	30.5	pW111J1	-0.1886	3.13	3.5		
	4	15.0	pW143E2	-0.1374	1.84			
	4	23.2	pO171J1	-0.1341	1.86			
	7	0.0	pO87E2	-0.1703	2.41			
Sinigrin -DR	3	72.1	pW172E2	-0.1151	1.91	2.8		
	5*	31.0	pO105J1	0.2833	7.19		24.5-39.9	47.18%
	5*	51.6	pN148E1	0.2173	3.24		21.0-53.6	27.76%
	5	86.3	AC-CACE02	-0.1406	2.13			

Using the designed data model excluding the major gene effect, significant QTLs were then mapped on LG3 and LG2 for progoitrin and sinigrin respectively (Map 2-A), utilizing the CIM analysis.

C2   **C3**

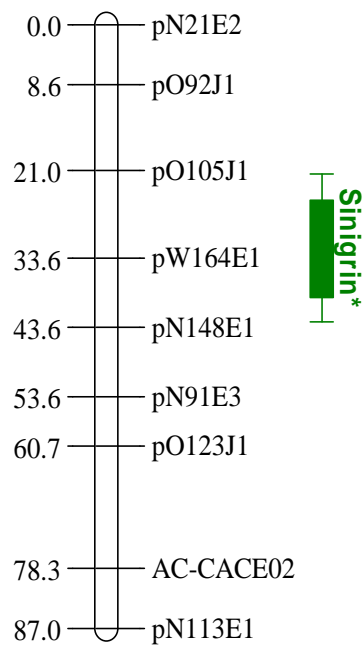|	
0.0 — pW116E1	0.0 — pW116J1
	7.5 — pW153J1
	15.8 — LEW6E1
	20.4 — pN102E1
	30.5 — pW111J1
41.7 — pN121E1	43.3 — pW112E1
49.4 — pN102E2	
	54.2 — pW143J1
59.9 — pN180E2	62.4 — pN213J2
69.2 — pN63E1	72.1 — pW172E2
79.3 — pR72J1	83.6 — pN207E1
88.6 — pO120E1	93.7 — pR6E1
	100.1 — pN96E1
104.7 — pW141E1	113.2 — flower
109.8 — pO119E2	121.2 — pW225E1
115.6 — AC-CTAE03	130.5 — pO43E1
Sinigrin*	140.1 — AA-CATE02
	Progoitrin*
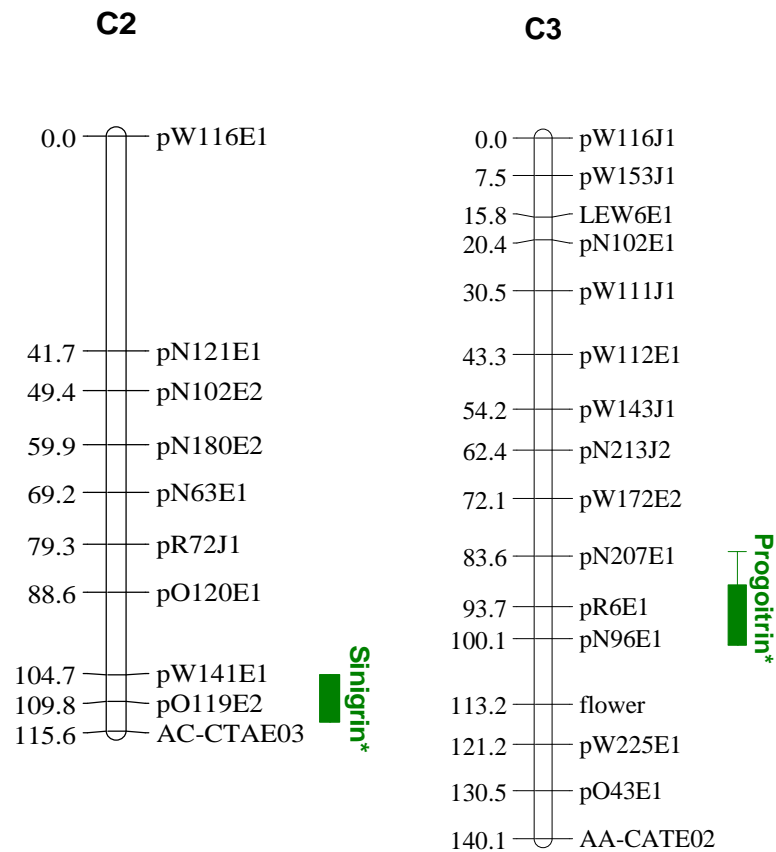
Map 2-A *Brassica oleracea* linkage map based on AGDH population (Unpublished revision of the Sebastian et al (2000) Integrated map by Graham Teakle) show QTLs detected for aliphatic glucosinoltes, idetified to be under the control of major gene effect, using Win QTL Car. CIM analysis, C: chromosomes 1-9, -DR indicating double recombinant plant lines were mapped as a missing value, significant QTL determined at *(p≤ 0.05)

A significant QTL for sinigrin was mapped on LG5 at the interval of (1.9-20.6) cM (Map 2-B) while another non-significant QTL was mapped at 33.6 cM (Table 21). Using Map QTL/ IM analysis, a significant QTL was mapped at the interval of (8.6-43.6) cM for the same trait (Table 2 in Appendix E), which could indicate the presence of several closely linked QTLs affecting this trait, but could also have resulted from a single underlying QTL. In the alternative model in which the DR lines were omitted, a significant QTL (p<0.05) was detected in the interval of (24.5-39.9) cM. This QTL had a subsidiary peak at 51.6 cM which would also have been significant (p<0.05) as an isolated peak, which might indicate the presence of more than one QTL. The evidence of at least one QTL in this region was supported by the IM analysis (Table 22) and the Map QTL/IM analysis (Table 2 in Appendix E) as each showed a single significant QTL in the intervals of (23.6-40.1) cM and (23-43) cM, respectively. Further work using backcrossed material RI lines would resolve this point.

C5
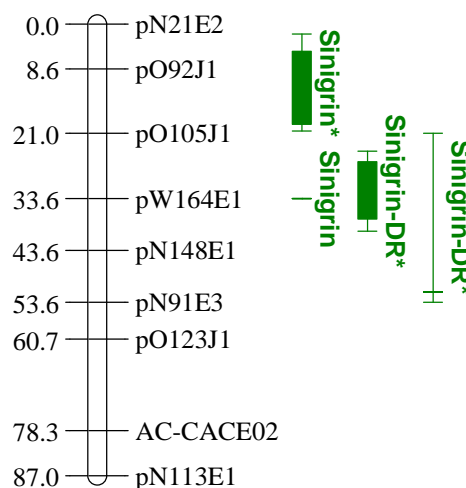
Continue Map 2-B

Table 22 QTLs detected for glucosinolates expected to be under the control of major gene effect in 89 AGDH segregating mapping population, sorted by trait type using the Win QTL Cartographer program with IM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. A missing value was used when a plant line was expressing the corresponding glucosinolate at concentrations less than the detection level and with the 9 double recombinant plant lines were indicated as (–DR). Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. Italic for QTL confirmed by CIM or by Map QTL analysis. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance, was calculated using Equation 2 for the significant QTLs only

Trait	LG	Position (cM)	Marker	Additive effect	LOD score	LOD Threshold	Two LOD support interval (cM)	% variation explained by QTL
Gluconapin	*7*	60.0	pCeriE3	-0.1693	1.80	2.9		
Progoitrin	*9*	31.4	pW137J1	-0.1482	1.55	2.6		
Progoitrin -DR	*3*	30.51	pW111J1	-0.1543	1.71	3.6		
Sinigrin -DR	*3*	83.6	pN207E1	-0.1553	1.62	3.0		
	5	16.6	pO92J1	0.1669	1.71			
	*5**	33.6	pW164E1	0.2393	4.42		23.6-40.1	33.66%

A significant QTL was mapped at the interval of 48.4-68.0 cM on LG7 for gluconapin (Map 2-C, Table 21) was also mapped within the same interval using IM analysis (Table 22). However, another significant QTL for the same trait was mapped within the same interval (but was 10 cM shifted from the maximum point) using the Map QTL/ IM analysis (Table 2 in Appendix E).

C7

0.0	pO87E2
5.9	pO79E1
9.1	pO131E2
18.4	pN86E1
32.3	pN20E2
36.6	pN64E2
50.3	pN97J2
60.0	pCeriE3
71.6	pO43J1

Gluconapin*
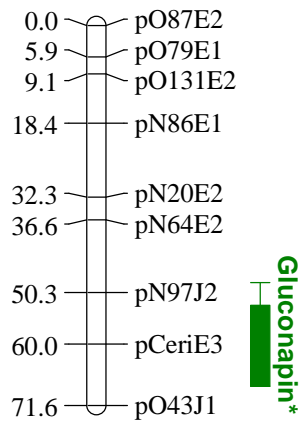
Continue Map 2-C

A QTL for progoitrin was mapped on LG9 (Map 2-D, Table 21), at the interval of (23.2-54.2) cM. However, the IM analysis showed the same QTL mapped at 31.4 cM with a LOD score below the threshold level (Table 22), this gives weak evidence for this QTL.

C9

0.0	pN52E2
12.9	pO125E1N
23.4	pW137J1
37.1	pW114E2
43.8	pO119J1
56.4	pW233J1
69.9	pO160E1
81.7	pO7E1
89.6	pN47E4NM
106.8	pW200J1

Progoitrin*

Continue Map 2-D

147
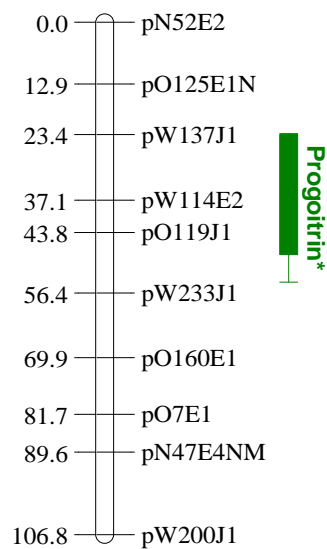
3.4.3.3 **QTL associated with indolic glucosinolates biosynthesis**

QTL mapping revealed the presence of 4 significant QTLs out of total 11 QTLs detected to affect indolic glucosinolate content, distributed on 4 out of 9 LG over the C genome of *B. oleracea,* as shown on Map 3.

In this section, we will be considering the significant QTLs (p≤ 0.05) only. The LOD scores at each marker position were calculated for individual and total indolic glucosinolates utilizing the CIM analysis, as shown in Table 23.

Table 23 QTLs detected for individual indolic glucosinolate and total indolic glucosinolates, in 89 AGDH segregating mapping population sorted by trait type using the Win QTL Cartographer program with CIM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) and ** (p≤0.01) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. Italic for QTL confirmed by IM or by Map QTL analysis. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance calculated using Equation 2 for the significant QTLs only

Trait	LG	Position (cM)	Marker	Additive effect	LOD score	LOD Threshold	Two LOD support interval (cM)	%variation explained by QTL
Glucobrassicin	*1**	97.6	pW216J1	-0.0913	2.89	2.5	91.4-101.6	10.39%
	7	2.0	pO87E2	0.0912	2.15			
	9	31.4	pW137J1	-0.0746	1.71			
	9	47.8	pO119J1	-0.0855	2.23			
Neoglucobrassicin	3	123.2	pW225E1	0.1162	1.73	2.8		
	4	89.6	pW139E1	-0.1412	2.77			
	*5**	20.6	pO92J1	-0.1997	3.19		8.5-23.5	19.35%
	5	39.6	pW164E1	0.2159	2.30			
	5	47.6	pN148E1	0.1542	1.89			
Total indolic glucosinolates	1	103.6	pW216J1	-0.0583	1.50	2.5		
	*2**	0.0	pW116E1	-0.0833	2.86		0.0-34.9	10.98%
	*9**	35.4	pW137J1	-0.0796	2.60		23.1-56.0	10.03%

For glucobrassicin, a QTL was mapped on LG1 (Map 3-A) at the interval of (91.4-101.6) cM, shifted by 10 cM when mapped at the same locus using the IM analysis (Table 24). However, the same QTL was mapped within the same interval using the Map QTL/ IM analysis (Table1 in Appendix E), with non-significant LOD score. Therefore, weak evidence for the presence of this QTL can be suggested.

C1
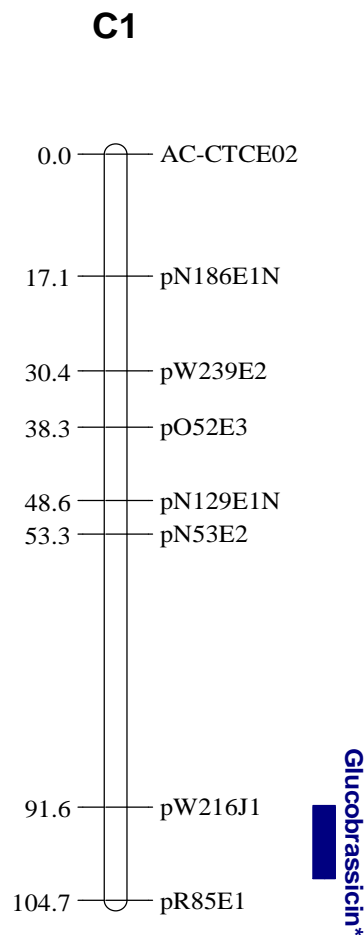
Map 3-A *Brassica oleracea* linkage map based on AGDH population (Unpublished revision of the Sebastian et al (2000) Integrated map by Graham Teakle), with QTLs detected for individual indolic glucosinolate and total indolic glucosinolate content using the Win QTL Car, and utilizing CIM analysis. C: chromosomes1-9, significant QTL were determined at *(p≤ 0.05) and ** (p≤ 0.001)

Table 24 QTLs detected for individual indolic glucosinolate and total indolic glucosinolates, in 89 AGDH segregating mapping population sorted by trait type using the Win QTL Cartographer program with IM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) determined by 1000 permutation test for each trait analysed. The maximum LOD point shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. Italic for QTL confirmed by CIM or by Map QTL analysis.

Trait	LG	Position (cM)	Marker	Additive effect	LOD score	LOD Threshold
Glucobrassicin	*1*	87.3	pN53E2	-0.0905	1.87	2.6
	1	99.6	pW216J1	-0.0868	1.98	
	7	0.0	pO87E2	-0.0821	1.58	
	9	47.8	pO119J1	-0.0949	2.16	
Neoglucobrassicin	*4*	89.6	pW139E1	-0.1502	2.47	2.7
Total Indolic glucosinolates	*2*	0.0	pW116E1	-0.0721	1.79	2.6
	9	37.1	pW114E2	-0.0708	1.77	

QTLs for the total content of indolic glucosinolates were mapped on LG2 and LG9 (Map 3-B), at the intervals of (0.0-34.9) cM and (23.1-56) cM, respectively (Table 23). These QTLs were mapped at the same intervals on the genome but with insignificant LOD scores using the IM analysis (Table 24) and Map QTL/ IM analysis (Table 1 in Appendix E), which indicats a weak evidence for the presence of these QTLs.
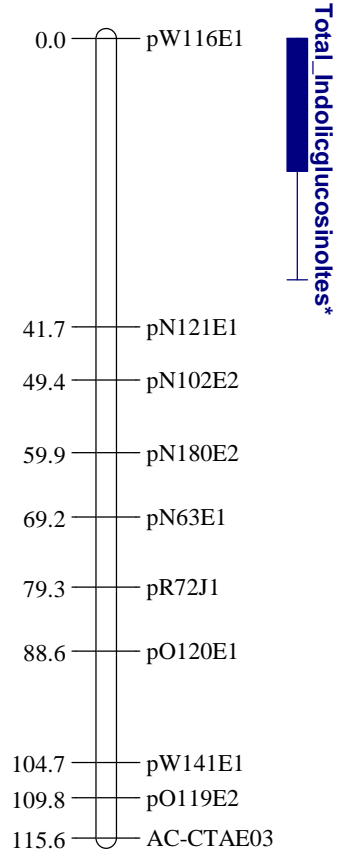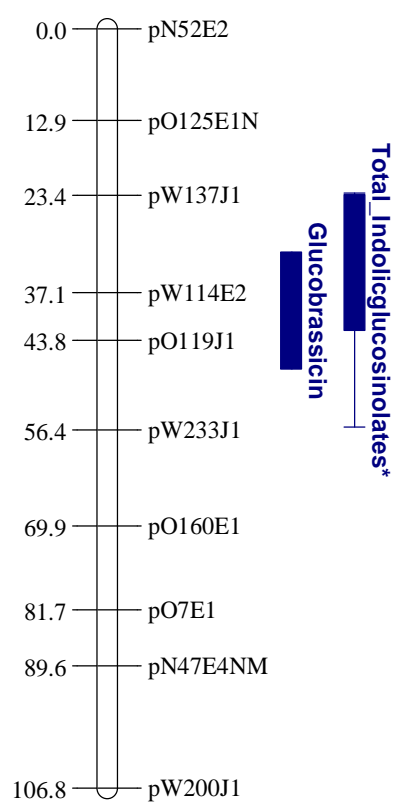
Continue Map 3-B

A significant QTL for neoglucobrassicin mapped on LG5 at the interval (8.5-23.5) cM (Map 3-C) (Table 23). A similar result was found using IM (Table 24) and the Map QTL/ IM analysis (Table 1 in Appendix E) with LOD score below the threshold level, suggesting there is a possible QTL on that locus.

C5

0.0	pN21E2
8.6	pO92J1
21.0	pO105J1
33.6	pW164E1
43.6	pN148E1
53.6	pN91E3
60.7	pO123J1
78.3	AC-CACE02
87.0	pN113E1

Neoglucobrassicin*
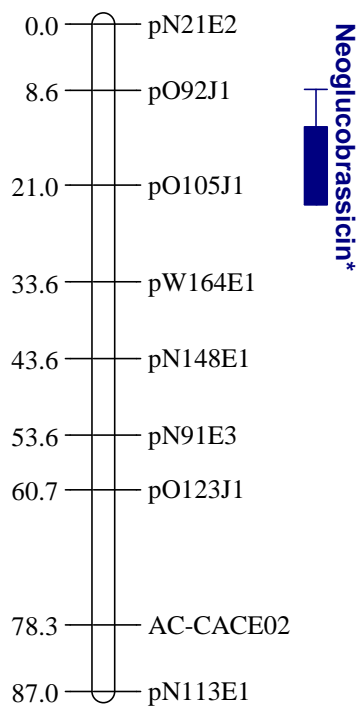
Continue Map 3-B

A non-significant QTL for glucobrassicin was mapped on LG9 at (48.8) cM using the CIM analysis (Table 23), IM analysis (Table 24) and Map QTL/ IM analysis (Table 2 in Appendix E). However, this QTL was mapped at the same locus using Map QTL/ MQM analysis with significant LOD score (p= 0.05) when the two markers (PW216J1 and PR85E1on LG1) were used as cofactors (Table 3 in Appendix E). These findings suggested a possible QTL on that locus.

3.4.4 Identifying the genes involved in the biosynthesis of aliphatic and indolic glucosinolates in *B. oleracea*

Aliphatic and indolic glucosinolates are synthesized through independent metabolic pathways (Halkier and Gershenzon, 2006; Mewis et al., 2006; Zang et al., 2009). Four different aliphatic glucosinolates were analysed in the AGDH plant lines, these have been classified according to the length of the R side chain (Magrath et al., 1994), as follows:

➢ The three-carbon R side chain glucosinolates, including singrin with an alkene bond.

➢ The four-carbon R side chain glucosinolates, including glucoraphanin with a sulphinyl group, gluconapin with an alkene bond and progoitrin with a hydroxyl group and an alkene bond.

The biosynthetic pathway for aliphatic glucosinolates identified in the AGDH population was predicted as shown in Figure 44. This biosynthetic pathway involves methionine; the precursor amino acid for the aliphatic glucosinolate synthesis undergoes an elongation phase before it can enter the core structure formation phase, is under genetic control and regulates the length of the R side chain for the aliphatic glucosinolates. Fine mapping of the *Gls-elong* loci on chromosome 5 in *A. thaliana*

(de Quiros et al., 2000) identified members of the *MAM* gene family; *MAM2* and *MAM1* coding for the synthesis of 3 and 4 carbon side chain which correspond to the production of 3-methylthiopropyl and 4-methylthiobutyl glucosinolates, respectively.

Methionine

MAM2   *MAM1*

3-methylthiopropyl   **4-methylthiobutyl**

Gls- oxid

**4-methylsulphinylbutyl
Glucoraphanin**

3-methylsulphinylpropyl

Gls-Alk

**2-propenyl
Sinigrin**

**3-butenyl
Gluconapin**

Gls- OH

**2-hydroxy-3-butenyl
Progoitrin**
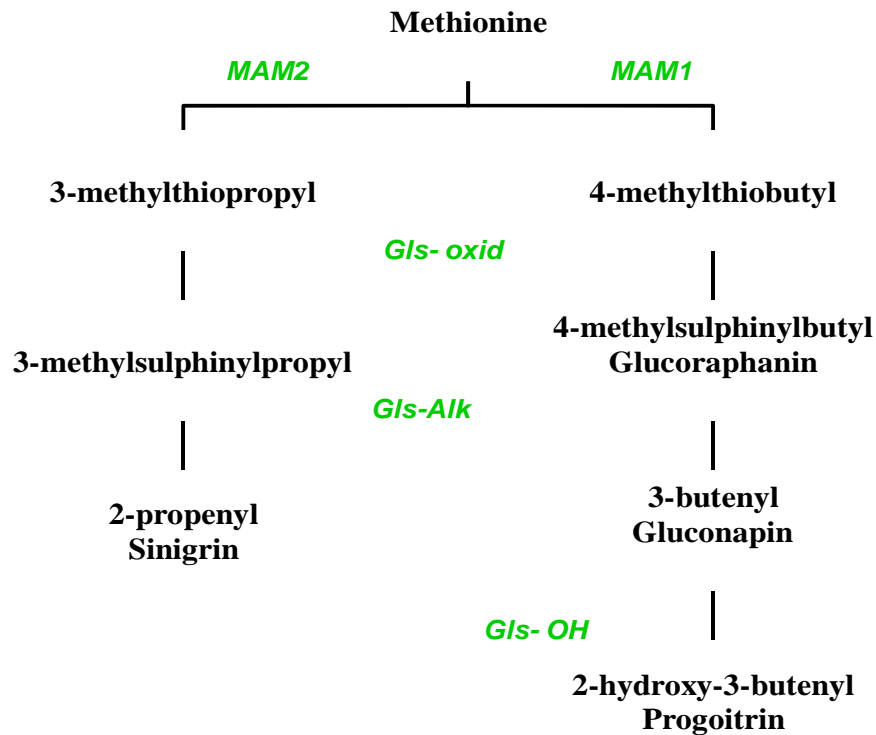
Figure 44 The biosynthetic pathway of aliphatic glucosinolates synthesis identified in the AGDH plant lines based on (Magrath et al., 1994), showed the elongation phase for methionine regulated by *MAM1* and *MAM2* genes, which control the R side chain length of glucosinolates. Further secondary modifications for the R side chain structure were under genetic control of *Gls-oxid, Gls-ALK* and *Gls-OH,* resulting in the observed diversity of aliphatic glucosinolate profiles

After the core structure formation, further secondary modifications on the R side chain for the aliphatic glucosinolates results in the production of the different individual glucosinolates synthesed within this biosynthetic pathway, which were known to be under genetic control.

Within this phase different chemical reactions can occur catalysed by enzymes, encoded by genes which were identified using QTL mapping in *Arabidopsis;* resulting in the identification of several loci involved in the modification step. First,

Gls-oxid, found to control the oxidation of the side chain of the thiol group and consequently the production of 3-methylsulfinylpropyl glucosinolate and 4-methylsulfinylbutyl glucosinolate (glucoraphanin) (Hansen et al., 2007). Second, the *Gls-ALK* locus involved in the formation of an alkenyl bond resulting in the synthesis of 2-propenyl glucosinolate (sinigrin) and 3-butenyl glucosinolate (gluconapin), and finally, the *Gls-OH* locus which controls the production of the 2-hydroxy-3-butenyl glucosinolate (progoitrin) (Kliebenstein et al., 2001b).

Three indolic glucosinolates were identified in the AGDH population. Their biosynthetic pathway was predicted as shown in Figure 45. Previous studies showed that glucobrassicin derived from tryptophan is the precursor for 4-methoxyglucobrassicin and neoglucobrassicin synthesis, within two different branches (Pfalz et al., 2009).

➤ Within the first branch, glucobrassicin is converted to neoglucobrassicin by the addition of a methoxy group at the nitrogen of the indole ring.

➤ In the second branch of the pathway; glucobrassicin is converted to 4-hydroxyglucobrassicin by the hydroxylation of the indole ring at position 4 and subsequently, the production of 4-methoxyglucobrassicin by methylation of the hydroxyl group which is catalyzed by an as yet unidentified enzyme (Pfalz et al., 2009; Zang et al., 2009).

These two sub-branches within the indolic glucosinolate synthesis pathway were independently regulated. A single gene identified as *CYP81F2* (encoding cytochrome P450 monooxygenase) has been mapped at the bottom of chromosome 5 in *Arabidosis,* this gene can affect the level of both 4-hydroxyglucobrassicin and 4-methoxyglucobrassicin, by regulating the hydroxylation reaction on glucobrassicin,

with no effect on the synthesis of neoglucobrassicin (Bednarek et al., 2009; Pfalz et al., 2009).

Tryptophan

|

3-indolylmethylglucosinolates
Glucobrassicin

CYP81F2

4-hydroxy-3-indolyl methylglucosinolate

1-methoxy-3-indolyl methylglucosinolate
Neoglucobrassicin

|

4-methoxy-3-indolyl methylglucosinolate
4-Methoxyglucobrassicin
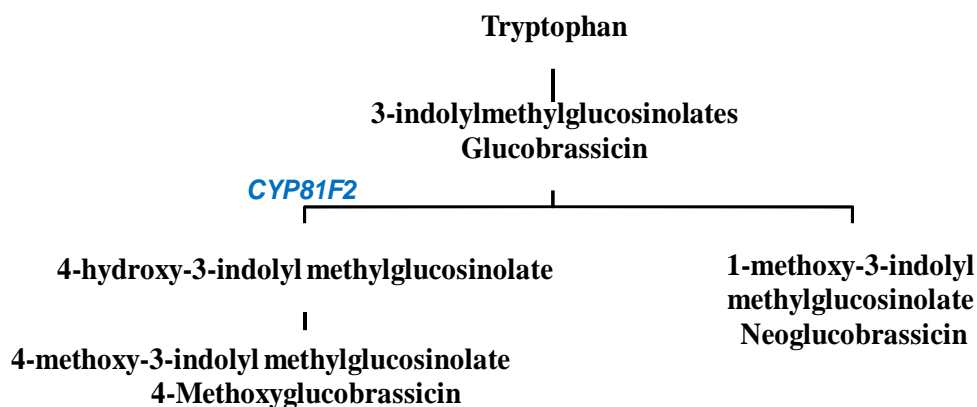
Figure 45 The biosynthetic pathway for indolic glucosinolates identified in the AGDH population based on (Pfalz et al., 2009), showed the genes regulating the modification of glucobrassicin (*CYP81F2)*, which control the subsequent production of either neoglucobrassicin or 4-methoxyglucobrassicin (Pfalz et al., 2009; Zang et al., 2009).

The composition of the glucosinolates can drastically vary in different plant species, depending on the transcription level of the genes encoding enzymes of the glucosinolate biosynthesis where other metabolites and co-factors are known to co-regulate the glucosinolate metabolic pathway. Consequently, the presence of factors regulating the glucosinolate biosynthtic pathway in the AGDH population can be expected.

Within the *MYB* family of gene regulators are known to be factors that control the transcription level of genes involved in the biosynthesis of glucosinolates, including *MYB28, MYB29* and *MYB76*, these are known to positively regulate the accumulation of aliphatic glucosinolates and their transcripts (Gigolashvili et al., 2007b). *MYB29* and *MYB76* control short chain aliphatic glucosinolates, while *MYB28* controls short

and long chain aliphatic glucosinolates (Sonderby et al., 2007). In contrast, it has been proposed that *MYB29* has no role in the regulation of aliphatic glucosinolates except in response to jasmonic acid (Hirai et al., 2007). QTLs for these regulatory genes have been mapped in *Arabidopsis* on chromosome 5 (Kliebenstein, 2009).

The transcription factors *MYB34, MYB51* and *MYB122* were identified in *B. rapa* and in *Arabidopsis* to regulate genes involved in the biosynthesis of indolic glucosinolates (Halkier and Gershenzon, 2006; Sonderby et al., 2007; Zang et al., 2009). In *Arabidopsis MYB34* has been mapped on chromosome 5, while *MYB51* and *MYB122* have been mapped on chromosome 1. All these regulators were found to specifically up regulate indolic glucosinolate biosynthetic genes via different roles (Gigolashvili et al., 2007a; Zang et al., 2009).

Other regulators were found in common between aliphatic and indolic glucosinolate biosynthetic pathways, and have been identified in *Arabidopsis* as follows:

- Sulfur limitation1 (SLIM1) regulates the catabolism of glucosinolates in response to sulfate deficiency (Maruyama-Nakashita et al., 2007).

- IQD1 (nuclear localized calmodulin binding protein) that up-regulates indolic glucosinolate synthesis and down-regulates aliphatic glucosinolate synthesis, has been mapped in *Arabidopsis* on chromosome 3 (Kliebenstein, 2009).

- The third regulator, AtDof1.1 (DNA binding with one finger) that increases the level of synthesis of indolic as well as aliphatic glucosinolates in response to wounding and herbivore attack, has been mapped in *Arabidopsis* on chromosome 3 and 1 (Gigolashvili et al., 2009; Kliebenstein, 2009; Maruyama-Nakashita et al., 2007).

The gene regulators involved in the biosynthesis of glucosinolates can be identified as candidate genes underlying QTLs mapped in the AGDH population, using a

comparative analysis approach with other related species that have been previously studied.

3.4.5 Comparison of QTL mapping for glucosinolates on the AGDH genetic map with corresponding regions on the genetic maps of *A. thaliana* and *B. rapa*

The quantitative analysis of glucosinolate content in the AGDH plant lines showed that glucosinolates segregated in the AGDH plant lines in a quantitative manner that lends itself to QTL analysis. Consequently, the genetic basis of glucosinolate content in *B. oleracea* was further resolved using this DH mapping population and the QTL approach. The genetic resources available in *B. oleracea* and exploitation of the synteny with that of *B. rapa* and *A. thaliana* mean that QTL analyses of glucosinolate biosynthesis offers the eventual prospect of identifying and characterizing the genes and gene regulators involved in their synthesis.

Towards the aim for determination of the previously identified factors underlying any QTL, a comparative genomic study for colinear regions between *B. oleracea, A. thaliana* and *B. rapa* was conducted. This approach is widely used to study conserved and rearranged regions between *Brassica* species and *Arabidopsis* (Lukens et al., 2003), in the sense of finding corresponding genes or gene regulators in relatively the same order and orientation, as every *B. oleracea* and *B. rapa* linkage group has significant collinear regions on at least one *A. thaliana* chromosome (Qiu et al., 2009). Despite the evidence of numerous genomic rearrangements, resulting in gene loss, fragmentation and duplications, this approach has been successfully used to aid understanding of factors involved in metabolite biosynthetic pathways (Gao et al., 2004).

In this study, the alignment between the C genome (n=9) of *B. oleracea* and the A genome (n=10) of *B. rapa* linkage groups was based on the linkage maps developed by (Luis Iniguez-Luy et al., 2009), as common loci and putative homologous regions between the two maps were identified through shared markers. The identification of the colinear regions with the At genome (n=5) of *A. thaliana* with these linkage maps was based on the same maps and was in agreement with other comparative analysis as described by (Lukens et al., 2003; Mun et al., 2009).

34.5.1 Comparative analysis of QTLs associated with aliphatic glucosinolates synthesis in the AGDH plant lines

Significant QTLs for gluconapin, progoitrin, and the sum of sinigrin and gluconapin, were co-localized at the middle region on LG3 (Map 4-A) which is co-linear with the top and middle regions on chromosome A3. QTLs controlling the side chain elongation and modification of aliphatic glucosinolates in *B. rapa* leaves, were previously mapped on this segment (Lou et al., 2008), showing co-linearity with the top region of chromosome At5 and the bottom region of chromosome At2, respectively.

The previous identification of *MAM 1, MAM2 and MAM3* genes, in addition to the regulator genes *MYB 29 and MYB76,*found them on the top of At5, while the gene *Gls-OH,* which controls the hydroxylation reaction for side chain modification, was identified at the bottom segment of At2 (Kliebenstein, 2009). This analysis can explain the QTLs mapped in this study, as the *MAM2* and *MAM1* elongation genes were involved in the synthesis of 3 and 4 carbons of the R side chains and resulting in the production of singrin and gluconapin, respectively. However, gluconapin can undergo further side chain modification reactions to produce progoitrin, via a

hydroxylation reaction controlled by the *Gls-OH* gene (as described previously in section 3.4.4).
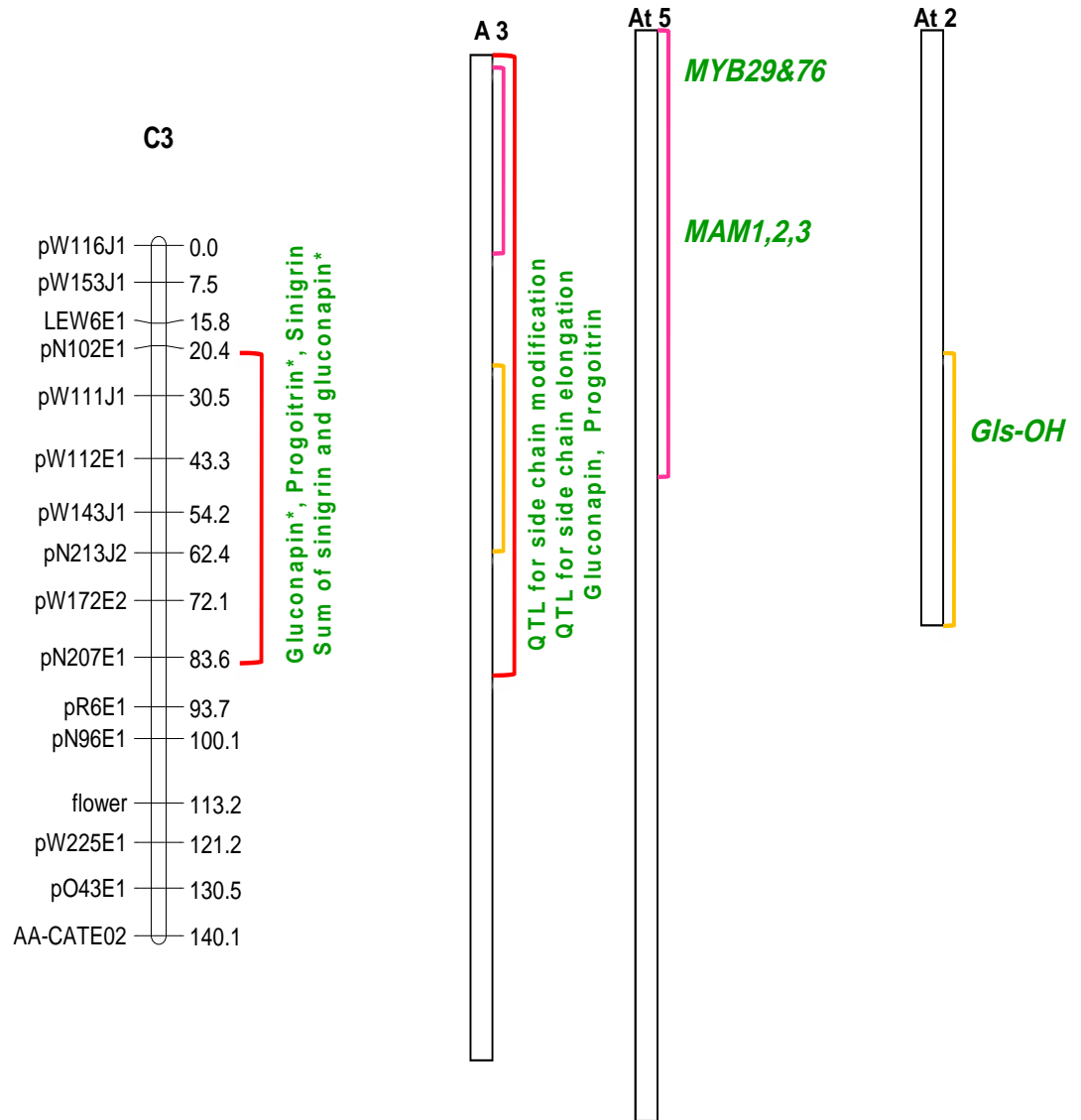
Map 4-A *Brassica oleracea* linkage map based on AGDH population (Unpublished revision of the Sebastian et al (2000) Integrated map by Graham Teakle), with QTLs for aliphatic glucosinolates. Alignment of conserved areas between the *B. oleracea* linkage map (C1-C9) and *B. rapa* map (A1-A10) with the *A. thaliana* map (At1-At5), shown in brackets to the right of each LG indicate the homologous segments between the three maps recognized by their colours, where a QTL was observed in *B. oleracea* and/ or *B. rapa*. All known genes controlling aliphatic glucosinolate content underlay the observed QTLs previously identified in the *A. thaliana* genome are shown next to the brackets. Markers positions in cM are shown to the left and to the right of each LG, respectively. Significant QTL determined at *(p≤ 0.05)

At the upper part of LG 4, a QTL for progoitrin was mapped on the region co-linear with the middle of chromosome A4 (Map 4-B). Previously a QTL for progoitrin level in leaves has been mapped to this region in *B. rapa* (Lou et al., 2008) and was explained by the presence of the *Gls-OH* gene responsible for hydroxyl group addition in the R side chain of progoitrin, which has been found at the bottom of chromosome At2.
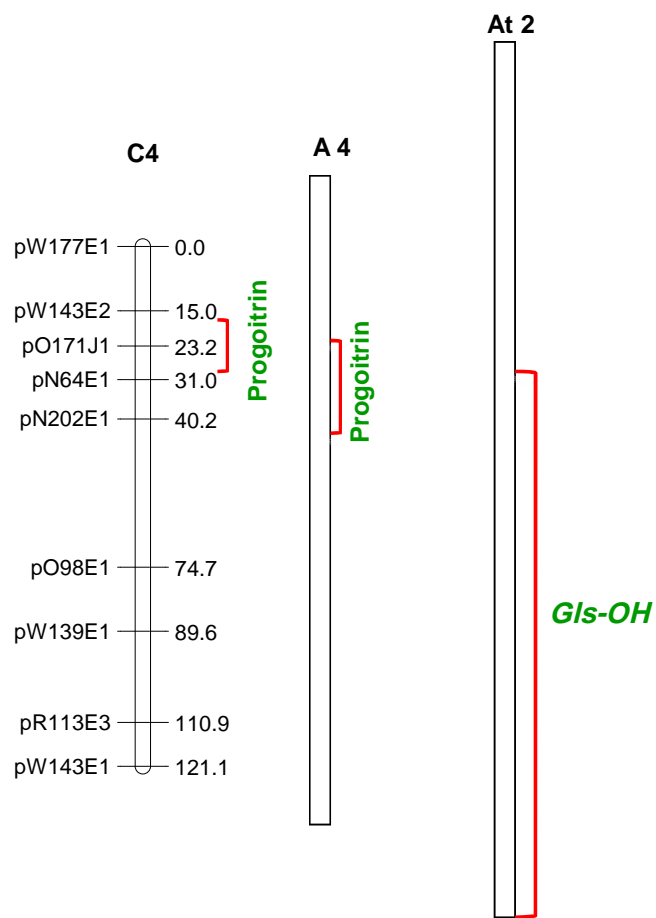
Continue Map 4-B

A *novel* QTL for sinigrin mapped to the middle region on LG5 (Map 4-C), which is co-linear with the upper-middle region on chromosome A5. Although a QTL for side modification of aliphatic glucosinolates has been mapped to this region for control of progoitrin and gluconapin in *B. rapa* leaves (Lou et al., 2008), no genes known to control sinigrin synthesis have been previously identified. This region is co-linear with the top region on chromosome At3 where the regulator IQD has been identified to down regulate aliphatic glucosinolate content (Gigolashvili et al., 2009; Kliebenstein, 2009; Maruyama-Nakashita et al., 2007). This raises the strong possibility that these loci control the synthesis of sinigrin.
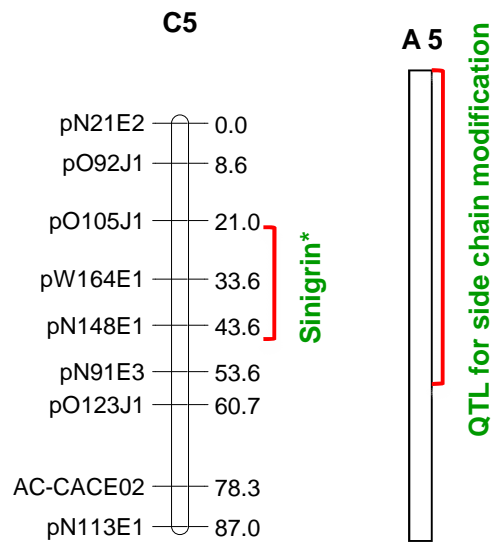
Continue Map 4-C

Significant QTLs for the total aliphatic glucosinolate content, the sum of glucoraphanin and progoitrin content, and for progoitrin were co-localized at the top region of LG7 (Map 4-D). This is known to be co-linear with the middle region on chromosome A6. A QTL for the total aliphatic glucosinolates; [the sum of progoitrin, gluconapin and glucobrassicanapin (an aliphatic glucosinolate with 5 carbon side chain)] in leaves were previously identified in this region in *B. rapa* (Lou et al., 2008). These findings suggest the presence of genes that control the R side chain modification, underly the QTLs mapped in this study.
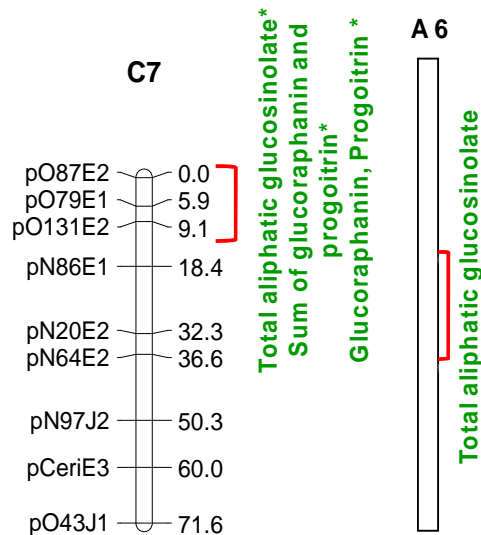
Continue Map 4-D

In this study, significant QTLs for the content of the total aliphatic glucosinolates, the sum of glucoraphanin and progoitrin, the sum of sinigrin and gluconapin and for gluconapin were mapped at the bottom region on LG8 (Map 4-E), which is co-linear with the middle region of chromosome A9. In this region, QTLs for gluconapin and progoitrin in leaves of *B. rapa* were previously mapped as described by (Lou et al., 2008). However; at the co-linear region on the middle of chromosome At1, genes controlling the core structure formation of aliphatic glucosinolates *(CYP79F1/F2)* have been identified in *Arabidopsis* (Kliebenstein, 2009). All aliphatic glucosinolates share this step in their biosynthetic pathway (as described in Chapter 1), the presence of genes controlling the synthesis of glucosinolates with different R side chain lengths at this QTL is possible.
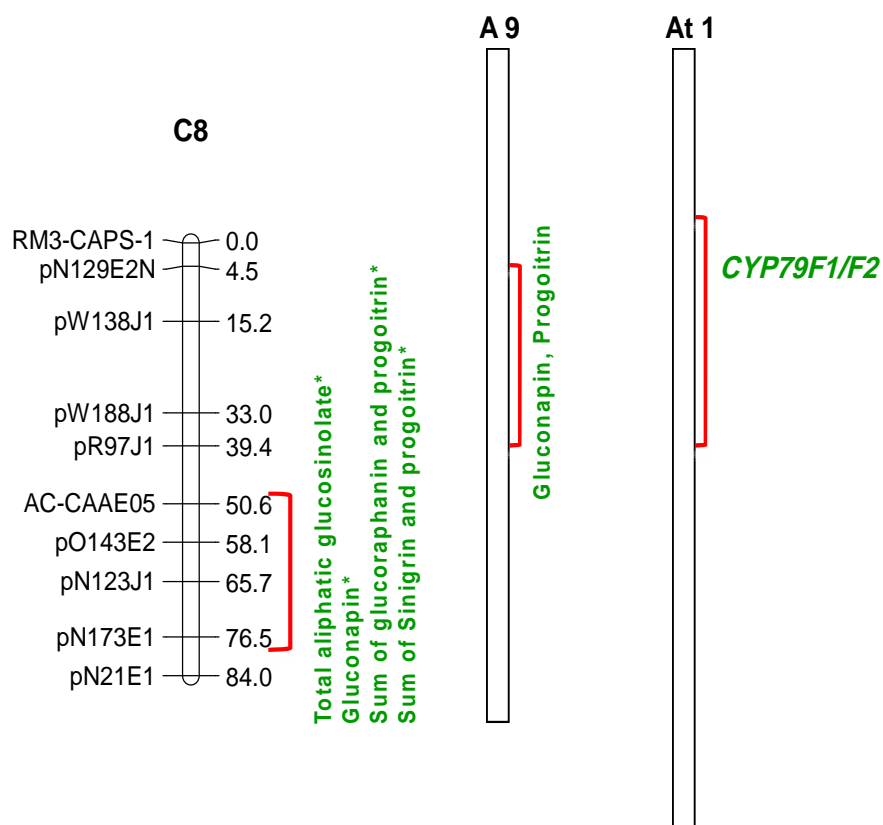
Continue Map 4-E

Within the top region at LG9 (Map 4-F), significant QTLs were mapped for the content of total aliphatic glucosinolates, the sum of glucoraphanin and progoitrin, the sum of sinigrin and gluconapin, sinigrin, glucoraphanin and gluconapin. QTLs mapped in this study, were confirmed by comparative analysis with the colinear region at the top of chromosome A10, where QTLs for the content of the sum of progoitrin, gluconapin and glucobrassicanipin, and for the individuals; gluconapin and progoitrin have been mapped in *B. rapa* as described by (Lou et al., 2008), but QTL for glucoraphanin and sinigrin were *novel*. This can be explained by the presence of *MAM1, MAM2* and *MAM3* genes. In addition, the regulator genes *MYB29* and *MYB76,* have been identified within the corresponding co-linear region at the top region of chromosome At5 in *Arabidopsis* (Kliebenstein, 2009).

C9

Marker	Position
pN52E2	0.0
pO125E1N	12.9
pW137J1	23.4
pW114E2	37.1
pO119J1	43.8
pW233J1	56.4
pO160E1	69.9
pO7E1	81.7
pN47E4NM	89.6
pW200J1	106.8

Total aliphatic glucosinolate*
Sum of glucoraphanin and progoitrin*
Sum of sinigrin and gluconapin**
Gluconapin**, Sinigrin*, Glucoraphanin*

A10

Sum of progoitrin, gluconapin and Glucobrassicanipin
Progoitrin, Gluconapin

At5

MYB29&76

MAM1,2,3

Continue Map 4-F

QTL mapping for glucosinolates expected to be under the control of the major gene effects as described under section 3.4.3.2), are shown in Map 5.

In this study, a *novel* QTL for sinigrin was mapped at the bottom region on LG2 (Map 5-A), this has not been previously identified in *Brassica*. This region shows co-linearity with the bottom region at chromosome At5, where the regulator gene *MYB28* has been identified (Kliebenstein, 2009). Therefore, genes controlling the synthesis of sinigrin at this locus are possible.
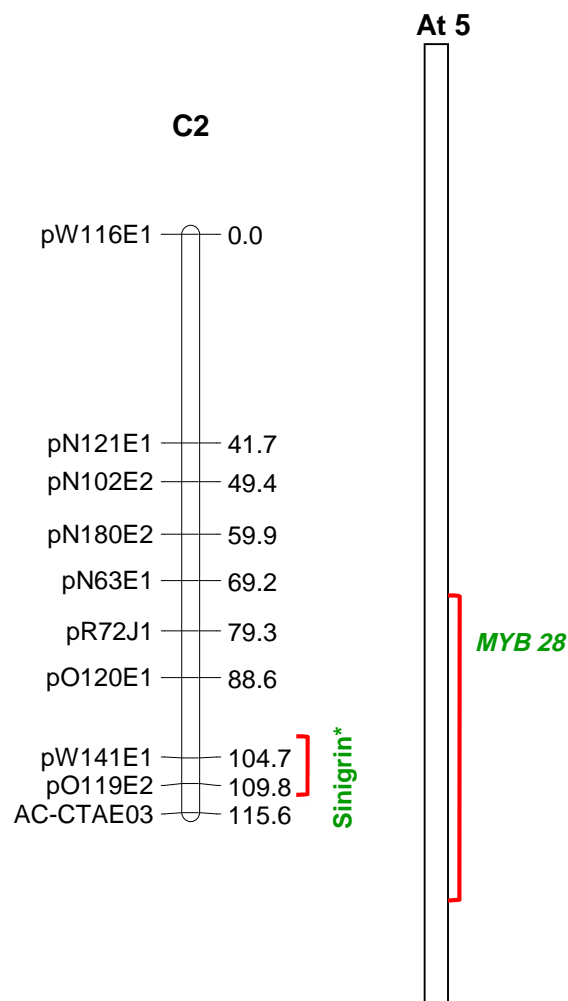
Map 5-A *Brassica oleracea* linkage map based on AGDH population (Unpublished revision of the Sebastian et al (2000) Integrated map by Graham Teakle), with QTLs for aliphatic glucosinoltes, that were expected to be under the control of major gene effect. Alignment of conserved areas between the *B. oleracea* linkage map (C1-C9) and *B. rapa* map (A1-A10) with the *A. thaliana* map (At1-At5), shown in brackets to the right of each LG indicate the homologous segments between the three maps recognized by their colours, where a QTL was observed in *B. oleracea* and/ or *B. rapa*. All known genes control aliphatic glucosinolates content underlay the observed QTLs previously identified in *A.thaliana* genome are shown next to the brackets. Markers positions in cM are shown to the left and to the right of each LG, respectively. Significant QTL determined at *(p≤ 0.05)

167

In this study, three QTLs for progoitrin, progoitrin-DR, and sinigrin-DR were co-localized at the middle region of LG3 (Map 5-B). Only the QTL controlling the content of progoitrin was significant. In addition, another QTL for progoitrin-DR were mapped to the middle segment on LG4 (Map 5-C). Using comparative analysis with the *B. rapa* genome, QTLs on LG3 show co-linearity with the top-middle region of the A3 LG, while QTL on LG4 shows co-linearity with the middle region of the A4 LG. Comparing these results with other research findings, QTLs controlling the content of progoitrin, the side chain modification step and the side chain elongation phase, have been mapped to A3 and A4 LG in *B. rapa* leaves as described by (Lou et al., 2008).

Interestingly, the top region of the A3 LG shows colinearity with the top region of the At5 LG in *Arabidopsis*, where *MAM 1, MAM2* and *MAM3* genes were previously mapped, and were known to regulate the R side chain length of aliphatic glucosinolates prior to the core structure formation of sinigrin (3 carbons) and progoitrin (4 carbons) (as described in Figure 44). In addition, the gene regulators *MYB 29* and *MYB76,* known to regulate the synthesis of aliphatic glucosinolates, have been mapped to the same region (Kliebenstein, 2009).

The middle region of A3 LG shows co-linearity with the bottom region of the At2 LG, where the *Gls-OH* gene controlling the conversion of gluconapin into progoitrin, was previously identified (Kliebenstein, 2009). Therefore, QTLs for aliphatic glucosinolates with different R side chain lengths and structures mapped at this region in our work are in agreement with previous studies.

C3

Marker	Position
pW116J1	0.0
pW153J1	7.5
LEW6E1	15.8
pN102E1	20.4
pW111J1	30.5
pW112E1	43.3
pW143J1	54.2
pN213J2	62.4
pW172E2	72.1
pN207E1	83.6
pR6E1	93.7
pN96E1	100.1
flower	113.2
pW225E1	121.2
pO43E1	130.5
AA-CATE02	140.1

Progoitrin*, Progoitrin-DR , Sinigrin-DR

A 3

QTL for side chain modification
QTL for side chain elongation
Progoitrin

At 5

MYB29&76

MAM1,2,3

At 2

Gls-OH

Continue Map 5-B

C4

pW177E1 — 0.0

pW143E2 — 15.0
pO171J1 — 23.2
pN64E1 — 31.0

pN202E1 — 40.2

pO98E1 — 74.7

pW139E1 — 89.6

pR113E3 — 110.9
pW143E1 — 121.1

Progoitrin-DR

A 4

Progoitrin

At 2

Gls-OH

Continue Map 5-C

In this study, two significant QTLs for sinigrin and sinigrin-DR were co-localized at the top-middle region of LG5 (Map 5-D), which shows co-linearity with the top-middle region at A5 chromosome. However, other research findings identified QTL for side chain modification for progoitrin and gluconapin at this region in *B. rapa* (Lou et al., 2008), indicating the QTL for sinigrin is *novel*. This region is found to be co-linear with the top region of At3 LG, where the regulator IQD1 has been mapped which down regulates aliphatic glucosinolate biosynthesis (Kliebenstein, 2009).
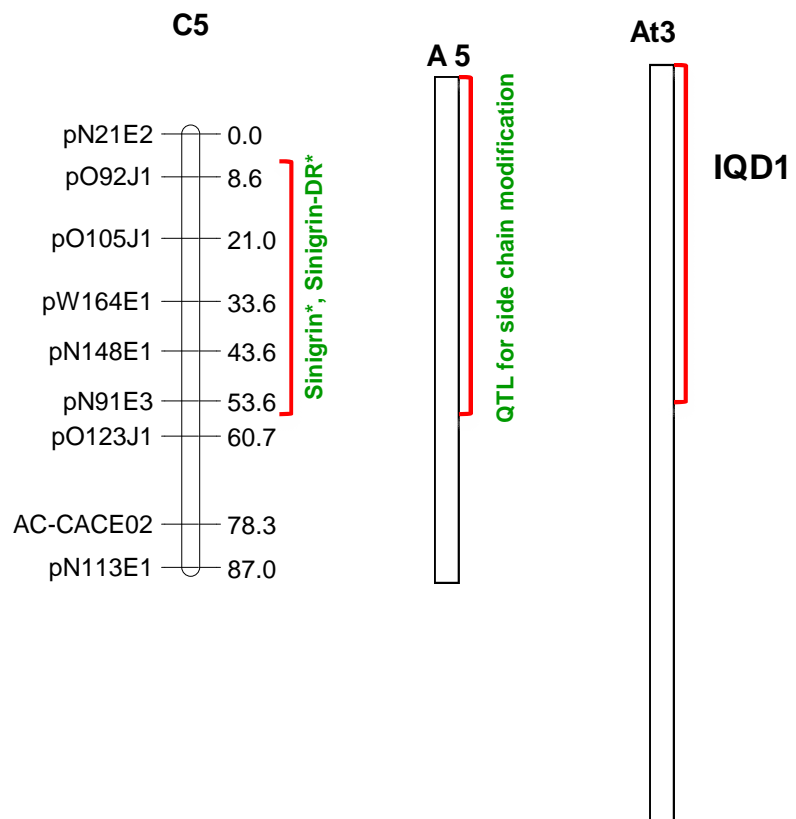
Continue Map 5-D

A significant QTL for gluconapin and a non-significant QTL for progoitrin, were co-localized at the bottom segment of LG7 (Map 5-E), which was co-linear with the middle region of the A7 LG. These results were in agreement with (Lou et al., 2008), where they had previously mapped a QTL for gluconapin on the co-linear region in *B. rapa* leaves. However, no indications for the presence of genes regulating these QTLs were found in any other *Brassica* species.
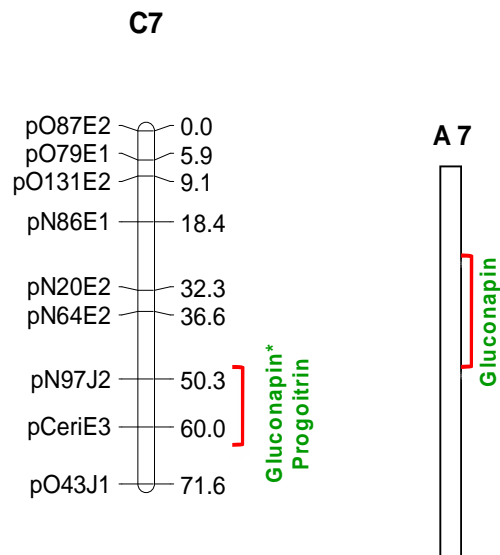
Continue Map 5-E

Interestingly, we were able to confirm a significant QTL for progoitrin which was mapped at the middle region of LG9 (Map 5-F) and shows co-linearity with the middle region of A10 LG in *A. thaliana*. These results agreed with (Lou et al., 2008), where they mapped a QTL for progoitrin to the same region in *B. rapa*. In addition, the *Gls-OH* gene, whose phenotype is the presence, or absence of progitrin, has been mapped to lie within this region by (Gao et al., 2007).
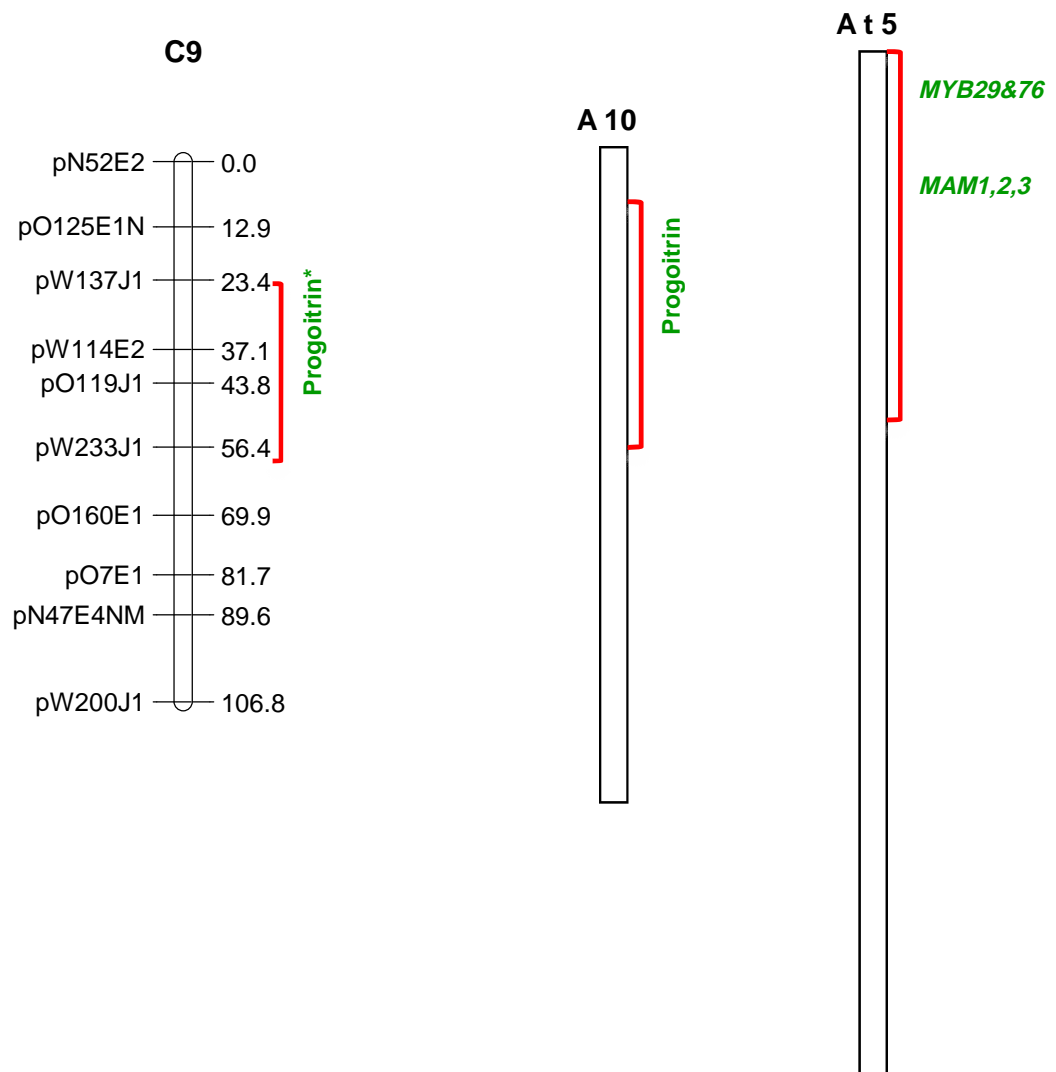
Continue Map 5-F

3.4.5.2 **Comparative analysis of QTLs associated with indolic glucosinolates in the AGDH plant lines**

In this study, two QTLs controlling the content of total indolic glucosinolates and glucobrassicin were co-localized at the bottom region of LG1 (Map 6-A). These QTLs have been identified within the colinear region at the bottom of At4 chromosome using recombinant inbred lines of the L*er* X Cvi mapping population (Kliebenstein et al., 2001a). In their study, they identified the gene controlling the core structure formation of indolic glucosinolates (*CYP83B1)* at the same region, where genes regulating the synthesis of different individual indolic glucosinolate are not known yet.
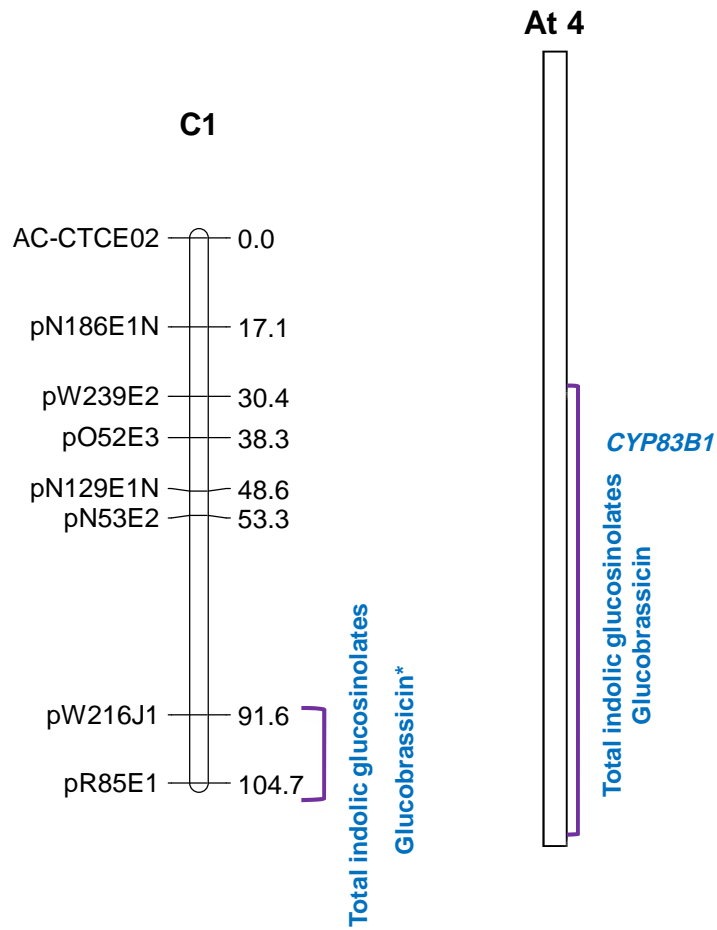
Map 6-A *Brassica oleracea* linkage map based on AGDH population (Unpublished revision of the Sebastian et al (2000) Integrated map by Graham Teakle). Alignment of the conserved areas between the *B. oleracea* linkage map (C1-C9) and *B. rapa* map (A1-A10) with the *A. thaliana* map (At1-At5) shown in brackets to the right of each LG indicate the homologous segments between the three maps recognized by their colours, with the QTLs observed in *B. oleracea* and/ or *B. rapa*. All known genes controlling indolic glucosinolates content underlay the observed QTLs previously identified in *A. thaliana* genome are shown next to the brackets. Markers positions in cM are shown to the left and to the right of each LG, respectively. Significant QTL determined at *(p≤ 0.05) and ** (p≤ 0.001)

175

As expected, a significant QTL controlling the content of total indolic glucosinolates mapped at the top region on LG2 (Map 6-B), this has been identified in the co-linear region on the top of A2 in *B. rapa* (Lou et al., 2008) and on the At5 chromosome in L*er* X Cvi recombinant inbred line (Kliebenstein et al., 2001a).
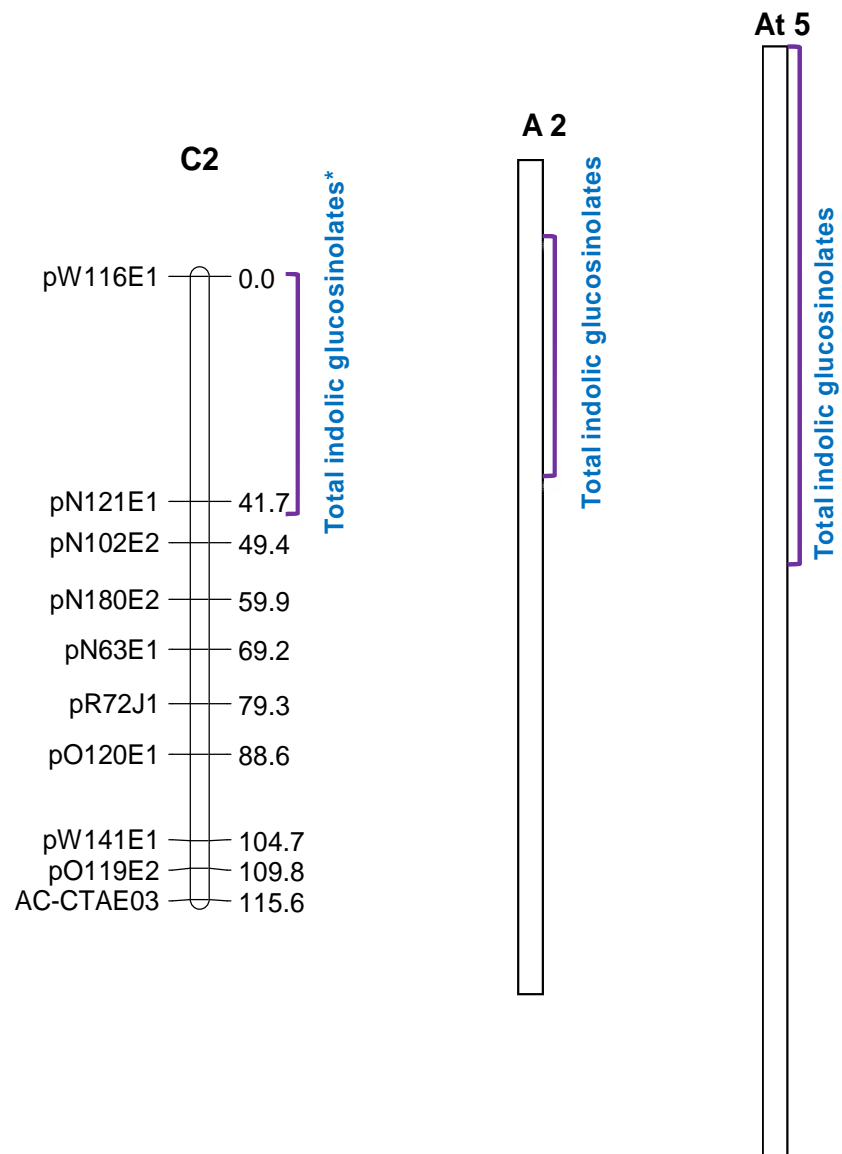
Continue Map 6-B

This study revealed the presence of a *novel* QTL for neoglucobrassicin mapped at the bottom region of LG3 (Map 6-C). Interestingly, this *novel* QTL mapped at the same position as the gene *CYP83B1* that controls the core structure formation of indolic glucosinolate biosynthesis, which has been mapped at the co-linear region on the bottom of At4 chromosome in *Arabidopsis* (Kliebenstein, 2009).

C3

Marker	Position
pW116J1	0.0
pW153J1	7.5
LEW6E1	15.8
pN102E1	20.4
pW111J1	30.5
pW112E1	43.3
pW143J1	54.2
pN213J2	62.4
pW172E2	72.1
pN207E1	83.6
pR6E1	93.7
pN96E1	100.1
flower	113.2
pW225E1	121.2
pO43E1	130.5
AA-CATE02	140.1

Neouglucobrassicin

At 4

CYP83B1
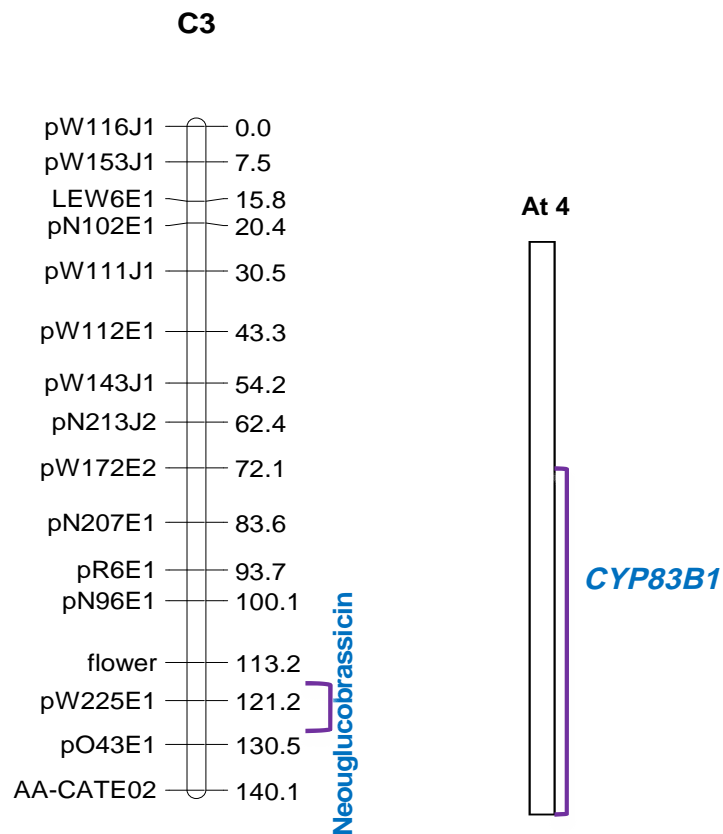
Continue Map 6-C

A *novel* QTL for neoglucobrassicin was mapped at the bottom region on LG4 (Map 6-D). This QTL has not been mapped with the co-linear region at the bottom of A2 LG in *Brassica* leaves. However, a QTL controlling glucosinolate content in seed has been identified at the top region of At2 in L*er* X Cvi recombinant inbred lines (Kliebenstein et al., 2001a).
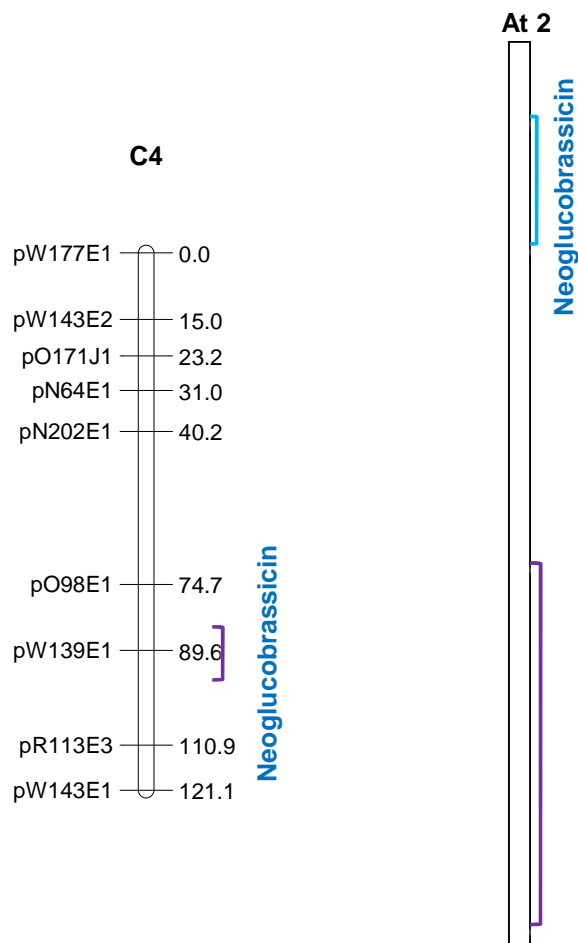
Continue Map 6-D

A third *novel* QTL for neoglucobrassicin was mapped to the top region on LG5 (Map 6-E), this QTL has not been mapped in *Brassica* species. By applying the co-linearity analysis with the other genetically related species, we were not able to identify any candidate genes underlying this QTL.

C5
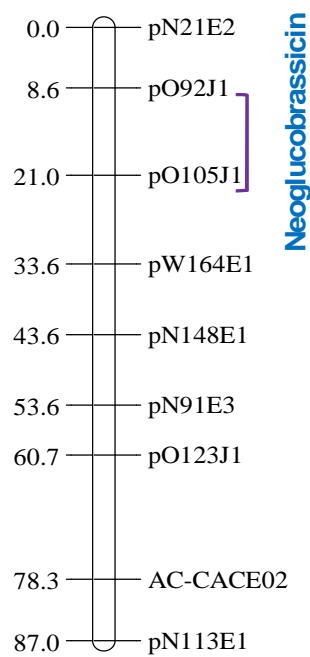
Continue Map 6-E

A *novel* QTL for glucobrassicin mapped at the top of LG7 (Map 6-F). The genes regulating the synthesis of different individual indolic glucosinolate are not known yet, and therefore the identification of candidate genes that control the synthesis of glucobrassicin at this position is not at present possible.

C7

Marker	Position
pO87E2	0.0
pO79E1	5.9
pO131E2	9.1
pN86E1	18.4
pN20E2	32.3
pN64E2	36.6
pN97J2	50.3
pCeriE3	60.0
pO43J1	71.6
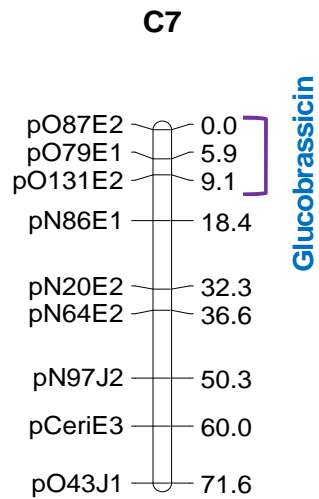
Glucobrassicin

Continue Map 6-F

It was not surprising to map two co-localized QTLs controlling the content of the total indolic glucosinolates and glucobrassicin in the middle region on LG9 (Map 6-G). QTL for total indolic glucosinolates and for glucobrassicin in leaves and seeds have been mapped on the co-linear region at the middle of At5 LG in a recombinant inbred line from the cross L*er* X Cvi (Kliebenstein et al., 2001a). The gene (*CYP81F2),* which is known to regulate the synthesis of 4-methoxyglucobrassicin has also been mapped to this region by (Bednarek et al., 2009; Pfalz et al., 2009), an equivilant QTL for 4-methoxyglucobrassicin synthesis was not mapped in our study.
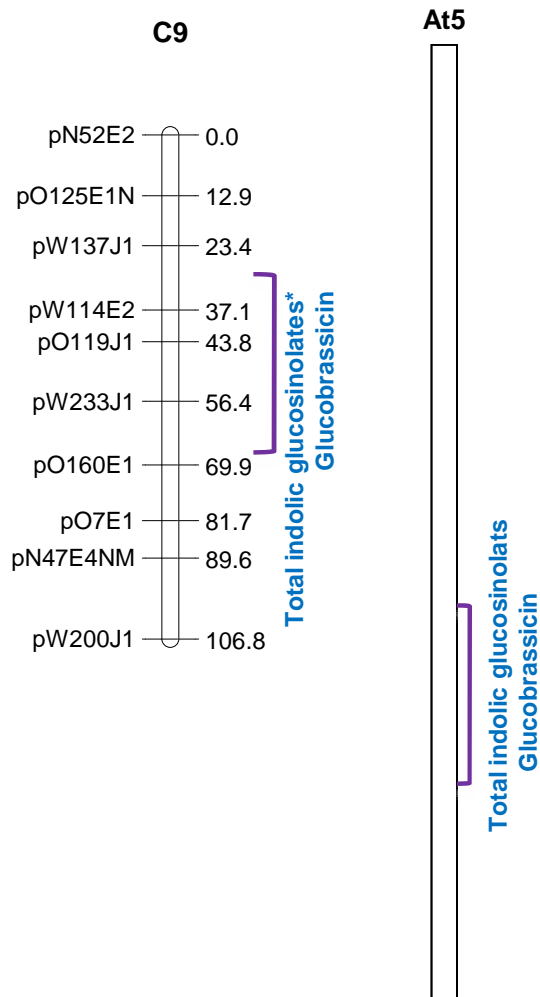
Continue Map 6-G

3.5. Conclusion

- The prediction of the biosynthetic pathways for 4 aliphatic and 3 indolic glucosinolates analysed in the AGDH populations and the identification of the genes and gene regulators controling their synthesis, was achieved based on known pathways combined with genetic analysis of quantitative data for metabolites.

- QTLs for gluconapin and progoitrin, at the top region of LG9, coincided with the known position of two genes, *Gls-OH* and *Gls-ALK* previously shown to have a major effect on aliphatic glucosinolate content.

- Three *novel* QTLs controlling sinigrin synthesis were identified in the AGDH genome, located on LG2, LG5 and LG9, while one *novel* QTL controlling glucoraphanin was located on LG9.

- Four *novel* QTLs controlling indolic glucosinolate synthesis were identified in the AGDH genome, of which three QTLs controlled neoglucobrassicin synthesis located on LG3, LG4 and LG5, while one QTL controlling glucobrassicin synthesis was found on LG7.

Summary of the project and its finding

The aim of this project was to increase understanding of the genetic control of glucosinolate biosynthesis in *B. oleracea*. Complementary approaches were implemented: as a first step; phenotype variation was assessed and investigated by analysis of glucosinolate profiles using *HPLC-UV/ESI-MS/MS* to analyse desulfoglucosinolates. Secondly; QTL mapping was performed using CIM and IM analysis. Finally, comparative genomic analysis of the significant QTLs with the *B. rapa* and *A. thaliana* genomes in order to determine where possible, potential candidate gene(s) within the QTL confidence interval that may determine the trait of interest.

4.1 Identification of glucosinolate profiles in the AGDH population

As glucosinolates are highly charged molecules, the *HPLC-MS* method used for extraction and analysis of intact glucosinolates was unsuccessful. However, a method was successfully developed for the analysis of desulfated glucosinolates, providing effective extraction and analysis of the glucosinolate from the plant extract in order to determine the glucosinolate content using *HPLC-UV/ESI-MS/MS* analysis.

A mass inclusion list was used for the identification of desulfated glucosinolates, which limited the glucosinolates observed to those on the list based on what is known in the literature. Alternatively, developing a natural loss driven acquisition method whereby novel glucosinolates could be observed, may reveal the presence of glucosinolates not previously identified in *Brassica*.

Despite the reproducibility of the desulfation method developed in this study and the high efficiency of the *HPLC-UV/ESI-MS-MS/MS* method used for separation and identification of desulfated glucosinolates, this protocol has limitations over the methods used for intact glucosinolate analysis. This protocol is time consuming, as for optimum desulfation; reaction for 18 hrs was required at defined working conditions. Moreover, optimization procedures must be applied when the experimental parameters or desulfation conditions were significantly modified. For example, the nature of the sample, type and amount of glucosinolates in the extract (including IS1), activity of the sulfatase enzyme, size of the ion exchange column, time and temperature of the enzymatic incubation.

For quantification of glucosinolate content in plant extracts relative to an IS, RRF were used to correct for differences in the *UV* absorbance between IS1 and desulfated glucosinolates. In this study, RRF were used as described in the standardised protocol (EEC, 1990) calculated relative to desulfosinigrin rather than desulfoglucotropaeolin (IS1). These glucosinolates are of different chemical structures and therefore, their maximum *UV* absorbance can vary significantly. Therefore, in the presence of pure glucosinolate reference materials, it would provide the tools for researchers to calculate response factors for the detected glucosinolates relative to an IS that they choose to work with. This would significantly improve the quantitative analysis of glucosinolate content. In addition, I presented the first use of a second internal standard (IS2). The peak are of IS2 was used as a base peak to correct for the variations in the autosampler injection volum, in order to improve the reproducibility of the quantitative measurements.

Improving the methods used for studying glucosinolate profiles from natural resources, in order to screen large number of samples for routine analysis using different protocols for sample preparation and analysis have been undertaken in the history of glucosinolate research. The need to commercialize pure standard materials of both intact and desulfated glucosinolates, through the synthesis of their analogues in large scales for research purposes is a promising approach that needs to be improved. By providing these resources, researchers can then use more advanced methods for analysis of glucosinolates with more reliable and accurate results. For example: the analysis of intact glucosinolate in order to obtain more accurate results and to decrease the time and laboratory work needed to desulfate glucosinolate during sample preparation prior to analysis. The quantification of glucosinolate profile for its absolute content rather than relative content using pure standard materials representing each glucosinolate in the sample, would improve the quantitative data obtained significantly.

4.2 Search for QTL affecting glucosinolates

The initial analysis showed that the plant lines making up the AGDH mapping population synthesis two different classes of glucosinolates with indolic and aliphatic side chains. No aromatic glucosinolate were observed in the extracts of the AGDH population. By comparing the average relative concentrations for a pair of glucosinolates within the same class or between the two different classes of aliphatic and indolic glucosinolates, using scatter plots matrix (Figure 42), it was possible to predict the putative biosynthetic pathways for these two distinct glucosinolate classes analysed in the AGDH population. This provided the basic to investigate the genetic factors controlling their content.

Glucosinolate content was found to segregate in the AGDH mapping population and the variation followed a continuous distribution with different contribution from the parental lines, which made this plant population suitable for QTL mapping.

All presented QTLs in this work were identified using the Win QTL Cart program; the Map QTL program confirmed these findings but could not resolve further significant regions controlling glucosinolate synthesis. Comparative linear analysis with the genomes of the related species *B. rapa* and *A. thaliana*, enabled the identification of potential genes and gene regulators as candidates for the control of the biosynthesis of glucosinolate at different levels of chain elongation, core structure formation and side chain modification, and allowed allocation to positions on the C genome when possible.

4.2.1 Identification of potential key genetic regions on the C genome controlling glucosinolates content

In this study, a number of *novel* QTLs controlling the content of glucosinolates were reported (Table 25):

- Three *novel* QTLs controlling the content of sinigrin were allocated on LG2, LG5 and LG9.

- One *novel* QTL controlling the content of glucoraphanin was allocated on LG9.

- Three *novel* QTLs controlling the content of neoglucobrassicin were allocated on LG3, LG4 and LG5.

- One *novel* QTL controlling the content of glucobrassicin was allocated on LG7.

186

The QTL identified on LG3, LG4, LG5, LG7 and LG9 were in *common* for both aliphatic and indolic glucosinolate production (Table 25). It is well documented that a common set of enzymes are involved in the core structure formation of all glucosinolates (Halkier and Gershenzon, 2006; Mithen, 2001). Our results obtained from the scatter plot matrix analysis (Figure 42), where positive linear relationships between pairs of glucosinolates from different classes were observed, confirmed this relationship. The detection of loci that control both aliphatic and indolic glucosinolate production is therefore not surprising and fits with the known biosynthesis.

Comparative analysis investigating QTLs that have been identified in *B. oleracea*, *B. rapa*, and *A. thaliana* re-inforced that the QTLs observed in this study were important in determining glucosinolate content and were in agreement with other published results (Gao et al., 2007; Hall et al., 2001; Kliebenstein, 2009; Kliebenstein et al., 2001a; Lou et al., 2008) (Table 25). In addition, previously identified genes and gene regulators in *Arabidopsis* provided potential genes underlying the QTLs which mapped to the same positions.

Table 25 Summary of QTLs mapped on the AGDH LG1-9 using Win QTL Cart. and CIM analysis assorted by linkage groups and were named by the glucosinolate QTL number. Alignment of theses QTLs on previously studied QTL for similar traits and genes or gene regulators underlying these QTLs in *B. rapa* and *A. thaliana* are shown.

Blue and green are for indolic and aliphatic glucosinolates respectively. Red for potential genes suggested at QTLs positions control the synthesis of similar traits previously identified at collinear regions in *Arabidopsis*. Bold, confirmed QTL by comparative analysis with previously mapped QTL for the same trait in other related species. Italic, QTL where the genes regulating the trait are know. * Novel QTL, not mapped

LG	AGDH plant lines QTL	Potenial genes	*B. oleracea*	*B. rapa*	*A. thaliana*
LG1	*Brassicin1* **Total indo 1**				Total indolic glucosinolate *CYP83B1*
LG2	**Total indo 2**	*CYP83B1*		Total indolic glucosinolate	Total indolic glucosinolate
LG3	**Prog1** **Napin1** *Sinig +Napin1* *Neoubr 1**	*ALK* gene		QTL for side chain elongation QTL for side chain modification Gluconapin progoitrin	*MAM 1,2&3* *MYB29&76* *Gls-OH* *CYP83B1*
LG4	Neoubr 2* **Prog 2**	*CYP83B1* *MAM/ ALK* gene		progoitrin	*Gls-OH*
LG5	Neoubr 3* Sinig 1*	*CYP83B1* *MAM / ALK* gene		QTL for side chain modification	
LG7	Prog 3 **Total Ali 1** Raphanin + Prog 1 Brassicin 2*	*MAM/ Gls-OH/* **ALK** gene *CYP79F1/F2* *CYP83B1*		Total aliphatic glucosinolates	
LG8	**Napin 2** *Total Ali 2* *Raphanin +Prog 2* *Sinig+ Napin 2*	**ALK** gene *Gls-OH* *MAM1,2&3*		Gluconapin Progoitrin	*CYP79F1/F2*
LG9	**Total Ali 3** *Raphanin +Prog 3* *Sinig +Napin 3* *Raphanin 1** **Napin 3** *Sinig 2** **Total indo 3** **Brassicin 3**	*CYP79F1/F2* *CYP83B1*	QTL for progoitrin on the map location of *Gls-OH* gene *Gls-ALK* gene	Sum of aliphatic glucosinolates Gluconapin Progoitrin	*MAM 1,2&3* *MYB29&76* Total indolic glucosinolates Glucobrassicin

By comparing position of genes known to code for enzymes catalysing specific steps in the biosynthesis of the glucosinolates that have been mapped in *Arabidopsis,* co-linear with the QTLs observed in this study in *B. oleracea* genome, (Kliebenstein, 2009), it was possible to propose potential genes underlying the QTLs in the C genome (Figure 46).

Methionine

MAM2   *MAM1*

CYP79F1/F2
Total Ali 1, Total Ali 2, Total Ali 3

3-methylthiopropyl   **4-methylthiobutyl**

Gls- oxid
Raphin 1

3-methylsulphinylpropyl   **4-methylsulphinylbutyl**
Glucoraphanin

Gls-Alk

Sinig 1, Sinig 2,Sinig 3,
Sinig 4,Sinig 5,
Napin 1,Napin 2, Napin 3, Napin 4

2-propenyl   **3-butenyl**
Sinigrin   **Gluconapin**

Gls- OH
Prog 1, Prog 2, Prog 3, Prog 4, Prog 5, Prog 6, Prog 7

2-hydroxy-3-butenyl
Progoitrin

Figure 46 The biosynthetic pathway of aliphatic glucosinolates synthesis identified in the AGDH plant lines based on (Magrath et al., 1994), showing the mapped QTLs for total and individual aliphatic glucosinolates synthesis, where the proposed potential genes (*CYP79F1/F2, Gls-oxid, Gls-ALK* and *Gls-OH)* for the core structre formation and secondary modifications on the R side chain, underlying the mapped QTLs on the C genome were hypotheized as described in Table 25 and Table 26.

The QTL (Total Alip 2) identified on LG8 which controls the content of aliphatic glucosinolates showed co-linearity with the position of *CYP79F1/F2* gene (coding for enzymes catalysing the core structure formation of aliphatic glucosinolates) (Kliebenstein, 2009) which has been previously identified at a similar position in

Arabidopsis (Luis Iniguez-Luy et al., 2009). In this study, QTLs controlling the content of aliphatic glucosinolates were mapped on LG7 (Total Alip 1) and on LG9 (Total Alip 3), suggesting the presence of potential *CYP79F1/F2* like gene at the same positions to underlay the observed QTLs (Figure 46).

The QTLs mapped on LG3 (Prog 1, Napin 1 and Sinig+Napin 1) and on LG9 (Raphanin+Prog 3, Sinig+Napin 3, Raphanin 1, Napin 3 and Sinig 2) that control the content of individual aliphatic glucosinolates, were co-linear with regions on the *Arabidopsis* genome where the *MAM* gene family (controlling the chain elongation step of methionine prior to core structure formation of aliphatic glucosinolates group), has been identified (Kliebenstein, 2009). Therefore, the QTLs mapped on LG4 (Prog 2) and on LG7 (Prog 3) control the synthesis of progoitrin, LG5 (Sinig 1) controls the synthesis of sinigrin, and on LG8 (Sinig+Napin 2) controls the synthesis of the sum of sinigrin and gluconapin, suggesting the presence of potential *MAM* like gene family, which have not been yet identified in the *Brassica* or *Arabidopsis* co-linear region (Figure 46).

The positions of the *Gls-ALK* and *Gls-OH* genes (coding for enzymes catalysing double bond formation and hydroxylation reactions on the R group side chain respectively) have been identified in the *Brassica* genome on LG9 (Gao et al., 2007; Kliebenstein, 2009). Comparative analysis of QTLs identified in this study, that control the synthesis of progoitrin mapped on LG3 (Prog 1) and on LG4 (Prog 2) showed co-linearity with regions on the *Arabidopsis* genome where the *Gls-OH* gene has been mapped. Similar QTLs mapped on LG7 (Prog 3) and LG8 (Raphanin+Prog 2) suggested the presence of potential *Gls-OH* like genes, which have not been yet identified in *Brassica* or *Arabidopsis* co-linear regions (Figure 46).

The *Gls-ALK* gene has been identified at the co-linear region for the QTLs mapped on LG9 (Sinig 2 and Napin 3), which control the synthesis of sinigrin and gluconapin. Therefore, similar QTLs mapped on LG3 (Prog 1 and Napin 1), LG4 (Prog 2), LG5 (Sinig 1), LG7 (Prog 3) and LG8 (Napin 2 and Sinig+Prog 2) control the synthesis of glucosinolates with a double bond in the side chain, suggesting the presence of potential *Gls-ALK* like genes underlaying these QTLs (Figure 46).

The QTLs mapped on LG1 (Brassicin 1 and Total indo 1) and on LG3 (Neubr 1) controling the synthesis of glucobrassicin, total indolic glucosinolates and neoglucobrassicin, showed co-linearity with regions on the *Arabidopsis* genome, where the *CYP83B1* gene (controlling the core structure formation of indolic glucosinolates group) has been identified (Kliebenstein et al., 2001a) (Table 25). Therefore, the QTLs mapped on LG2 (Total indo 2) controls the synthesis of total indolic glucosinolates, LG4 (Neubr 2) and on LG5 (Neubr 3) control the synthesis of neoglucobrassicin, LG7 (Brassicin 2) and on LG9 (Tatal indo 3 and Brassicin 3) control the synthesis of total indolic glucosinolates and glucobrassicin, suggesting the presence of potential *CYP83B1* like genes to occur at the same positions (Figure 47). In this study, not any QTL for the synthesis of 4-methoxyglucobrassicin was mapped. QTLs for the synthesis of nuoglucobrassicin mapped on LG3 (Neubr 1), LG4 (Neubr 2) and LG5 (Neubr 3), where the genen(s) controls conversion of glucobrassicin into neuglucobrassicin has not been yet identified.

Tryptophan

CYP83B1
Total indo 1, Total indo 2, Total indo 3

3-indolylmethylglucosinolates
Glucobrassicin

Brassicin 1, Brassicin 2, Brassicin 3

Unknown gene?? Neobra 1, Neobra 2, Neobra 3

CYP81F2

4-hydroxy-3-indolylmethylglucosinolate

1-methoxy-3-indolyl methylglucosinolate
Neoglucobrassicin

4-methoxy-3-indolylmethylglucosinolate
4-Methoxyglucobrassicin
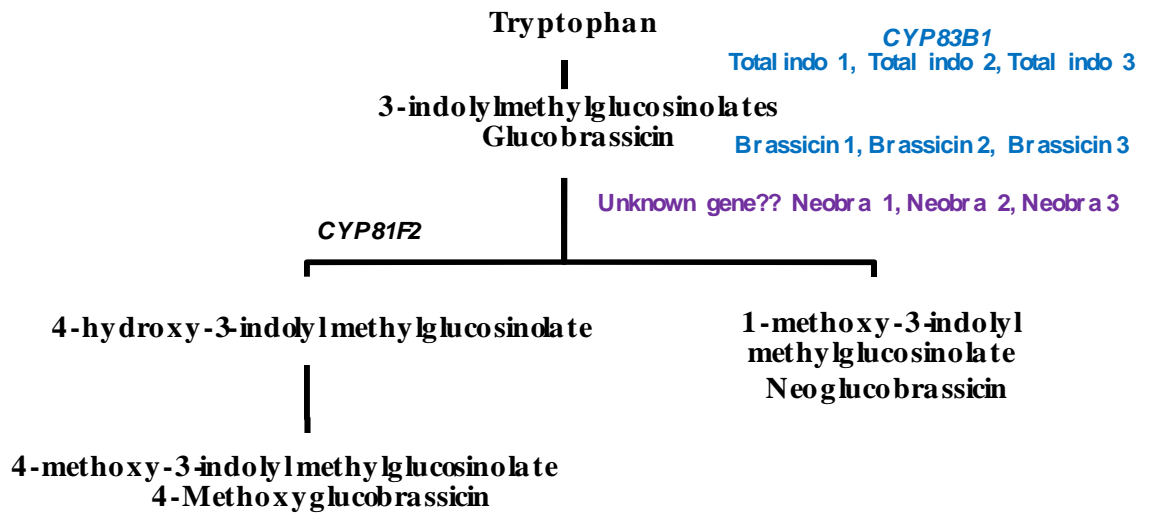
Figure 47 The biosynthetic pathway of indolic glucosinolates synthesis identified in the AGDH plant lines based on (Pfalz et al., 2009), showing the mapped QTLs for total and individual indolic glucosinolates synthesis, where only the proposed potential gene *(CYP83B1)* for the core structre formation of indolic glucosinolates, underlying the mapped QTLs on the C genome was hypothesized as described in Table 25.

4.2.2 QTL mapped on LG9 revealed potential major gene effect controlling aliphatic glucosinolates content

A major QTL for gluconapin at the top region of LG9 (Napin 3) was observed. The frequency distribution within the trait analysed, suggested the presence of a potential gene with a major effect on this trait underling this QTL. Therefore, a possibility of having the same effect on the other glucosinolates linked to the same biosynthetic pathway was proposed.

The initial quantitative data agreed with this hypothesis (Chapter 2). By comparing the concentrations of gluconapin with those of progoitrin, two expression patterns of these traits were observed, in which either progoitrin was expressed at much higher concentrations than that of its precursor gluconapin, or not being expressed at all. It is proposed that this depending on whether the gene *Gls-OH* was functional or not.

Using the data model described under section (3.4.3.2), once the major gene effect was removed, a QTL (Prog 7) for progoitrin was observed within the same interval of the previously identified QTL (Napin 3) for gluconapin (Table 26). These result were expected as it was in agreement with that described by (Mithen et al., 1995), where the *Gls-ALK* gene has been mapped on chromosome At4 in *A. thaliana* as a single dominant Mendelian gene. The previously mapped QTL for progoitrin in *B. oleracea* by (Gao et al., 2007) was at a position co-linear with the *Gls-OH* gene position on chromosome At2 in *A. thaliana* whose phenotype is the presence or absence of progoitrin. These two potential genes were proposed at a region co-linear with the top region of LG9 in the AGDH linkage map used in this study (Kliebenstein, 2009; Luis Iniguez-Luy et al., 2009; Mun et al., 2009; Sebastian et al., 2000) which supported the prediction of a potential major gene underling the observed QTLs.

Table 26 Summary of QTLs mapped on the AGDH LG1-9 using Win QTL Cart. and CIM analysis assorted by linkage groups and were named by the glucosinolate QTL number, for **major gene effect**. Alignment of theses QTLs on previously studied QTL for similar traits and candidate genes or gene regulators underlying these QTLs in *B. rapa* and *A. thaliana* are shown.

Red for potential genes suggested at QTLs positions control the synthesis of similar traits previously identified at collinear regions in *Arabidopsis*. Bold, confirmed QTL by comparative analysis with previously mapped QTL for the same trait in other related species. Italic, QTL where the genes regulating the trait are know.

C	AGDH plant lines QTL	Potenial genes	*B. oleracea*	*B. rapa*	*A. thaliana*
LG2	Sinig 3*	***MAM / ALK* gene**			*MYB28*
LG3	**Prog 4** **Prog-DR 4** Sinig-DR 4	***ALK* gene**		Progoitrin QTL for side chain modification QTL for side chain elongation	*MAM 1,2&3* *MYB29&76* *Gls-OH*
LG4	Prog-DR 5	***MAM/ ALK* gene**		progoitrin	*Gls-OH*
LG5	Sinig 5* Sinig-DR 5	***MAM / ALK* gene**		QTL for side chain modification	
LG7	Prog 6 **Napin 4**	***MAM / Gls-OH* *ALK* gene**		Gluconapin	
LG9	**Prog 7**		QTL for progoitrin on the map location of *Gls-OH* gene *Gls-ALK* gene	Progoitrin	*MAM 1,2&3* *MYB29&76*

The comparative analysis discussed earlier in the previous section (4.2.1), was applied for the investigation of QTLs that control the glucosinolates content, where a potential major gene effect was proposed. Consequently, QTLs mapped on LG2 (Sinig 3), LG4 (Prog 4 and Sinig 4), LG5 (Sinig 5) and LG7 (Prog 6 and Napin 4) coding for the content of glucosinolates with 3 or 4 carbon side chains, proposed potential *MAM* gene like family, which has not been yet identified in the *Brassica* or *Arabidopsis* co-linear region, to underly these QTLs (Figure 46).

194

The QTLs mapped on LG2 (Sinig 3), LG3 (Prog 4 and Sinig 4), LG4 (Prog 5), LG5 (Sinig 5) and LG7 (Prog 6 and Napin 4), which control the synthesis of glucosinolates with alkenyl side chains, proposed potential *Gls-ALK* like genes at similar positions underlying these QTLs, while a QTL controls the synthesis of progoitrin mapped on LG7 (Prog 6), proposed potential *Gls-OH* like genes at a similar position (Figure 46).

4.3 Future work

Glucosinolates are well known to have significant roles in plant defence which indeed affect the agricultural field, crop production, economy and the ecosystem as well (Hopkins et al., 2009). In addition, several studies focused on their toxicological and pharmacological potential, as a fungicidal, bactericidal or nematocidal agents (Gimsing and Kirkegaard, 2009) and their most attractive role as anticancer agents (Bellostas et al., 2007a; Traka and Mithen, 2009).

The identification of genes regulating indolic glucosinolate biosynthesis are of great importance in the medical field as they were known for their anti cancer-activity (Hecht, 2000), and also for agricultural applications increasing aphid resistance and antifungal activity of plants (Bednarek et al., 2009; Pfalz et al., 2009).

Crops finding popular culinary use, especially those for consumption in salads, are the target for genetic engineering and breeding programs to enhance their glucosinolate profile in order to obtain the required variations in their glucosinolate content without adversely affecting their taste. Studying the factors affecting the gene expression or the signals regulating the biosynthesis pathway of different metabolites will significantly contribute to our ability to produce healthy food with

195

high fitness towards environmental stress (Hirai et al., 2007; Keurentjes et al., 2006; Zang et al., 2009).

This project has confirmed some previous findings (QTL for glucosinolates previously mapped in *Brassica* and *Arabidopsis*), and provided novel results that need to be taken forward to yield practical knowledge for more efficient plant breeding.

The consequences for the study of glucosinolates in *B. oleracea,* is the identification of QTL in other *Brassica* species. Therefore, these results need to be improved before being used in the crop species. Thus additional phenotyping experiments will be required, where the absolute concentrations of intact glucosinolates are determined, in order to select only the plant lines with consistent results with the quantification measurements of desulfated glucosinolates obtained in this study. Subsequently, theses line can then be taken forward for more precise genetic analysis performing biological rather than technical replicates.

Analysis of QTL affecting glucosinolate synthesis in crop plants is a promising approach that can be undertaken and used to develop crops for specific purposes, utilizing different methods:

- In this study, the QTLs identified with major gene effect control the biosynthesis of gluconapin and progoitrin on LG9, were proposed to locate at the *Gls-ALK* and *Gls-OH* genes locus. Both glucosinolates are derived from methionine and therefore breeders aiming to produce *Brassica* crops of high composition of the health benefit glucosinolates, e.g. glucoraphanin, as functional foods is can apply the advanced genetic tools in order to down regulate *Gls-ALK* gene (Cartea and Velasco, 2008). Cloning the *Gls-elong* and *Gls-ALK* gene have been already done, and therefore, up regulating the

specific genes that enhance the accumulation of sinigrin for biological control, as anti-fungal and anti-nematode activity for agricultural applications is possible (Li and Quiros, 2003). Such achievements will enable breeders to utilize *Brassica* species for the production of varieties with high levels of specific glucosinolates and the production of crops with lower levels of toxic glucosinolates such as progoitrin (Li and Quiros, 2002).

- The identification of the genes underlying the identified QTL will ultimately provide a greater understanding of the scenario occurring during the evolution process affecting the metabolic pathway of quantitative traits in plants. This is the first step towards fine mapping of QTL and identification of the genes that regulate the trait in the studied population in order to identify candidate genes or defined chromosomal regions controlling the expression of these genes. This may provide further knowledge of the plant genetics and insight into the control of glucosinolate synthesis genes (Kearsey, 1998).

- Markers associated with these QTL can be utilized in marker assisted breeding programs in order to develop crops with the desired profile in the agronomical important crops (*Brassicaceae*). The comparative analysis performed between *Arabidopsis* and *B. oleracea* will enable the transfer of knowledge gained in this study, to crop breeding programmes and the practical development of varieties with defined glucosinolate contents for a varity of end uses.

- Back-cross and selfing approaches can be used to refine the size of the QTL bearing the gene of interest into a small interval for candidate genes identification.

In conclusion, the research presented in this thesis can be used to enhance our ability to breed crops for economic and agricultural interest and for the production of functional foods, which can contribute to a healthier diet. Moreover, the ability of performing precise analysis of glucosinolate content in plant extracts will increase the potential for medical applications, and the design of pharmaceutical formulations for complementary supplements of phytochemicals known for their activity in prevention and treatment of cancer, or for use as natural biofumigants.

REFERANCES

Abercrombie, J. M., Farnham, M. W., Rushing, J. W., 2005. Genetic combining ability of glucoraphanin level and other horticultural traits of broccoli. Euphytica 143, 145-151.

Agerbirk, N., Petersen, B. L., Olsen, C. E., Halkier, B. A., Nielsen, J. K., 2001. 1,4-Dimethoxyglucobrassicin in Barbarea and 4-hydroxyglucobrassicin in Arabidopsis and Brassica. Journal of Agricultural and Food Chemistry 49, 1502-1507.

Andersson, D., Chakrabarty, R., Bejai, S., Zhang, J. M., Rask, L., Meijer, J., 2009. Myrosinases from root and leaves of Arabidopsis thaliana have different catalytic properties. Phytochemistry 70, 1345-1354.

Antonious, G. F., Bomford, M., Vincelli, P., 2009. Screening Brassica species for glucosinolate content. Journal of Environmental Science and Health Part B-Pesticides Food Contaminants and Agricultural Wastes 44, 311-316.

Babula, D., Misztal, L. H., Jakubowicz, M., Kaczmarek, M., Nowak, W., Sadowski, J., 2006. Genes involved in biosynthesis and signalisation of ethylene in Brassica oleracea and Arabidopsis thaliana: identification and genome comparative mapping of specific gene homologues. Theoretical and Applied Genetics 112, 410-420.

Bak, S., Tax, F. E., Feldmann, K. A., Galbraith, D. W., Feyereisen, R., 2001. CYP83B1, a cytochrome P450 at the metabolic branch paint in auxin and indole glucosinolate biosynthesis in Arabidopsis. Plant Cell 13, 101-111.

Barker, G. C., Larson, T. R., Graham, I. A., Lynn, J. R., King, G. J., 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, 1827-1842.

Bednarek, P., Pislewska-Bednarek, M., Svatos, A., Schneider, B., Doubsky, J., Mansurova, M., Humphry, M., Consonni, C., Panstruga, R., Sanchez-Vallet, A., Molina, A., Schulze-Lefert, P., 2009. A Glucosinolate Metabolism Pathway in Living Plant Cells Mediates Broad-Spectrum Antifungal Defense. Science 323, 101-106.

Bellostas, N., Kachlicki, P., Sorensen, J. C., Sorensen, H., 2007a. Glucosinolate profiling of seeds and sprouts of B-oleracea varieties used for food. Scientia Horticulturae 114, 234-242.

Bellostas, N., Sorensen, J. C., Sorensen, H., 2007b. Profiling glucosinolates in vegetative and reproductive tissues of four Brassica species of the U-triangle for their biofumigation potential. Journal of the Science of Food and Agriculture 87, 1586-1594.

Benderoth, M., Pfalz, M., Kroymann, J., 2009. Methylthioalkylmalate synthases: genetics, ecology and evolution. Phytochemistry Reviews 8, 255-268.

Bennett, R. N., Mellon, F. A., Kroon, P. A., 2004. Screening crucifer seeds as sources of specific intact glucosinolates using ion-pair high-performance liquid chromatography negative ion electrospray mass spectrometry. Journal of Agricultural and Food Chemistry 52, 428-438.

Bennett, R. N., Rosa, E. A. S., Mellon, F. A., Kroon, P. A., 2006. Ontogenic profiling of glucosinolates, flavonoids, and other secondary metabolites in Eruca sativa (salad rocket), Diplotaxis erucoides (wall rocket), Diplotaxis tenuifolia (wild rocket), and Bunias orientalis (Turkish rocket). Journal of Agricultural and Food Chemistry 54, 4005-4015.

Bohuon, E. J. R., Keith, D. J., Parkin, I. A. P., Sharpe, A. G., Lydiate, D. J., 1996. Alignment of the conserved C genomes of Brassica oleracea and Brassica napus. Theoretical and Applied Genetics 93, 833-839.

Bones, A. M., Rossiter, J. T., 2006. The enzymic and chemically induced decomposition of glucosinolates. Phytochemistry 67, 1053-1067.

Bor, M., Ozkur, O., Ozdemir, F., Turkan, I., 2009. Identification and Characterization of the Glucosinolate-Myrosinase System in Caper (Capparis ovata Desf.). Plant Molecular Biology Reporter 27, 518-525.

Borowski, J., Szajdek, A., Borowska, E. J., Ciska, E., Zielinski, H., 2008. Content of selected bioactive components and antioxidant properties of broccoli (Brassica oleracea L.). European Food Research and Technology 226, 459-465.

Brown, P. D., Tokuhisa, J. G., Reichelt, M., Gershenzon, J., 2003. Variation of glucosinolate accumulation among different organs and developmental stages of Arabidopsis thaliana. Phytochemistry 62, 471-481.

Cartea, M., Velasco, P., 2008. Glucosinolates in Brassica foods: bioavailability in food and significance for human health. Phytochemistry Reviews 7, 213-229.

Cartea, M. E., Velasco, P., Obregon, S., Padilla, G., de Haro, A., 2008. Seasonal variation in glucosinolate content in Brassica oleracea crops grown in northwestern Spain. Phytochemistry 69, 403-410.

Charron, C. S., Saxton, A. M., Sams, C. E., 2005. Relationship of climate and genotype to seasonal variation in the glucosinolate-myrosinase system. II. Myrosinase activity in ten cultivars of Brassica oleracea grown in fall and spring seasons. Journal of the Science of Food and Agriculture 85, 682-690.

Chen, S. X., Glawischnig, E., Jorgensen, K., Naur, P., Jorgensen, B., Olsen, C. E., Hansen, C. H., Rasmussen, H., Pickett, J. A., Halkier, B. A., 2003. CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in Arabidopsis. Plant Journal 33, 923-937.

Cieslik, E., Leszczynska, T., Filipiak-Florkiewicz, A., Sikora, E., Pisulewski, P. M., 2007. Effects of some technological processes on glucosinolate contents in cruciferous vegetables. Food Chemistry 105, 976-981.

Davidson, N. E., Rutherford, T. J., Botting, N. P., 2001. Synthesis, analysis and rearrangement of novel unnatural glucosinolates. Carbohydrate Research 330, 295-307.

de Quiros, H. C., Magrath, R., McCallum, D., Kroymann, J., Scnabelrauch, D., Mitchell-Olds, T., Mithen, R., 2000. alpha-keto acid elongation and glucosinolate biosynthesis in Arabidopsis thaliana. Theoretical and Applied Genetics 101, 429-437.

Dias, J. S., 1995. Genetic relationships of Portuguese coles and other close related Brassica genotypes using nuclear RFLPs. Genetic Resources and Crop Evolution 42, 363-369.

EEC, 1990. Oil seeds- Determination of glucosinolates. High performance liquid chromatography. Official jornal of European communities VIII, L 170: 127-134.

Fahey, J. W., Zalcmann, A. T., Talalay, P., 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56, 5-51.

Farnham, M. W., Wilson, P. E., Stephenson, K. K., Fahey, J. W., 2004. Genetic and environmental effects on glucosinolate content and chemoprotective potency of broccoli. Plant Breeding 123, 60-65.

Fenwick, G. R., Heaney, R. K., Mullin, W. J., 1983. Glucosinolates and their breakdown products in food and food plants. Crc Critical Reviews in Food Science and Nutrition 18, 123-201.

Fiebig, H. J., 1991. Desulfation of glucosinolates - potential pitfalls of the hplc method for double zero rapeseed. Fett Wissenschaft Technologie-Fat Science Technology 93, 264-267.

Field, B., Cardon, G., Traka, M., Botterman, J., Vancanneyt, G., Mithen, R., 2004. Glucosinolate and amino acid biosynthesis in Arabidopsis. Plant Physiology 135, 828-839.

Font, R., Del Rio-Celestino, M., Rosa, E., Aires, A., De Haro-Bailon, A., 2005. Glucosinolate assessment in Brassica oleracea leaves by near-infrared spectroscopy. Journal of Agricultural Science 143, 65-73.

Francisco, M., Moreno, D. A., Cartea, M. E., Ferreres, F., Garcia-Viguera, C., Velasco, P., 2009. Simultaneous identification of glucosinolates and phenolic compounds in a representative collection of vegetable Brassica rapa. Journal of Chromatography A 1216, 6611-6619.

Gao, M. Q., Li, G. Y., Yang, B., McCombie, W. R., Quiros, C. F., 2004. Comparative analysis of a Brassica BAC clone containing several major aliphatic glucosinolate genes with its corresponding Arabidopsis sequence. Genome 47, 666-679.

Gao, M. Q., Li, G. Y., Yang, B., Qiu, D., Farnham, M., Quiros, C., 2007. High-density Brassica oleracea linkage map: identification of useful new linkages. Theoretical and Applied Genetics 115, 277-287.

Gigolashvili, T., Berger, B., Flugge, U. I., 2009. Specific and coordinated control of indolic and aliphatic glucosinolate biosynthesis by R2R3-MYB transcription factors in Arabidopsis thaliana. Phytochemistry Reviews 8, 3-13.

Gigolashvili, T., Berger, B., Mock, H. P., Muller, C., Weisshaar, B., Fluegge, U. I., 2007a. The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in Arabidopsis thaliana. Plant Journal 50, 886-901.

Gigolashvili, T., Yatusevich, R., Berger, B., Muller, C., Flugge, U. I., 2007b. The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in Arabidopsis thaliana. Plant Journal 51, 247-261.

Gimsing, A. L., Kirkegaard, J. A., 2009. Glucosinolates and biofumigation: fate of glucosinolates and their hydrolysis products in soil. Phytochemistry Reviews 8, 299-310.

Gols, R., Bukovinszky, T., van Dam, N. M., Dicke, M., Bullock, J. M., Harvey, J. A., 2008. Performance of generalist and specialist herbivores and their endoparasitoids differs on cultivated and wild Brassica populations. Journal of Chemical Ecology 34, 132-143.

Graser, G., Oldham, N. J., Brown, P. D., Temp, U., Gershenzon, J., 2001. The biosynthesis of benzoic acid glucosinolate esters in Arabidopsis thaliana. Phytochemistry 57, 23-32.

Griffiths, A., Miller, J., Suzuki, D., Lewontin, R., Gelbart, W., 1996. An introduction to genetic analysis.

Griffiths, D. W., Bain, H., Deighton, N., Botting, N. P., Robertson, A. A. B., 2000. Evaluation of liquid chromatography-atmospheric pressure chemical ionisation-mass spectrometry for the identification and quantification of desulphoglucosinolates. Phytochemical Analysis 11, 216-225.

Grubb, C. D., Abel, S., 2006. Glucosinolate metabolism and its control. Trends in Plant Science 11, 89-100.

Halkier, B. A., Gershenzon, J., 2006. Biology and biochemistry of glucosinolates. Annual Review of Plant Biology 57, 303-333.

Hall, C., McCallum, D., Prescott, A., Mithen, R., 2001. Biochemical genetics of glucosinolate modification in Arabidopsis and Brassica. Theoretical and Applied Genetics 102, 369-374.

Hansen, B. G., Kliebenstein, D. J., Halkier, B. A., 2007. Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in Arabidopsis. Plant Journal 50, 902-910.

Hansen, C. H., Du, L. C., Naur, P., Olsen, C. E., Axelsen, K. B., Hick, A. J., Pickett, J. A., Halkier, B. A., 2001. CYP83B1 is the oxime-metabolizing enzyme in the glucosinolate pathway in Arabidopsis. Journal of Biological Chemistry 276, 24790-24796.

Hecht, S. S., 2000. Inhibition of carcinogenesis by isothiocyanates. Drug Metabolism Reviews 32, 395-411.

Higdon, J. V., Delage, B., Williams, D. E., Dashwood, R. H., 2007. Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis. Pharmacological Research 55, 224-236.

Hirai, M. Y., Sugiyama, K., Sawada, Y., Tohge, T., Obayashi, T., Suzuki, A., Araki, R., Sakurai, N., Suzuki, H., Aoki, K., Goda, H., Nishizawa, O. I., Shibata, D., Saito, K., 2007. Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proceedings of the National Academy of Sciences of the United States of America 104, 6478-6483.

Hopkins, R. J., van Dam, N. M., van Loon, J. J. A., 2009. Role of Glucosinolates in Insect-Plant Relationships and Multitrophic Interactions. Annual Review of Entomology 54, 57-83.

Hounsome, N., Hounsome, B., Tomos, D., Edwards-Jones, G., 2008. Plant metabolites and nutritional quality of vegetables. Journal of Food Science 73, R48-R65.

Hrncirik, K., Velisek, J., Davidek, J., 1998. Comparison of HPLC and GLC methodologies for determination of glucosinolates using reference material. Zeitschrift Fur Lebensmittel-Untersuchung Und-Forschung a-Food Research and Technology 206, 103-107.

Hull, A. K., Vij, R., Celenza, J. L., 2000. Arabidopsis cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis. Proceedings of the National Academy of Sciences of the United States of America 97, 2379-2384.

Jahangir, M., Abdel-Farid, I. B., Kim, H. K., Choi, Y. H., Verpoorte, R., 2009. Healthy and unhealthy plants: The effect of stress on the metabolism of Brassicaceae. Environmental and Experimental Botany 67, 23-33.

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

Jeffery, E. H., Araya, M., 2009. Physiological effects of broccoli consumption. Phytochemistry Reviews 8, 283-298.

Jin, J., Koroleva, O. A., Gibson, T., Swanston, J., Magan, J., Zhang, Y., Rowland, I. R., Wagstaff, C., 2009. Analysis of Phytochemical Composition and Chemoprotective Capacity of Rocket (Eruca sativa and Diplotaxis tenuifolia) Leafy Salad Following Cultivation in Different Environments. Journal of Agricultural and Food Chemistry 57, 5227-5234.

Jones, C. G., Firn, R. D., 1991. On the evolution of plant secondary chemical diversity. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 333, 273-280.

Kearsey, M. J., 1998. The principles of QTL analysis (a minimal mathematics approach). Journal of Experimental Botany 49, 1619-1623.

Keurentjes, J. J. B., Fu, J. Y., de Vos, C. H. R., Lommen, A., Hall, R. D., Bino, R. J., van der Plas, L. H. W., Jansen, R. C., Vreugdenhil, D., Koornneef, M., 2006. The genetics of plant metabolism. Nature Genetics 38, 842-849.

Kiddle, G., Bennett, R. N., Botting, N. P., Davidson, N. E., Robertson, A. A. B., Wallsgrove, R. M., 2001. High-performance liquid chromatographic separation of natural and synthetic desulphoglucosinolates and their chemical validation by UV, NMR and chemical ionisation-MS methods. Phytochemical Analysis 12, 226-242.

Kim, J. K., Chu, S. M., Kim, S. J., Lee, D. J., Lee, S. Y., Lim, S. H., Ha, S. H., Kweon, S. J., Cho, H. S., Variation of glucosinolates in vegetable crops of Brassica rapa L. ssp pekinensis. Food Chemistry 119, 423-428.

Kim, J. K., Chu, S. M., Kim, S. J., Lee, D. J., Lee, S. Y., Lim, S. H., Ha, S. H., Kweon, S. J., Cho, H. S., 2009. Variation of glucosinolates in vegetable crops of Brassica rapa L. ssp pekinensis. Food Chemistry 119, 423-428.

Kim, M. K., Park, J. H. Y., 2009. Cruciferous vegetable intake and the risk of human cancer: epidemiological evidence. Proceedings of the Nutrition Society 68, 103-110.

Kliebenstein, D. J., 2004. Secondary metabolites and plant/environment interactions: a view through Arabidopsis thaliana tinged glasses. Plant Cell and Environment 27, 675-684.

Kliebenstein, D. J., 2009. A quantitative genetics and ecological model system: understanding the aliphatic glucosinolate biosynthetic network via QTLs. Phytochemistry Reviews 8, 243-254.

Kliebenstein, D. J., Gershenzon, J., Mitchell-Olds, T., 2001a. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in Arabidopsis thaliana leaves and seeds. Genetics 159, 359-370.

Kliebenstein, D. J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., Mitchell-Olds, T., 2001b. Genetic control of natural variation in Arabidopsis glucosinolate accumulation. Plant Physiology 126, 811-825.

Kliebenstein, D. J., Kroymann, J., Mitchell-Olds, T., 2005. The glucosinolate-myrosinase system in an ecological and evolutionary context. Current Opinion in Plant Biology 8, 264-271.

Kushad, M. M., Brown, A. F., Kurilich, A. C., Juvik, J. A., Klein, B. P., Wallig, M. A., Jeffery, E. H., 1999. Variation of glucosinolates in vegetable crops of Brassica oleracea. Journal of Agricultural and Food Chemistry 47, 1541-1548.

Kusznierewicz, B., Bartoszek, A., Wolska, L., Drzewiecki, J., Gorinstein, S., Namiesnik, J., 2008. Partial characterization of white cabbages (Brassica oleracea var capitata f alba) from different regions by glucosinolates, bioactive compounds, total antioxidant activities and proteins. Lwt-Food Science and Technology 41, 1-9.

Lai, R. H., Keck, A. S., Walling, M. A., West, L. G., Jeffery, E. H., 2008. Evaluation of the safety and bioactivity of purified and semi-purified glucoraphanin. Food and Chemical Toxicology 46, 195-202.

Lander, E. S., Botstein, D., 1989. Mapping mendelian factors underlying quantitative traits using rflp linkage maps. genetics 121, 185-199.

Lanner, C., Bryngelsson, T., Gustafsson, M., 1997. Relationships of wild Brassica species with chromosome number 2n=18, based on RFLP studies. Genome 40, 302-308.

Leoni, O., Iori, R., Haddoum, T., Marlier, M., Wathetet, J. P., Rollin, P., Palmieri, S., 1998. Approach to the use of immobilized sulfatase for analytical purposes and for the production of desulfo-glucosinolates. Industrial Crops and Products 7, 335-343.

Li, G., Quiros, C. F., 2003. In planta side-chain glucosinolate modification in Arabidopsis by introduction of dioxygenase Brassica homolog BoGSL-ALK. Theoretical and Applied Genetics 106, 1116-1121.

Li, G. Y., Quiros, C. F., 2002. Genetic analysis, expression and molecular characterization of BoGSL-ELONG, a major gene involved in the aliphatic glucosinolate pathway of Brassica species. Genetics 162, 1937-1943.

Lou, P., Zhao, J., He, H., Hanhart, C., Del Carpio, D. P., Verkerk, R., Custers, J., Koornneef, M., Bonnema, G., 2008. Quantitative trait loci for glucosinolate accumulation in Brassica rapa leaves. New Phytologist 179, 1017-1032.

Luis Iniguez-Luy, F., Lukens, L., Farnham, M. W., Amasino, R. M., Osborn, T. C., 2009. Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling Brassica rapa and B. oleracea. Theoretical and Applied Genetics 120, 31-43.

Lukens, L., Zou, F., Lydiate, D., Parkin, I., Osborn, T., 2003. Comparison of a Brassica oleracea genetic map with the genome of Arabidopsis thaliana. Genetics 164, 359-372.

Mackay, D. S. and Flaconer , 1996. Intoduction to quantitative genetics.

Magrath, R., Bano, F., Morgner, M., Parkin, I., Sharpe, A., Lister, C., Dean, C., Turner, J., Lydiate, D., Mithen, R., 1994. Genetics of aliphatic glucosinolates .1. Side-chain elongation in brassica-napus and arabidopsis-thaliana. Heredity 72, 290-299.

Maruyama-Nakashita, A., Nakamura, Y., Saito, K., Takahashi, H., 2007. Identification of a novel cis-acting element in SULTR1;2 promoter conferring sulfur deficiency response in Arabidopsis roots. Plant and Cell Physiology 48, S34-S34.

Matthaus, B., Luftmann, H., 2000. Glucosinolates in members of the family Brassicaceae: Separation and identification by LC/ESI-MS-MS. Journal of Agricultural and Food Chemistry 48, 2234-2239.

Mewis, I., Tokuhisa, J. G., Schultz, J. C., Appel, H. M., Ulrichs, C., Gershenzon, J., 2006. Gene expression and glucosinolate accumulation in Arabidopsis thaliana in response to generalist and specialist herbivores of different feeding guilds and the role of defense signaling pathways. Phytochemistry 67, 2450-2462.

Meyer, M., Adam, S. T., 2008. Comparison of glucosinolate levels in commercial broccoli and red cabbage from conventional and ecological farming. European Food Research and Technology 226, 1429-1437.

Michaud, D. S., Spiegelman, D., Clinton, S. K., Rimm, E. B., Willett, W. C., Giovannucci, E. L., 1999. Fruit and vegetable intake and incidence of bladder cancer in a male prospective cohort. Journal of the National Cancer Institute 91, 605-613.

Mithen, R., 2001. Glucosinolates - biochemistry, genetics and biological activity. Plant Growth Regulation 34, 91-103.

Mithen, R., Clarke, J., Lister, C., Dean, C., 1995. Genetics of aliphatic glucosinolates .3. side-chain structure of aliphatic glucosinolates in arabidopsis-thaliana. Heredity 74, 210-215.

Mun, J. H., Kwon, S. J., Yang, T. J., Seol, Y. J., Jin, M., Kim, J. A., Lim, M. H., Kim, J. S., Baek, S., Choi, B. S., Yu, H. J., Kim, D. S., Kim, N., Lim, K. B., Lee, S. I., Hahn, J. H., Lim, Y. P., Bancroft, I., Park, B. S., 2009. Genome-wide comparative analysis of the Brassica rapa gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. Genome Biology 10.

Newton, E. L., Bullock, J. M., Hodgson, D. J., 2009. Glucosinolate polymorphism in wild cabbage (Brassica oleracea) influences the structure of herbivore communities. Oecologia 160, 63-76.

Olsen, O., Sorensen, H., 1981. Recent advances in the analysis of glucosinolates. Journal of the American Oil Chemists Society 58, 857-865.

Pappa, G., Lichtenberg, M., Iori, R., Barillari, J., Bartsch, H., Gerhauser, C., 2006. Comparison of growth inhibition profiles and mechanisms of apoptosis induction in human colon cancer cell lines by isothiocyanates and indoles from Brassicaceae. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis 599, 76-87.

Pfalz, M., Vogel, H., Kroymann, J., 2009. The Gene Controlling the Indole Glucosinolate Modifier1 Quantitative Trait Locus Alters Indole Glucosinolate Structures and Aphid Resistance in Arabidopsis. Plant Cell 21, 985-999.

Poelman, E. H., van Dam, N. M., van Loon, J. J. A., Vet, L. E. M., Dicke, M., 2009. Chemical diversity in Brassica oleracea affects biodiversity of insect herbivores. Ecology 90, 1863-1877.

Pratt, C., Pope, T. W., Powell, G., Rossiter, J. T., 2008. Accumulation of glucosinolates by the cabbage aphid Brevicoryne brassicae as a defense against two coccinellid species. Journal of Chemical Ecology 34, 323-329.

Prestera, T., Fahey, J. W., Holtzclaw, W. D., Abeygunawardana, C., Kachinski, J. L., Talalay, P., 1996. Comprehensive chromatographic and spectroscopic methods for the separation and identification of intact glucosinolates. Analytical Biochemistry 239, 168-179.

Qiu, D., Gao, M. Q., Li, G. Y., Quiros, C., 2009. Comparative sequence analysis for Brassica oleracea with similar sequences in B-rapa and Arabidopsis thaliana. Plant Cell Reports 28, 649-661.

Quinsac, A., Ribaillier, D., 1987. Optimization of glucosinolate desulfation before hight performance liquid chromatography. World crops: production, utilization, description. Advances in the production and utilization of cruciferous crops, volum 11.

Rae, A. M., Howell, E. C., Kearsey, M. J., 1999. More QTL for flowering time revealed by substitution lines in Brassica oleracea. Heredity 83, 586-596.

Rochfort, S., Caridi, D., Stinton, M., Trenerry, V. C., Jones, R., 2006. The isolation and purification of glucoraphanin from broccoli seeds by solid phase extraction and preparative high performance liquid chromatography. Journal of Chromatography A 1120, 205-210.

Schonhof, I., Krumbein, A., Bruckner, B., 2004. Genotypic effects on glucosinolates and sensory properties of broccoli and cauliflower. Nahrung-Food 48, 25-33.

Schonhof, I., Krumbein, A., Schreiner, M., Gutezeit, B., 1999. Bioactive subestances in cruciferous products. Agriculture and food quality 229, 222-226.

Sebastian, R. L., Howell, E. C., King, G. J., Marshall, D. F., Kearsey, M. J., 2000. An integrated AFLP and RFLP Brassica oleracea linkage map from two morphologically distinct doubled-haploid mapping populations. Theoretical and Applied Genetics 100, 75-81.

Sonderby, I. E., Hansen, B. G., Bjarnholt, N., Ticconi, C., Halkier, B. A., Kliebenstein, D. J., 2007. A Systems Biology Approach Identifies a R2R3 MYB Gene Subfamily with Distinct and Overlapping Functions in Regulation of Aliphatic Glucosinolates. Plos One 2.

Song, L. J., Lori, R., Thornalley, P. J., 2006. Purification of major glucosinolates from Brassicaceae seeds and preparation of isothiocyanate and amine metabolites. Journal of the Science of Food and Agriculture 86, 1271-1280.

Song, L. J., Morrison, J. J., Botting, N. P., Thornalley, P. J., 2005. Analysis of glucosinolates, isothiocyanates, and amine degradation products in vegetable extracts and blood plasma by LC-MS/MS. Analytical Biochemistry 347, 234-243.

Song, L. J., Thornalley, P. J., 2007. Effect of storage, processing and cooking on glucosinolate content of Brassica vegetables. Food and Chemical Toxicology 45, 216-224.

Souli, E., Machluf, M., Morgenstern, A., Sabo, E., Yannai, S., 2008. Indole-3-carbinol (I3C) exhibits inhibitory and preventive effects on prostate tumors in mice. Food and Chemical Toxicology 46, 863-870.

Suzuki, T., Grellet, F., Potter, D., Li, G. Y., Quiros, C. F., 2003. Structure, sequence, and phylogeny of the members of the Ck1 gene family in Brassica oleracea and Arabidopsis thaliana (Brassicaceae). Plant Science 164, 735-742.

Tanksley, S. D., 1993. Mapping polygenes. Annual Review of Genetics 27, 205-233.

Textor, S., de Kraker, J. W., Hause, B., Gershenzon, J., Tokuhisa, J. G., 2007. MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in arabidopsis. Plant Physiology 144, 60-71.

Thompson, C. A., Habermann, T. M., Wang, A. H., Vierkant, R. A., Folsom, A. R., Ross, J. A., Cerhan, J. R., 2010. Antioxidant intake from fruits, vegetables and other sources and risk of non-Hodgkin's lymphoma: the Iowa Women's Health Study. International Journal of Cancer 126, 992-1003.

Traka, M., Mithen, R., 2009. Glucosinolates, isothiocyanates and human health. Phytochemistry Reviews 8, 269-282.

U, Nagaharu, 1935. Genome analysis in Brassica with special referance to the experimental formation of B. napus and peculiar mode of fertilization. Japanese journal of botany 7, 389-452.

Van Ooijen, J. W., 1999. LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity 83, 613-624.

Van Ooijen, J. W., M.P. Boer, R.C. Jansen, C. Maliepaard, 2002. MapQTL 4.0, Software for the calculation of QTL positions on genetic maps. Plant research international, Wageningen, the Netherlands.

Velasco, P., Cartea, M. E., Gonzalez, C., Vilar, M., Ordas, A., 2007. Factors affecting the glucosinolate content of kale (Brassica oleracea acephala group). Journal of Agricultural and Food Chemistry 55, 955-962.

Verhoeven, D. T. H., Verhagen, H., Goldbohm, R. A., vandenBrandt, P. A., vanPoppel, G., 1997. A review of mechanisms underlying anticarcinogenicity by brassica vegetables. Chemico-Biological Interactions 103, 79-129.

Volden, J., Borge, G. I. A., Hansen, M., Wicklund, T., Bengtsson, G. B., 2009. Processing (blanching, boiling, steaming) effects on the content of glucosinolates and antioxidant-related parameters in cauliflower (Brassica oleracea L. ssp botrytis). Lwt-Food Science and Technology 42, 63-73.

Voorrips, R. E., 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93, 77-78.

Wathelet, J.-P., Iori, R., Onofrio, L., Rollin, P., Quinsac, A., Palmieri, S., 2004. Guidelines for glucosinolte analysis in green tissues used for biofumigation. Agroindustria 3, 257-266.

Windsor, A. J., Reichelt, M., Figuth, A., Svatos, A., Kroymann, J., Kliebenstein, D. J., Gershenzon, J., Mitchell-Olds, T., 2005. Geographic and evolutionary diversification of glucosinolates among near relatives of Arabidopsis thaliana (Brassicaceae). Phytochemistry 66, 1321-1333.

Wittstock, U., Halkier, B. A., 2002. Glucosinolate research in the Arabidopsis era. Trends in Plant Science 7, 263-270.

Zang, Y. X., Kim, H. U., Kim, J. A., Lim, M. H., Jin, M., Lee, S. C., Kwon, S. J., Lee, S. I., Hong, J. K., Park, T. H., Mun, J. H., Seol, Y. J., Hong, S. B., Park, B. S., 2009. Genome-wide identification of glucosinolate synthesis genes in Brassica rapa. Febs Journal 276, 3559-3574.
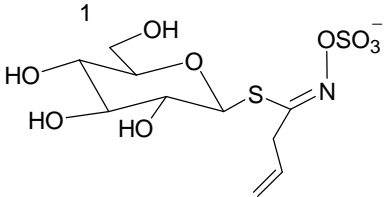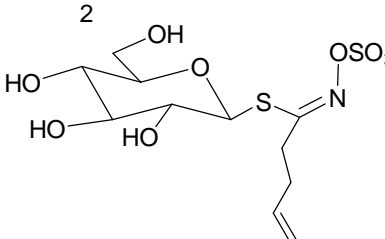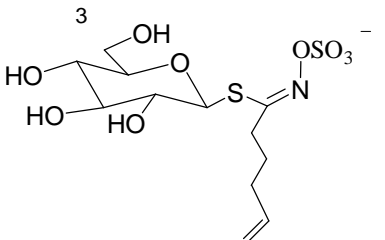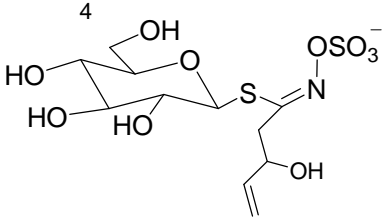
Zeng, Z. B., 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proceedings of the National Academy of Sciences of the United States of America 90, 10972-10976.

Zhu, C. Y., Loft, S., 2001. Effects of Brussels sprouts extracts on hydrogen peroxide-induced DNA strand breaks in human lymphocytes. Food and Chemical Toxicology 39, 1191-1197.

Zimmermann, N. S., Gerendas, J., Krumbein, A., 2007. Identification of desulphoglucosinolates in Brassicaceae by LC/MS/MS: Comparison of ESI and atmospheric pressure chemical ionisation-MS. Molecular Nutrition & Food Research 51, 1537-1546.

Appendix A

The chemical structures, common names, chemical names and masses for IS1 and intact glucosinolates expected in *B. oleracea.* (Bellostas et al., 2007a; Cartea and Velasco, 2008; Fahey et al., 2001; Kiddle et al., 2001; Song and Thornalley, 2007; Zimmermann et al., 2007).

Structure	Glucosinolate	Chemical name	$[M+H]^+$ (m/z)
	[1] Sinigrin	2-Propenyl glucosinolate	359
	[2] Gluconapin	3-Butenyl glucosinolate	373
	[3] Glucobrassicanipin	4-Pentenyl glucosinolate	387
	[4] Progoitrin	2(R)-2-Hydroxy-3-butenyl glucosinolate	389

211

[5] Glucotropaeolin (IS1)	Benzyl glucosinolate	409
[6] Glucoerucin	4-(Methylthio) butyl glucosinolate	421
[7] Gluconasturtiin	2-Phenyethyl glucosinolate	423
[8] Glucoraphanin	4-(Methylsulfinyl) butyl glucosinolate	437
[9] Glucobrassicin	3-Indolylmethyl glucosinolate	448
[10] Glucoalyssin	5-(Methylsulfinyl) Pentyl glucosinolate	451

212
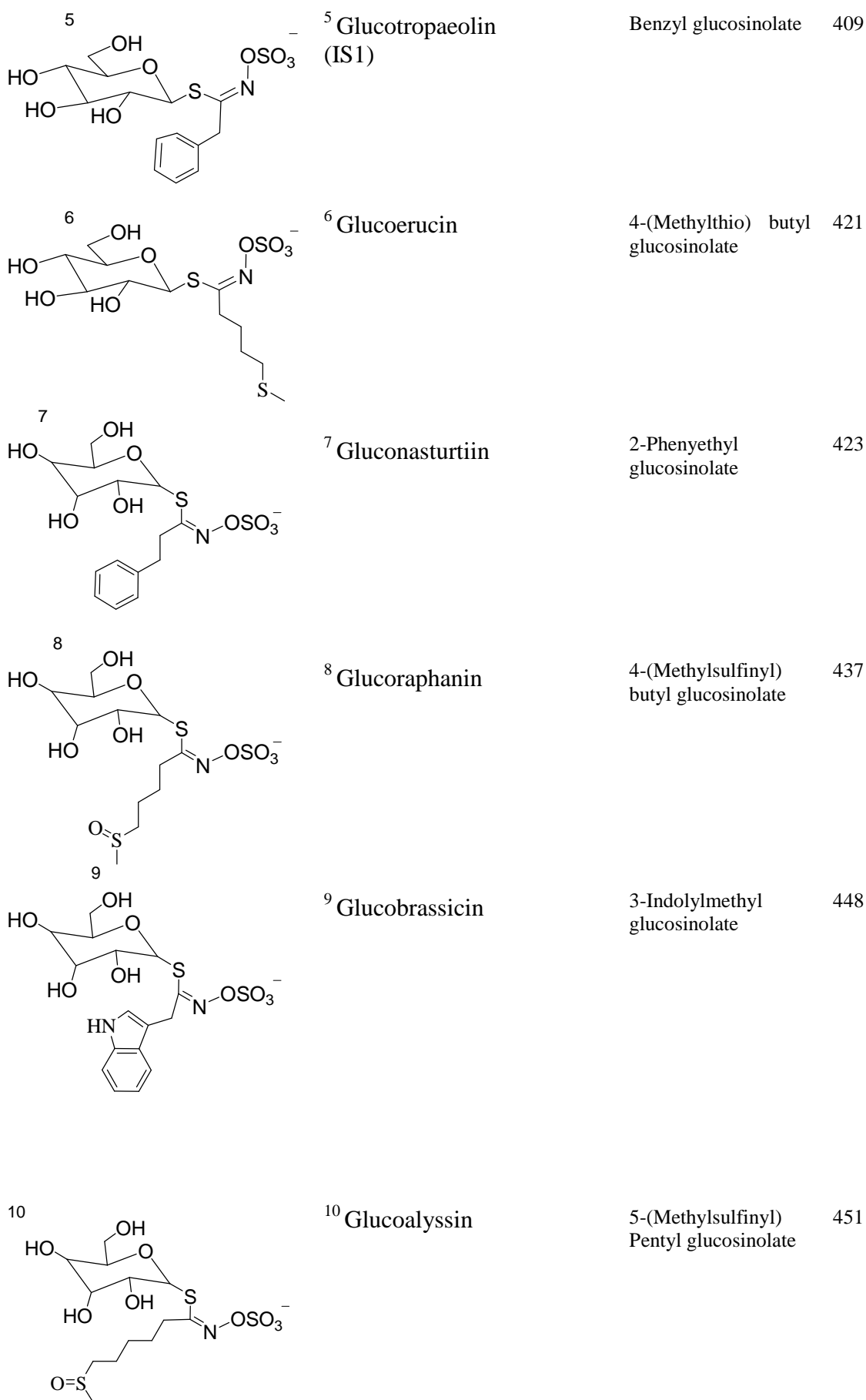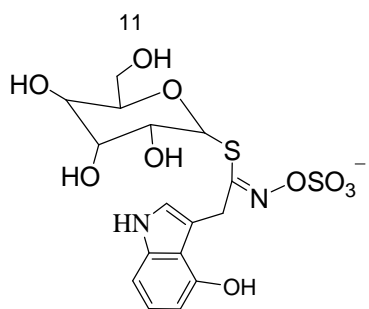
	[11] 4-Hydroxyglucobrassicin	4-Hydroxy-3-indolylmethyl glucosinolate	464
	[12] Glucohesperin	6-(Methylsulfinyl) glucosinolate	465
	[13] 4-Methoxyglucobrassicin	4-Methoxy-3-indolylmethyl glucosinolate	478
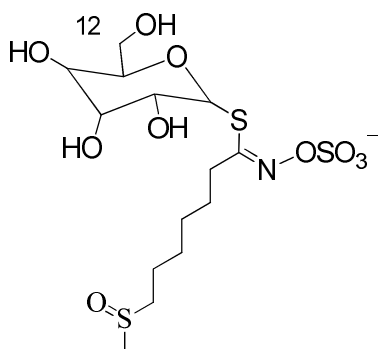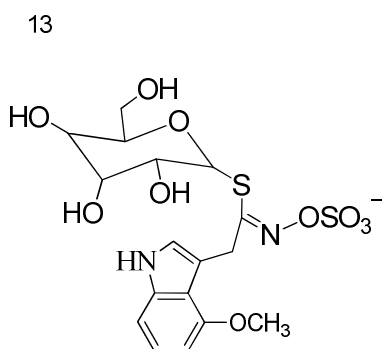
213

14
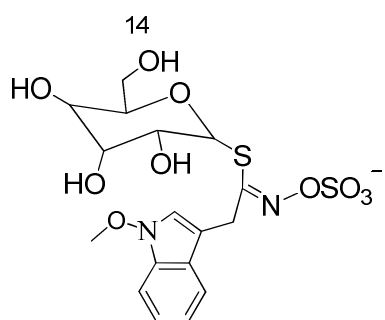
[14] Neoglucobrassicin — 1-Methoxy-3-indolylmethyl glucosinolate — 478

Appendix B

Composition of the mixture used for tuning Bruker Daltonics *ESI-MS (HCT plus)*

(G2431A/G2431-60001)

Neat Material	Gravimetric Conc. mg/kg	Neat Material Purity
Hexamethoxyphosphazene	0.60	99. 9%
Hexakis(2,2-Difluoroethoxy)Phosphazene	5.85	99.0%
Hexakis(1H, 1H, 3H-Tetrafluoropropoxy)Phosphazene	23.09	97 .0%
Hexakis(1H, 1H, 5H-Octafluoropentoxy)Phosphazene	38.15	99.0%
Hexakis(1H, 1H, 7H-Dodecafluoroheptoxy)Phosphazene	53.23	99.0%
Hexakis(1H, 1H, 9H-Perfluorononyloxy)Phosphazene	68.28	99.0%
Tris(Heptafluoropropyl)-S-Triazine	5.50	99. 1%
Betaine	5.89	98.0%
Trifluoroacetic Acid, Ammonium Salt	328.90	99.0%

Appendix C

Composition of the mixture used for tuning Thermo Fisher Scientific *LTQ XL ESI-MS w*ith linear ion trap mass analyser

The *ESI* calibration solution was purchased from *Thermo electron*. It contains 200 μL of 1 mg/mL stock solution of caffeine in 100% methanol, 100 μL of the stock solution of 166.7 pmol/μL MRFA (L-methionyl-arginyl-phenylalanyl-alanine acetate•H2O) in 50:50 methanol: water and 100 μL of 0.1% Ultramark 1621 in acetonitrile. 100 μL of glacial acetic acid and 5mL of acetonitrile were added to the previous mixture. The volume of the solution was adjusted up to 10 mL with 50:50 methanol: water.

Appendix D

Tables of glucosinolate content in 89 AGDH plant lines (µmoles/g dry leaf material)

Frequency distribution histograms of glucosinolates analyzed in 89AGDH plant lines

Line name	D-glucoraphanin					D-progoitrin					D-sinigrin					D-gluconapin				
	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD
GD33	2.44	1.94	1.97	2.11	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.64	1.52	1.57	1.58	0.06	2.06	2.26	1.96	2.09	0.15
AGDH1002	1.11	1.24	1.35	1.23	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1004	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1010	0.00	0.00	0.00	0.00	0.00	3.96	3.84	3.52	3.77	0.23	2.21	2.00	2.15	2.12	0.11	4.41	4.43	4.39	4.41	0.02
AGDH1017	2.31	2.22	2.10	2.21	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1019	0.50	0.47	0.48	0.49	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1020	0.66	0.76	0.64	0.69	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1035	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.05	3.09	2.79	2.98	0.16	4.62	4.59	4.74	4.65	0.08
AGDH1036	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.48	0.48	0.51	0.49	0.02	0.00	0.00	0.00	0.00	0.00
AGDH1042	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.55	0.55	0.55	0.55	0.00
AGDH1047	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2056	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2066	0.00	0.00	0.00	0.00	0.00	22.71	27.55	24.77	25.01	2.43	0.00	0.00	0.00	0.00	0.00	1.07	1.15	1.10	1.10	0.04
AGDH2075	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.58	2.38	2.50	2.49	0.10	2.83	2.76	2.83	2.81	0.04
AGDH2081	1.40	1.39	1.53	1.44	0.08	4.34	4.53	4.32	4.40	0.12	0.00	0.00	0.00	0.00	0.00	0.63	0.68	0.68	0.66	0.03
AGDH2185	0.69	0.78	0.67	0.71	0.06	4.11	4.15	3.94	4.07	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2186	2.86	2.18	2.07	2.37	0.43	11.17	11.77	12.45	11.80	0.64	1.94	2.00	2.01	1.98	0.04	0.89	1.01	0.91	0.94	0.07
AGDH2190	0.38	0.48	0.44	0.44	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2206	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2221	0.00	0.00	0.00	0.00	0.00	1.38	1.21	1.16	1.25	0.12	0.30	0.31	0.34	0.31	0.02	0.00	0.00	0.00	0.00	0.00
AGDH2224	0.00	0.00	0.00	0.00	0.00	4.13	3.97	4.34	4.15	0.18	1.38	1.02	1.68	1.36	0.33	0.28	0.31	0.26	0.28	0.03
AGDH2236	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.03	1.03	1.07	1.05	0.02	2.46	2.38	2.44	2.43	0.04
AGDH2270	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.53	1.00	1.23	1.25	0.26	0.64	0.67	0.73	0.68	0.05
AGDH3015	0.56	0.57	0.61	0.58	0.03	5.38	5.14	5.42	5.31	0.15	0.00	0.00	0.00	0.00	0.00	1.44	1.45	1.40	1.43	0.02
AGDH3016	1.01	0.97	0.95	0.98	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH3066	0.00	0.00	0.00	0.00	0.00	8.85	8.83	8.27	8.65	0.33	0.00	0.00	0.00	0.00	0.00	1.15	1.06	1.17	1.13	0.06
AGDH3070	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.76	0.60	0.66	0.67	0.08	0.00	0.00	0.00	0.00	0.00
AGDH3078	0.44	0.53	0.50	0.49	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.01	3.01	3.09	3.04	0.05
AGDH3081	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.51	4.98	4.98	4.82	0.27	2.93	3.05	3.04	3.01	0.06
AGDH3083	0.00	0.00	0.00	0.00	0.00	7.31	7.46	7.67	7.48	0.18	0.91	0.86	0.77	0.85	0.07	1.71	1.60	1.57	1.63	0.07
AGDH3123	0.00	0.00	0.00	0.00	0.00	5.57	5.51	5.48	5.52	0.05	2.14	2.21	2.16	2.17	0.03	1.05	1.05	1.05	1.05	0.00
AGDH3130	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.63	0.58	0.54	0.11
AGDH3238	0.00	0.00	0.00	0.00	0.00	5.61	6.25	6.19	6.02	0.35	0.00	0.00	0.00	0.00	0.00	1.64	1.57	1.61	1.61	0.04
AGDH4029	6.26	6.06	5.94	6.09	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4031	0.00	0.00	0.00	0.00	0.00	8.18	8.28	9.70	8.72	0.85	0.00	0.00	0.00	0.00	0.00	1.18	1.11	1.19	1.16	0.04
AGDH4034	0.23	0.36	0.26	0.28	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4035	0.00	0.00	0.00	0.00	0.00	79.94	81.00	74.80	78.58	3.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4054	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4061	0.68	0.84	0.76	0.76	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4199	0.00	0.00	0.00	0.00	0.00	5.74	6.12	5.47	5.77	0.33	0.71	0.67	0.36	0.58	0.19	0.12	0.10	0.11	0.11	0.01
AGDH4252	0.50	0.56	0.41	0.49	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5005	0.37	0.40	0.41	0.39	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5007	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.50	1.24	1.91	1.55	0.34	0.00	0.00	0.00	0.00	0.00
AGDH5008	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5010	0.00	0.00	0.00	0.00	0.00	5.76	4.77	5.22	5.25	0.50	2.13	2.25	2.19	2.19	0.06	2.00	1.99	2.02	2.00	0.01
AGDH5012	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5075	0.45	0.52	0.48	0.48	0.03	0.00	0.00	0.00	0.00	0.00	1.21	1.18	1.35	1.24	0.09	4.00	3.92	3.99	3.97	0.05

Line name	D-glucoraphanin					D-progoitrin					D-sinigrin					D-gluconapin				
	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD
AGDH5077	0.00	0.00	0.00	0.00	0.00	5.05	4.94	4.90	4.96	0.08	5.35	5.21	4.89	5.15	0.23	2.31	2.26	2.25	2.27	0.03
AGDH5079	1.04	0.91	0.74	0.90	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5081	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5119	1.58	1.27	1.49	1.45	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH6016	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.80	0.75	0.69	0.75	0.05
AGDH6036	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.87	1.06	0.79	0.91	0.14
AGDH6044	0.00	0.00	0.00	0.00	0.00	3.37	3.59	3.36	3.44	0.13	0.72	0.85	0.64	0.74	0.11	0.26	0.29	0.24	0.26	0.03
AGDH6098	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.40	3.09	3.66	3.39	0.29
AGDH6106	1.17	1.09	1.11	1.12	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH6131	0.00	0.00	0.00	0.00	0.00	8.83	8.67	8.50	8.66	0.16	3.67	3.70	4.02	3.80	0.19	2.88	3.10	2.99	2.99	0.11
AGDH6150	0.69	0.50	0.50	0.56	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1015	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1025	0.00	0.00	0.00	0.00	0.00	21.12	21.70	20.87	21.23	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1027	0.00	0.00	0.00	0.00	0.00	3.29	3.24	3.41	3.32	0.09	0.00	0.00	0.00	0.00	0.00	0.71	0.71	0.71	0.71	0.00
AGDH1038	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1039	0.00	0.00	0.00	0.00	0.00	67.54	64.10	65.01	65.55	1.78	16.95	14.52	14.54	15.34	1.39	7.55	7.20	6.77	7.17	0.39
AGDH1049	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1058	10.79	9.96	9.46	10.07	0.67	6.04	6.75	7.32	6.70	0.64	0.00	0.00	0.00	0.00	0.00	4.35	4.29	4.18	4.27	0.09
AGDH1059	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1060	0.21	0.28	0.25	0.25	0.03	0.00	0.00	0.00	0.00	0.00	1.58	1.89	1.49	1.65	0.21	3.18	3.17	3.11	3.16	0.04
AGDH1064	7.39	8.24	7.53	7.72	0.45	21.36	22.33	22.56	22.08	0.64	0.00	0.00	0.00	0.00	0.00	1.00	0.95	0.82	0.92	0.09
AGDH2069	0.00	0.00	0.00	0.00	0.00	46.77	54.86	54.50	52.04	4.57	0.00	0.00	0.00	0.00	0.00	2.13	2.67	3.16	2.65	0.51
AGDH2078	1.98	1.62	1.66	1.75	0.20	3.83	4.77	4.08	4.23	0.49	0.00	0.00	0.00	0.00	0.00	0.69	0.73	0.72	0.71	0.02
AGDH2134	1.91	2.11	2.01	2.01	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2187	1.06	1.41	1.45	1.30	0.21	7.21	7.49	7.23	7.31	0.15	1.37	1.66	1.86	1.63	0.25	0.54	0.58	0.60	0.57	0.03
AGDH2208	0.53	0.45	0.59	0.52	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2223	0.76	0.74	0.96	0.82	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH3013	0.00	0.00	0.00	0.00	0.00	11.35	11.59	11.26	11.40	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH3088	0.00	0.00	0.00	0.00	0.00	9.62	9.72	8.84	9.39	0.48	0.00	0.00	0.00	0.00	0.00	4.06	4.06	4.00	4.04	0.03
AGDH3235	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4051	2.49	2.56	2.41	2.49	0.07	0.00	0.00	0.00	0.00	0.00	3.07	2.97	2.41	2.82	0.35	0.44	0.39	0.47	0.43	0.04
AGDH4052	0.55	0.50	0.53	0.53	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4056	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.53	1.35	1.58	1.49	0.12	0.57	0.55	0.61	0.58	0.03
AGDH4201	2.17	2.38	2.38	2.31	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5047	0.00	0.00	0.00	0.00	0.00	4.45	4.22	4.05	4.24	0.20	1.64	1.36	1.28	1.43	0.19	0.53	0.52	0.50	0.52	0.02
AGDH5071	0.69	0.78	0.68	0.72	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5076	0.98	0.91	1.10	1.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5080	4.71	3.92	3.72	4.11	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5145	0.00	0.00	0.00	0.00	0.00	5.77	4.76	4.09	4.87	0.85	0.00	0.00	0.00	0.00	0.00	0.74	0.55	0.52	0.60	0.12
AGDH5147	0.00	0.00	0.00	0.00	0.00	4.43	4.71	4.01	4.38	0.35	0.00	0.00	0.00	0.00	0.00	1.35	1.35	1.33	1.35	0.01
AGDH6024	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
AGDH6031	0.00	0.00	0.00	0.00	0.00	7.30	7.09	8.89	7.76	0.98	0.00	0.00	0.00	0.00	0.00	4.13	3.80	3.99	3.97	0.16
AGDH6105	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.92	0.79	0.90	0.10	2.90	2.88	2.90	2.89	0.01

Line name	D-glucobrassicin					D-4-methoxyglucobrassicin					D-neoglucobrassicin				
	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD
GD33	0.95	0.91	0.95	0.94	0.02	1.15	1.16	1.12	1.14	0.02	0.10	0.20	0.10	0.13	0.06
A12	0.45	0.43	0.43	0.44	0.01	0.61	0.65	0.63	0.63	0.02	0.00	0.00	0.00	0.00	0.00
AGDH1002	0.34	0.34	0.36	0.35	0.01	0.36	0.37	0.37	0.37	0.00	0.08	0.06	0.06	0.07	0.01
AGDH1004	0.71	0.79	0.69	0.73	0.05	0.18	0.18	0.16	0.17	0.02	0.20	0.21	0.18	0.20	0.01
AGDH1010	1.28	1.27	1.26	1.27	0.01	0.24	0.25	0.25	0.25	0.01	0.15	0.17	0.15	0.16	0.01
AGDH1017	0.36	0.42	0.41	0.40	0.03	0.56	0.61	0.60	0.59	0.03	0.00	0.00	0.00	0.00	0.00
AGDH1019	0.64	0.61	0.62	0.62	0.01	0.18	0.17	0.17	0.17	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1020	2.02	2.02	1.94	1.99	0.05	0.37	0.39	0.35	0.37	0.02	0.46	0.44	0.44	0.45	0.01
AGDH1035	1.36	1.35	1.40	1.37	0.03	0.34	0.32	0.35	0.34	0.01	0.00	0.00	0.00	0.00	0.00
AGDH1036	0.39	0.39	0.38	0.39	0.01	0.18	0.19	0.18	0.18	0.00	0.16	0.14	0.14	0.15	0.01
AGDH1042	0.43	0.41	0.43	0.42	0.01	0.22	0.23	0.23	0.22	0.01	0.07	0.06	0.03	0.05	0.02
AGDH1047	0.76	0.72	0.75	0.75	0.02	0.95	0.94	0.93	0.94	0.01	0.00	0.00	0.00	0.00	0.00
AGDH2056	1.36	1.39	1.24	1.33	0.08	0.68	0.67	0.54	0.63	0.08	0.61	0.49	0.45	0.52	0.08
AGDH2066	1.09	1.07	1.12	1.09	0.03	0.44	0.43	0.48	0.45	0.02	0.53	0.50	0.50	0.51	0.02
AGDH2075	0.44	0.39	0.46	0.43	0.04	0.61	0.67	0.62	0.63	0.03	0.18	0.14	0.12	0.14	0.03
AGDH2081	0.76	0.76	0.76	0.76	0.00	0.16	0.16	0.17	0.16	0.01	0.00	0.00	0.00	0.00	0.00
AGDH2185	0.83	0.82	0.83	0.83	0.00	0.07	0.06	0.06	0.06	0.01	0.00	0.00	0.00	0.00	0.00
AGDH2186	1.37	1.46	1.43	1.42	0.05	0.56	0.58	0.66	0.60	0.05	0.20	0.22	0.21	0.21	0.01
AGDH2190	0.50	0.50	0.51	0.50	0.01	0.18	0.17	0.18	0.18	0.00	0.18	0.17	0.18	0.18	0.00
AGDH2206	1.28	1.47	1.51	1.42	0.12	2.78	3.08	3.09	2.98	0.18	0.00	0.00	0.00	0.00	0.00
AGDH2221	0.54	0.54	0.54	0.54	0.00	0.22	0.23	0.23	0.23	0.00	0.20	0.22	0.22	0.21	0.01
AGDH2224	0.37	0.36	0.37	0.36	0.00	0.16	0.16	0.15	0.16	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2236	0.57	0.58	0.56	0.57	0.01	0.33	0.33	0.31	0.33	0.01	0.08	0.08	0.09	0.08	0.01
AGDH2270	1.45	1.39	1.40	1.41	0.03	0.28	0.27	0.27	0.27	0.01	0.00	0.00	0.00	0.00	0.00
AGDH3015	0.43	0.43	0.44	0.43	0.00	0.10	0.12	0.11	0.11	0.01	0.18	0.18	0.18	0.18	0.00
AGDH3016	0.52	0.52	0.56	0.54	0.02	0.21	0.22	0.27	0.23	0.03	0.09	0.08	0.10	0.09	0.01
AGDH3066	1.94	1.90	1.90	1.91	0.02	0.30	0.27	0.27	0.28	0.02	0.00	0.00	0.00	0.00	0.00
AGDH3070	0.14	0.16	0.16	0.15	0.01	0.26	0.27	0.24	0.26	0.01	0.00	0.00	0.00	0.00	0.00
AGDH3078	0.40	0.41	0.40	0.40	0.01	0.30	0.28	0.27	0.28	0.01	0.00	0.00	0.00	0.00	0.00
AGDH3081	2.02	1.59	1.72	1.78	0.22	1.75	1.81	1.80	1.79	0.03	0.00	0.00	0.00	0.00	0.00
AGDH3083	0.95	0.89	0.87	0.90	0.04	0.38	0.36	0.34	0.36	0.02	0.29	0.25	0.26	0.27	0.02
AGDH3123	1.80	1.79	1.81	1.80	0.01	0.21	0.21	0.21	0.21	0.00	0.91	0.90	0.91	0.90	0.00
AGDH3130	2.43	2.28	2.84	2.51	0.29	0.70	1.10	1.01	0.94	0.21	0.91	0.87	0.79	0.86	0.06
AGDH3238	0.48	0.51	0.50	0.49	0.02	0.08	0.11	0.08	0.09	0.02	0.00	0.00	0.00	0.00	0.00
AGDH4029	0.47	0.48	0.47	0.47	0.00	0.91	0.85	0.78	0.85	0.07	0.00	0.00	0.00	0.00	0.00
AGDH4031	0.40	0.45	0.43	0.43	0.02	0.20	0.18	0.19	0.19	0.01	0.00	0.00	0.00	0.00	0.00
AGDH4034	0.56	0.56	0.55	0.56	0.00	0.03	0.03	0.03	0.03	0.00	0.12	0.13	0.14	0.13	0.01
AGDH4035	0.00	0.00	0.00	0.00	0.00	2.70	2.14	2.27	2.37	0.30	0.00	0.00	0.00	0.00	0.00
AGDH4054	0.79	0.79	0.78	0.79	0.01	0.15	0.15	0.15	0.15	0.00	0.21	0.18	0.20	0.20	0.01
AGDH4061	0.82	0.84	0.83	0.83	0.01	0.33	0.33	0.33	0.33	0.00	0.14	0.15	0.18	0.16	0.02
AGDH4199	0.28	0.29	0.29	0.29	0.01	0.13	0.13	0.13	0.13	0.00	0.00	0.00	0.00	0.00	0.00
AGDH4252	0.92	0.93	0.92	0.93	0.00	0.12	0.13	0.12	0.13	0.01	0.21	0.20	0.21	0.21	0.01
AGDH5005	0.62	0.63	0.62	0.63	0.01	0.12	0.13	0.12	0.12	0.00	0.15	0.14	0.14	0.14	0.01
AGDH5007	0.63	0.69	0.68	0.67	0.03	0.40	0.44	0.43	0.42	0.02	0.00	0.00	0.00	0.00	0.00
AGDH5008	0.37	0.37	0.38	0.37	0.00	0.08	0.07	0.08	0.08	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5010	1.22	1.22	1.22	1.22	0.00	0.18	0.19	0.19	0.19	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5012	0.16	0.17	0.13	0.15	0.02	0.35	0.31	0.31	0.32	0.02	0.00	0.00	0.00	0.00	0.00
AGDH5075	0.31	0.32	0.31	0.31	0.00	0.11	0.11	0.11	0.11	0.00	0.19	0.17	0.20	0.19	0.01

Line name	D-glucobrassicin					D-4-methoxyglucobrassicin					D-neoglucobrassicin				
	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD	T1	T2	T3	AVR	SD
AGDH5077	0.54	0.55	0.57	0.55	0.02	0.11	0.10	0.11	0.11	0.01	0.00	0.00	0.00	0.00	0.00
AGDH5079	0.49	0.50	0.47	0.49	0.01	0.21	0.21	0.21	0.21	0.00	0.00	0.00	0.00	0.00	0.00
AGDH5081	0.00	0.00	0.00	0.00	0.00	2.76	2.95	3.34	3.01	0.30	0.00	0.00	0.00	0.00	0.00
AGDH5119	1.34	1.40	1.32	1.35	0.04	0.50	0.50	0.50	0.50	0.00	0.00	0.00	0.00	0.00	0.00
AGDH6016	0.29	0.31	0.28	0.30	0.01	0.32	0.36	0.35	0.34	0.02	0.00	0.00	0.00	0.00	0.00
AGDH6036	0.38	0.42	0.56	0.45	0.10	0.38	0.42	0.38	0.40	0.02	0.00	0.00	0.00	0.00	0.00
AGDH6044	1.19	1.18	1.15	1.17	0.02	0.16	0.15	0.16	0.15	0.01	0.00	0.00	0.00	0.00	0.00
AGDH6098	0.63	0.64	0.77	0.68	0.08	0.53	0.59	0.67	0.60	0.07	0.00	0.00	0.00	0.00	0.00
AGDH6106	0.31	0.31	0.31	0.31	0.00	0.17	0.20	0.16	0.18	0.02	0.07	0.15	0.08	0.10	0.05
AGDH6131	0.77	0.80	0.79	0.79	0.02	0.37	0.37	0.37	0.37	0.00	0.12	0.11	0.14	0.12	0.01
AGDH6150	1.25	1.27	1.24	1.25	0.01	0.14	0.12	0.16	0.14	0.02	0.32	0.34	0.36	0.34	0.02
AGDH1015	0.59	0.56	0.62	0.59	0.03	0.26	0.29	0.32	0.29	0.03	0.00	0.00	0.00	0.00	0.00
AGDH1025	0.24	0.37	0.32	0.31	0.07	0.53	0.57	0.48	0.53	0.05	0.00	0.00	0.00	0.00	0.00
AGDH1027	0.38	0.38	0.39	0.39	0.01	0.07	0.07	0.06	0.07	0.01	0.00	0.00	0.00	0.00	0.00
AGDH1038	0.63	0.62	0.58	0.61	0.03	0.66	0.60	0.58	0.61	0.04	0.00	0.00	0.00	0.00	0.00
AGDH1039	0.60	0.58	0.52	0.57	0.04	2.23	2.13	1.96	2.11	0.13	0.00	0.00	0.00	0.00	0.00
AGDH1049	0.33	0.31	0.30	0.31	0.02	0.11	0.12	0.11	0.11	0.01	0.00	0.00	0.00	0.00	0.00
AGDH1058	0.51	0.49	0.48	0.49	0.01	0.17	0.16	0.15	0.16	0.01	0.00	0.00	0.00	0.00	0.00
AGDH1059	1.21	1.17	1.19	1.19	0.02	0.37	0.36	0.36	0.36	0.01	0.08	0.06	0.05	0.06	0.02
AGDH1060	0.60	0.60	0.58	0.59	0.01	0.05	0.04	0.04	0.04	0.00	0.00	0.00	0.00	0.00	0.00
AGDH1064	0.26	0.38	0.50	0.38	0.12	0.54	0.94	0.58	0.69	0.22	0.00	0.00	0.00	0.00	0.00
AGDH2069	0.52	0.55	0.63	0.57	0.05	0.69	0.73	0.80	0.74	0.05	0.00	0.00	0.00	0.00	0.00
AGDH2078	1.03	1.00	1.00	1.01	0.02	0.11	0.11	0.11	0.11	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2134	0.63	0.72	0.74	0.70	0.06	0.52	0.58	0.53	0.54	0.03	0.15	0.12	0.16	0.14	0.02
AGDH2187	0.32	0.30	0.28	0.30	0.02	0.16	0.17	0.17	0.17	0.00	0.00	0.00	0.00	0.00	0.00
AGDH2208	1.04	1.03	1.01	1.03	0.01	0.24	0.25	0.24	0.24	0.01	0.00	0.00	0.00	0.00	0.00
AGDH2223	0.60	0.57	0.57	0.58	0.02	0.23	0.20	0.21	0.21	0.02	0.00	0.00	0.00	0.00	0.00
AGDH3013	0.60	0.58	0.55	0.58	0.03	0.20	0.21	0.24	0.22	0.02	0.00	0.00	0.00	0.00	0.00
AGDH3088	0.33	0.32	0.32	0.33	0.00	0.25	0.24	0.25	0.25	0.01	0.00	0.00	0.00	0.00	0.00
AGDH3235	1.46	1.45	1.44	1.45	0.01	0.20	0.19	0.19	0.20	0.01	0.27	0.27	0.28	0.27	0.01
AGDH4051	0.54	0.60	0.57	0.57	0.03	0.34	0.37	0.42	0.38	0.04	0.18	0.18	0.17	0.18	0.01
AGDH4052	0.67	0.66	0.66	0.66	0.00	0.08	0.07	0.07	0.07	0.01	0.12	0.10	0.11	0.11	0.01
AGDH4056	0.42	0.40	0.39	0.40	0.02	0.21	0.19	0.16	0.19	0.03	0.00	0.00	0.00	0.00	0.00
AGDH4201	0.64	0.68	0.53	0.62	0.08	0.50	0.48	0.43	0.47	0.04	0.00	0.00	0.00	0.00	0.00
AGDH5047	1.42	1.40	1.40	1.41	0.01	0.24	0.24	0.24	0.24	0.00	0.24	0.23	0.25	0.24	0.01
AGDH5071	0.59	0.61	0.63	0.61	0.02	0.22	0.20	0.20	0.21	0.01	0.00	0.00	0.00	0.00	0.00
AGDH5076	0.44	0.44	0.45	0.44	0.01	0.09	0.09	0.09	0.09	0.00	0.16	0.11	0.11	0.13	0.03
AGDH5080	0.78	0.89	0.78	0.82	0.06	0.52	0.48	0.47	0.49	0.03	0.10	0.11	0.12	0.11	0.01
AGDH5145	0.78	0.56	0.50	0.61	0.14	0.23	0.16	0.16	0.18	0.04	0.22	0.18	0.16	0.19	0.03
AGDH5147	0.45	0.45	0.45	0.45	0.00	0.14	0.14	0.17	0.15	0.02	0.20	0.14	0.22	0.19	0.04
AGDH6024	0.77	0.71	0.73	0.74	0.03	0.49	0.44	0.51	0.48	0.04	0.24	0.29	0.16	0.23	0.07
AGDH6031	0.31	0.34	0.30	0.31	0.02	0.36	0.35	0.33	0.35	0.01	0.12	0.12	0.18	0.14	0.03
AGDH6105	0.59	0.59	0.58	0.59	0.01	0.19	0.19	0.19	0.19	0.00	0.16	0.16	0.16	0.16	0.00

Log$_{10}$ of average relative concentration of glucoraphanin

Log$_{10}$ of average relative concentration of progoitrin

Log$_{10}$ of average relative concentration of 4-methoxyglucobrassicin

Log$_{10}$ of average relative concentration of neoglucobrassicin
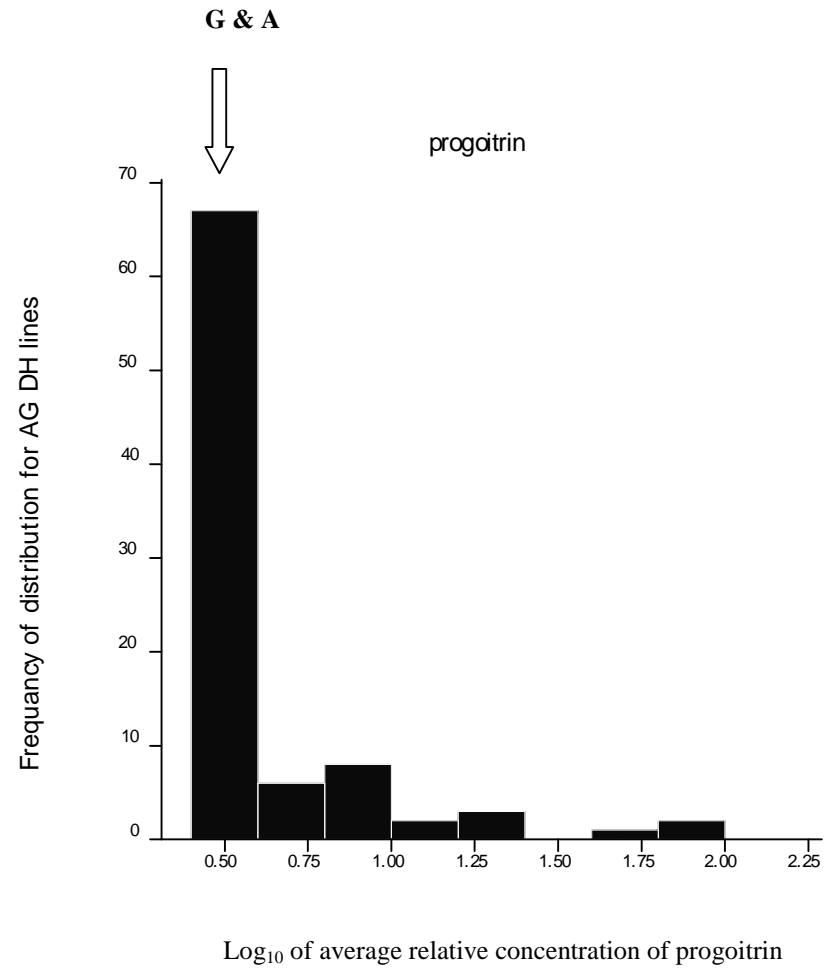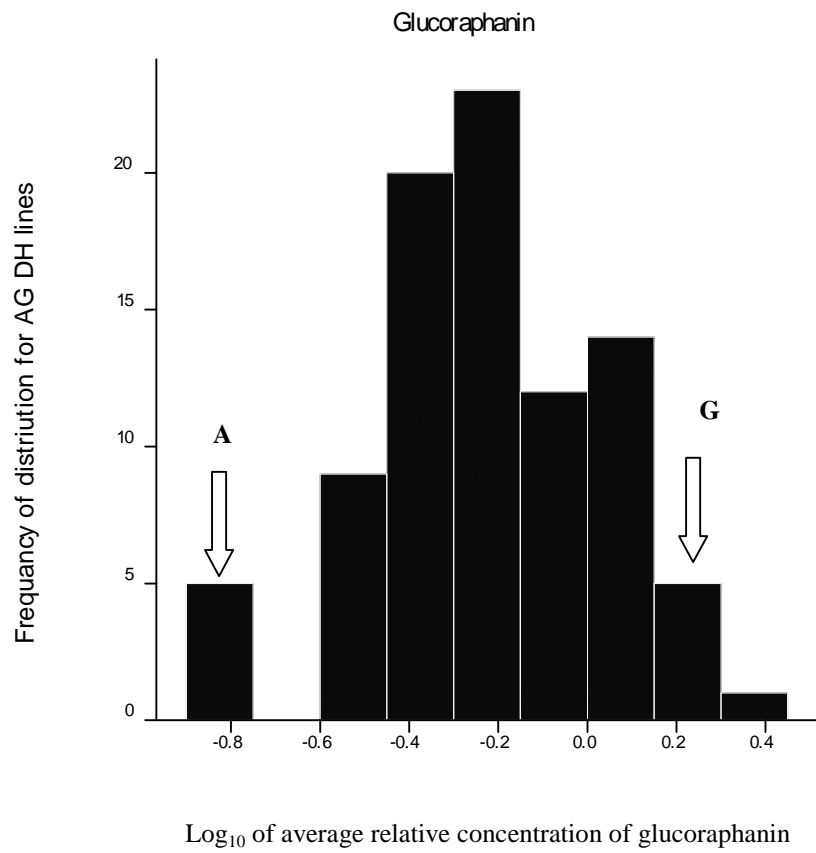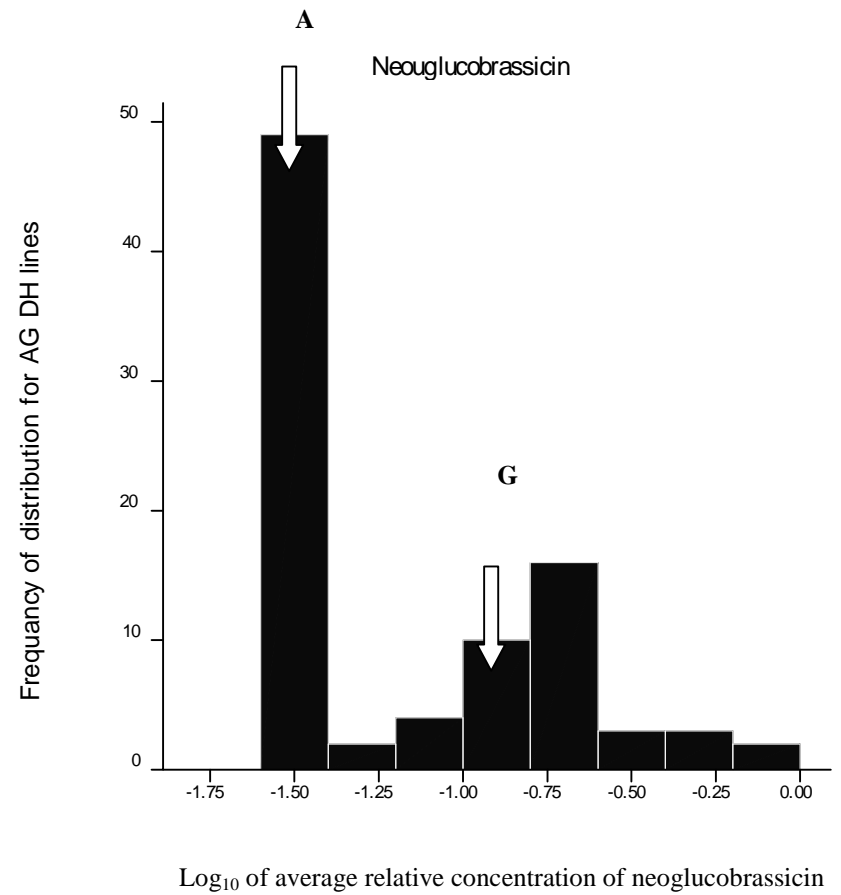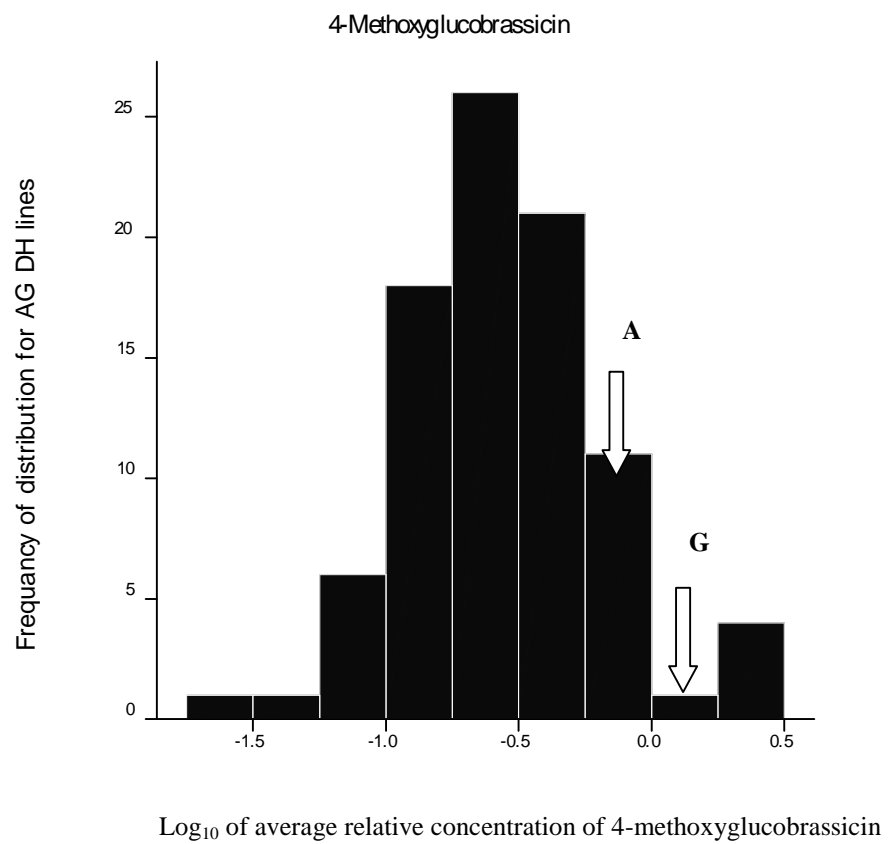
2.

Appendix E

Map QTL tables

2.

Table 1 QTL detected for each glucosinolate, classes of glucosinolates and sub classes of aliphatic glucosinolates, in 89 AGDH segregating mapping population sorted by trait type using the Map QTL program with IM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) and ** (p≤0.01) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance, genome wide significant QTL defined as these with P-value≤ 0.05

Trait	LG	Position (cM)	Marker	Additive effect	Genome wide significant	LOD score	P- value	Tow LOD support interval (cM)	% variation explained by QTL
Progoitrin	3	30.5	PW111J1	-0.1044	2.4	1.84	0.23		9.4%
	4	23.2	P0171J1	-0.0967		1.63	0.32		8.1%
Sinigrin	3	83.6	PN207E1	-0.0976	2.6	1.79	0.22		8.9%
	9	5.0	PN52E2-P0125E1N	0.1121		2.06	0.13		11.4%
Gluconapin	3	54.2	PW143J1	-0.1482	2.7	2.15	0.13		10.6%
	7	50.3	PN97J2	0.1371		1.84	0.28		9.1%
	9**	10.0	PN52E2-P0125E1N	0.2872		7.76	0.01	0.0-22.9	38.2%
Glucobrassicin	1	96.6	PW216J1	-0.0865	2.6	1.97	0.18		10.3%
	1	101.6	PR85E1	-0.0859		1.97	0.18		10.1%
	9	48.8	P0119J1-PW233J1	-0.0956		2.18	0.13		12.9%
Neoglucobrassicin	4	89.6	PW139E1	-0.1504	2.7	2.49	0.08		12.1%
Total aliphatic glucosinolates	1	73.3	PN53E2-PW216J1	-0.1440	2.5	1.55	0.33		19.5%
	9	10.0	PN52E2-P0125E1N	0.1022		1.81	0.23		10.1%
Total indolic glucosinolates	2	0	PW116E1	-0.0725	2.7	1.83	0.28		9.4%
	9	37.1	PW114E2	-0.0708		1.77	0.33		9.0%
Sum of glucoraphanin and progoitrin	4	23.2	P0171J1	-0.0414	2.7	1.58	0.34		7.8%
	7	50.3	PN97J2	0.0426		1.72	0.33		8.5%
	8	58.1	P0143E2	-0.0507		1.7	0.33		8.6%
	9	10.0	PN52E2-P0125E1N	0.0578		2.57	0.06		14.7%
Sum of sinigrin and gluconapin	8	58.1	P0143E2	-0.1194	2.6	1.55	0.40		7.9%
	9**	5	PN52E2-P0125E1N	0.2125		6.53	0.01	0.0-23.4	33.3%

225

2.

Table 2 QTL detected for glucosinolates expected to be under the control of major gene effect, in 89 AGDH segregating mapping population sorted by trait type using the Map QTL program with IM analysis. The QTLs are shown related to the molar concentration/ g dry plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) and ** (p≤0.01) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance, genome wide significant QTL defined as these with P-value≤ 0.05

Trait	LG	Position (cM)	Marker	Additive effect	Genome wide significant	LOD score	P-value	Two LOD support interval (cM)	% variation explained by QTL
Gluconapin	7*	50.3	PN97J2	0.2056	2.8	2.92	0.03	32.3-72.0	26.8%
Sinigrin	3	83.6	PN207E1	-0.1620	2.7	2.20	0.13		18.4%
	5*	31.0	P0105J1-PW164E1	0.1993		3.15	0.02	8.6-43.6	27.8%
Progoitrin-DR	3	30.5	PW111J1	-0.1545	2.7	1.72	0.29		16.9%
Sinigrin-DR	3	83.6	PN207E1	-0.1552	2.6	1.63	0.41		16.1%
	5**	33.6	PW164E1	0.2392		4.42	0.01	23.0-43.0	37.7%

226

2.

Table 3 QTL detected for each glucosinolate, classes of glucosinolates and sub classes of aliphatic glucosinolates, in 89 AGDH segregating mapping population sorted by trait type using the Map QTL program with MQM analysis. The QTLs are shown related to the molar concentration/ g dried plant material. Map positions expressed relative to an integrated map in bold for significant QTLs defined as these with LOD scores above the threshold level were significant at * (p≤0.05) and ** (p≤0.01) determined by 1000 permutation test for each trait analysed. The maximum LOD point and the two LOD support interval are shown for each QTL in centi-Morgans (cM) with the nearest markers allocated at these points where applicable. Additive effects indicated for each trait, with positive effect associated with A12DH and negative effect associated with GD33DH parents. % variation of trait explained by QTL equal to the additive effect squared as a proportion of the line variance, genome wide significant QTL defined as these with P-value≤ 0.05, markers were chosen as cofactors where a significant QTL was mapped in the IM analysis

Trait	LG	Position (cM)	Marker	Additive effect	Genome wide significant	LOD score	P-value	Two LOD support interval (cM)	% variation explained by QTL	Cofactors
Sinigrin	5*	33.6	PW164E1	0.1120	2.4	2.55	0.03	5.0-48.6	11.0%	C9: PN52E2-P0125E1N
Glucobrassicin	9*	48.8	P0199J1-PW233J1	-0.0965	2.5	2.5	0.05	10.0-74.9	13.1%	C1: PW216J1-PR85E1
Sum of glucoraphanin and progoitrin	9*	10.0	PN52E2-P0125E1N	0.0576	2.6	2.92	0.02	0.0-42.1	14.6%	C8: P0143E2
Sum of sinigrin and gluconapin	3*	54.2	PW143J1	-0.1212	2.6	2.53	0.05	35.5-98.7	11.4%	C8: P0143E2
	9**	5.0	PN52E2-P0125E1N	0.2076		7.15	0.01	0.0-22.0	31.1%	