PLANT SCIENCE MEETS FOOD SCIENCE:

genetic effects of glucosinolate degradation during food processing in *Brassica*

Kristin Hennig

PROPOSITIONS

1. Breeding for retention of phytochemicals during processing is a novel approach to improve food quality (this thesis).

2. Novel metabolomics techniques are powerful tools to propose new hypotheses, but conventional analytical methods are still required to confirm the findings (this thesis).

3. Various compounds contribute to the beneficial effects of fruits and vegetables on human health (Martin et al. 2013); therefore increased concentrations of specific phytochemicals cannot compensate for lower consumption.
   Martin, C; Zhang, Y; Tonelli, C; Petroni, K (2013). Plants, diet, and health. *Annual Review of Plant Biology* **64**: 19-46.

4. Chronic sleep restriction in our 24-h society is associated with increased obesity and diabetes occurrence (Knutson, 2007); hence strategies to reduce the incidence of obesity and related diseases should also tackle the problem of sleep deprivation.
   Knutson, K L; Spiegel, K; Penev, P; Van Cauter, E (2007). The metabolic consequences of sleep deprivation. *Sleep Medicine Reviews* **11** (3): 163-178

5. A shorter food supply chain by using local and seasonal food products will reduce food fraud.

6. We should be happy about the things we have rather than complaining about the things we don't have.

Propositions belonging to the thesis, entitled:
"Plant science meets food science: genetic effects of glucosinolate degradation during food processing in *Brassica*"
Kristin Hennig
Wageningen, 30 August 2013

PLANT SCIENCE MEETS FOOD SCIENCE:

genetic effects of glucosinolate degradation during food processing in *Brassica*

Kristin Hennig

Thesis committee

Promotor

Prof. dr. ir. Martinus A. J. S. van Boekel
Professor, Product Design and Quality Management
Wageningen University, Wageningen, NL

Co-promotors

Dr. ir. Guusje A. B. Bonnema
Assistant Professor, Laboratory of Plant Breeding
Wageningen University, Wageningen, NL

Dr. ir. Matthijs Dekker
Associate Professor, Food Quality and Design
Wageningen University, Wageningen, NL

Dr. ing. Ruud Verkerk
Assistant Professor, Food Quality and Design
Wageningen University, Wageningen, NL

Other members

Prof. dr. Harry J. Wichers, Wageningen University, NL

Prof. dr. Richard Mithen, Institute of Food Research, UK

Prof. dr. ir. Maarten Koornneef, Wageningen University, NL

Prof. dr. Raoul J. Bino, Wageningen University, NL

This research was conducted under the auspices of the Graduate schools of Experimental Plant Sciences.

PLANT SCIENCE MEETS FOOD SCIENCE:
genetic effects of glucosinolate degradation during food processing in *Brassica*

Kristin Hennig

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
on Friday, 30 August 2013
at 11 a.m. in the Aula.

Table of Contents

LIST OF ABBREVIATIONS

AA	ascorbic acid
Acetyl-CoA	acetyl coenzyme A
ALFP	amplified fragment length polymorphism
ANOVA	analysis of variance
AOP	genes encoding α-ketoglutarate-dependent dioxygenases
CYP	cytochromes P450
Conc.	concentration
DH	doubled haploid
DW	dry weight
ESP	epithiospecifier protein
FW	fresh weight
GB	glucobrassicin
GLs	glucosinolates
HPLC	high pressure liquid chromatography
IB	glucoiberin
ITC	isothiocyanate
LC	liquid chromatography
LOD	logarithm of odds
MAM	methylthioalkylmalate
MAS	marker-assisted selection
MS	mass spectrometry
NAP	gluconapin
PRO	progoitrin
QTL	quantitative trait loci
Q-TOF-MS	quadrupole time-of-flight mass spectrometry
RAPH	glucoraphanin
RFLP	restricted fragment length polymorphism
SIN	sinigrin
UV	ultra violet
WHO	World Health Organisation
4ME-GB	4-methoxyglucobrassicin
4OH-GB	4-hydroxyglucobrassicin

CHAPTER

General Introduction

1.1 Introduction

The World Health Organization (WHO) identified nutrition as the major modifiable factor for the occurrence of chronic diseases that increased dramatically over the last decades [1]. Several studies report epidemiological evidence for a reduced risk of disease through a higher consumption of fruits and vegetables [2-4]. Fruits and vegetables contain high amounts of dietary fibre and potassium, low amounts of sodium, a broad spectrum of phytochemicals and are of low energy density [4]. Phytochemicals are non-essential nutrients occurring in plant-based foods, which have been linked to reduced incidence and progression of a range of diseases. Phytochemicals supplied in purified form as dietary supplements either lose their bioactivity or behave differently compared to whole foods [2, 3]. Carotenoids, flavonols and more complex polyphenols, saponins, phytosterols and glucosinolates belong to the most important groups of phytochemicals [5].

1.2 The glucosinolate – myrosinase system

Glucosinolates (GLs) are a group of phytochemicals that exclusively occur in *Cruciferous* plants. Important vegetable crops belong to the subfamily *Brassicaceae*, like turnips, cabbage, broccoli, cauliflower, kale, Chinese cabbage, Chinese kale, but also horseradish and mustard [6]. Chemically, GLs are *β*-thioglucoside-*N*-hydroxysulphates with a sulphur-linked *β*-D-glucopyranose moiety. Up to 120 different GLs have been identified with a common core structure, but differing in their side chain (aliphatic, aromatic, and indolic). However, only a restricted number (3-10 GLs) occurs in commonly consumed vegetables. Chemical names, structures and occurrence of some GLs are given in Table 1.1.

GLs coexist with an endogenous enzyme myrosinase (*β*-thioglucosidase, E.C. 3.2.1.147) in the plant tissue, but are located in different cells and are therefore physically separated. Upon tissue damage, GLs are hydrolysed by myrosinase giving rise to a range of breakdown products (isothiocyanates, nitriles, epithionitriles, thiocyanates, oxazolidine-2-thiones) (Figure 1.1). The structure of the GL side chain, the reaction conditions (e.g. pH), presence of additional cofactors (e.g. Fe^{2+}) and proteins (e.g. epithiospecifier protein (ESP) and thiocyanate-forming protein) determine the type of hydrolysis product formed [7-10].

Table 1.1: Chemical names, structures, occurrence and concentrations of glucosinolates commonly found in *Brassica* vegetables and some *Brassicaceae* seeds (modified from [6, 13, 14]).

General glucosinolate structure					

Trivial name	Chemical name side chain	Structure side-chain (R=)	Molecular weight [g/mol][1]	Main source	Conc. range[2] [μmol/100g FW]
Aliphatic					
Glucoibervirin	3-Methylthiopropyl-		407.48	green cauliflower white cauliflower	0 – 11.8 1.5 – 7.1
Glucoerucin	4-Methylthiobutyl-		421.51	rocket	52 – 109
Glucoiberin	3-Methylsulfinylpropyl-		423.48	broccoli sprouts savoy cabbage	59 – 181 24 – 50
Glucoraphanin	4-Methylsulfinylbutyl-		437.51	broccoli sprouts broccoli	233 – 676 24 – 285
Sinigrin	Prop-2-enyl-		359.37	Brussels sprouts white cauliflower	46 – 91 57 – 121
Gluconapin	But-3-enyl-		373.40	pak choi	24 – 157
Gluco-brassicanapin	Pent-4-enyl-		387.42	Chinese cabbage pak choi	2.3 – 25 27 - 69
Progoitrin	(2R)-2-Hydroxybut-3-enyl-		390.41	turnip Chinese broccoli	18 – 41 49
Indolic					
Glucobrassicin	Indol-3-ylmethyl-		448.47	many vegetables, e.g.: broccoli cauliflower	13 – 29 11 – 33
4-Hydoxy-glucobrassicin	4-Hydroxy-indol-3-ylmethyl-		465.47	many vegetables, e.g.: broccoli cauliflower	0.1 – 3.3 0.2 – 2.8
4-Methoxy-glucobrassicin	4-Methoxy-indol-3-ylmethyl-		479.50	many vegetables, e.g.: broccoli cauliflower	0.9 – 2.8 0.7 – 3.2
Neo-glucobrassicin	N-methoxyindol-3-ylmethyl-		479.50	many vegetables, e.g.: broccoli cauliflower	1.8 – 13 0.9 – 3.0
Aromatic					
Glucotropaeolin	Benzyl-		396.41	garden cress	
Gluconasturtiin	Phenylethyl-		410.44	water cress	

[1] Calculated with protonated sulphate group

[2] Conversions from dry weight to fresh weight calculated assuming 10% dry matter

GL hydrolysis products are involved in the defence system of the plant, they have toxic and deterring properties to herbivores (reviewed by Textor and Gershenzon [11]). GLs determine the typical flavour of *Brassica* vegetables; sinigrin and gluconapin are described as bitter, whereas glucoiberin, glucoraphanin, glucoibervirin and glucoerucin do not contribute to the bitterness [12]. Furthermore, GL hydrolysis products show antifungal, antibacterial and insecticidal properties and positive effects on human health [5].
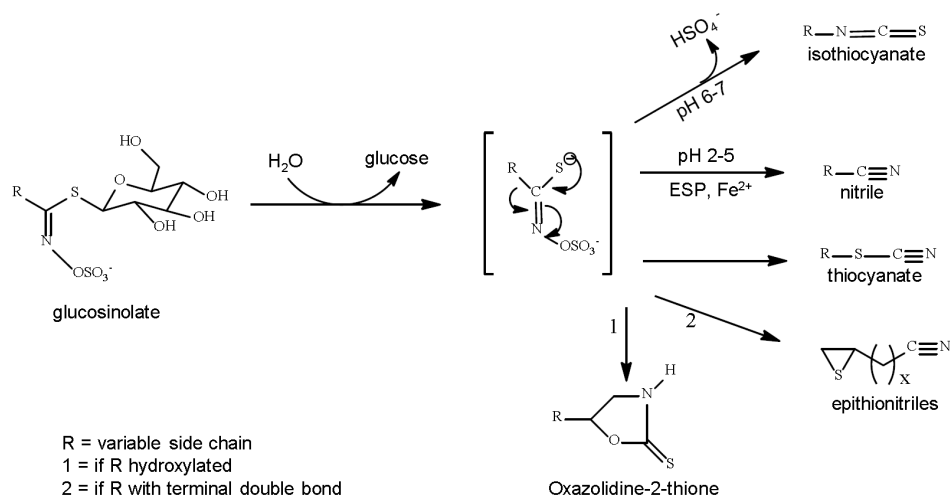

Figure 1.1: Enzymatic breakdown of glucosinolates. The type of breakdown product formed depends on the reaction environment, for example enzymatic cofactors (epithiospecifier protein (ESP), ferrous ions (Fe^{2+}) (adjusted from Fahey *et al.* [8] and Latte, Apple & Lampen [10]).

1.3 Health effects of glucosinolates

Epidemiological studies suggest an inverse association between consumption of cruciferous vegetables and the risk of lung, stomach, colorectal, breast, bladder and prostate cancer. Studies on cell and animal models support the epidemiological evidence. GLs are not bioactive themselves, their isothiocyanate (ITC) breakdown products are bioactive and act on the process of carcinogenesis [15]. In contrast, nitriles don't show bioactivity at dietary levels [16]. Sulforaphane, the ITC derived from the aliphatic glucoraphanin, which is the most abundant GL in broccoli, is the most widely studied and appears to be the most bioactive substance. Limited evidence for protective effects exists also for iberin-ITC (derived from glucoiberin), phenylethyl-ITC (derived from gluconasturtiin), benzyl-ITC (from glucotropaeolin) and for allyl-ITC (from sinigrin)

[10, 15]. Also indoles (indole-3-carbinol derived from glucobrassicin) as degradation products have shown anticarcinogenic properties [10, 17].

The described mechanisms affecting key regulators of carcinogenesis for ITCs are:
a) protection against external factors by affecting xenobiotic metabolism, b) targeting of tumour cell growth and c) increasing antioxidant activities of cells. Other biological activities of ITCs that are indirectly linked to cancer are anti-inflammatory effects and anti-bacterial properties against *Helicobacter pylori* which cause gastritis and enhance the risk of gastric cancer [15, 18]. ITCs alter the xenobiotic metabolism by inhibition of phase 1 enzymes and by induction of phase 2 enzymes. Phase 1 enzymes catalyse hydrolytic, oxidation or reduction reactions to form more polar compounds that can be further metabolised by phase 2 enzymes. The products of phase 1 enzyme reactions are possibly highly reactive intermediates that can be harmful by binding to DNA, RNA and proteins. Phase 2 enzymes convert the intermediates into inactive products which are then excreted from the body [5, 15]. ITCs influence tumour cell growth by induction of apoptosis (programmed cell death), inhibition of cell cycle progression, and inhibition of angiogenesis (playing an important role in maintaining tumour growth and size). ITCs affect the antioxidant levels in the cells by an increase of glutathione levels, which maintains cellular oxidation-reduction balance [15, 16]. Indole-3-carbinol and its major metabolite diindoylmethane has been shown to induce alteration of the oestrogen metabolism, which may cause the prevention or breast cancer in mice [16]. Furthermore, a few studies suggest that indole-3-carbinol and sulforaphane may protect against inflammatory diseases [15, 16, 18, 19].

Besides the health promoting properties, some negative effects of GL breakdown products have been described. For some GL-derived degradation products goitrogenic effects (swelling of the thyroid gland) have been observed in animal studies. Mutagenic potential of broccoli extract, especially in fresh juice, was found in a limited number of *in vitro* and experimental animal studies. Mutagenic products were detected after incubation of the indolic-GL neoglucobrassicin with myrosinase. Animal experiments showed that some nitriles might cause damages in specific organs when applied in unphysiological high concentrations. However, these toxic properties have not been reported in humans [10].

Brassica vegetables are commonly consumed after cooking, where myrosinase has been inactivated. GLs can be converted into beneficial ITCs by myrosinases of the intestinal

1

flora, although the conversion rate is about 3-10 times lower compared to ingestion of uncooked *Brassica* vegetables [19-21]. Furthermore, the conversion rate varies enormously (1%-40%) between individuals [22]. Due to the conversion of GLs into ITCs by the intestinal flora, high levels of intact GLs, in the absence of active myrosinase at the stage of consumption, are considered to exhibit health promoting effects as well. Since the main intake of GLs in the diet is through products that do not contain active myrosinase, it is likely that the observed health effects in epidemiological studies are mainly the result of the conversion by the intestinal flora.

1.4 Effects of food processing

Brassica vegetables are commonly consumed after cooking, which leads to changes in the GL myrosinase system. Mechanisms affecting the GL-myrosinase system that could take place in the vegetables during cooking have been identified as a) enzymatic degradation of GLs, b) (partial) inactivation of myrosinase, c) leaching of GLs, breakdown products and myrosinase into the cooking water and d) thermal degradation of GLs and breakdown products as illustrated in Figure 1.2. GL losses during food processing strongly depend on the cooking method and time-temperature profile applied [23]. During cooking, the various mechanisms could take place either sequentially or simultaneously that involve (bio)chemical reactions, heat and mass transfer [24].
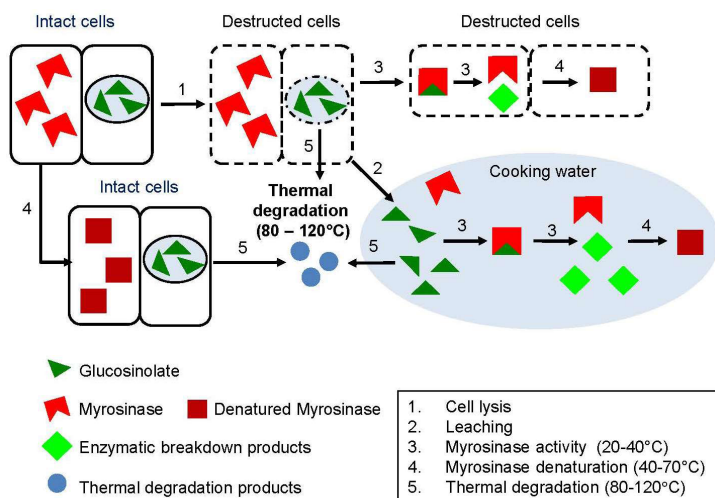

Figure 1.2: Possible changes in the glucosinolate-myrosinase system during food processing. Depending on the processing conditions, the mechanisms could take place sequentially or simultaneously and to different extent (adapted from Dekker *et al.* [23]).

1.4.1 Inactivation of myrosinase

Enzymatic degradation of GLs during cooking appears to be a minor mechanism during food processing, since myrosinase is inactivated at temperatures below common processing temperatures (100°C). Purified myrosinase from mustard seeds (*sinapis alba L.*) was thermally stable during 10 min heat treatment at 60°C, but was quickly inactivated at higher temperatures [25]. Slightly lower thermostability after heat treatment for 10 min was reported in broccoli juice where myrosinase was thermally stable up to 40°C, at 60° the myrosinase activity was reduced by 90% and at 70°C no myrosinase activity was detected anymore [26]. Partially purified myrosinase from broccoli did not show inactivation up to 30°C, however at 40°C significant inactivation occurred [27]. Similar results were reported from partially purified myrosinase from green cabbage, which was stable for 10 min up to 35°C and 90% of the activity was lost after 10 min at 60°C [28]. These data show slight differences in stability of myrosinase among different vegetables; however, inactivation occurs rapidly at temperatures below the cooking temperature of 100°C. Steaming of vegetables for a short time can result in residual myrosinase activity, for example after steaming of red cabbage for 7 min, 10% of myrosinase activity remained [29] and steaming of broccoli for less than 6 min preserved myrosinase activity partially [30].

1.4.2 Leaching

Leaching of GL is the major cause of GL losses during boiling and blanching (with submersion into water) and it is strongly dependent on the ratio of vegetable to water, the heating time and method, and the type and geometrical shape of the vegetable tissues. GL losses of 25-75% are typically expected during boiling. During steaming, diffusion and leaching are lower compared to boiling, because the vegetable is not in direct contact with a large pool of water. Steaming is generally regarded as the mildest treatment since only slightly decreased or even increased GL concentrations are reported, possibly due to increased extractability [6, 24].

1.4.3 Thermal degradation

Thermal degradation of GLs refers to the degradation induced by heat. The chemical structure of a GL and the time-temperature profile applied influence GL thermal degradation. Furthermore, differences in thermal stability depending on the pH of the reaction environment were observed [31, 32]. Ferrous ions (Fe^{2+}) promote thermal

degradation releasing ferric ions (Fe^{3+}) when added in molar excess with respect to the GL [33]. Increased GL thermal degradation was observed in the presence of the combination of vitamin C and ferrous ions (Fe^{2+}) in a buffer solution compared to the sole presence of ferrous ions, due to redox-cycling of ferric ions (Fe^{3+}) by vitamin C; however, degradation in the buffer solution was still lower when compared to a broccoli sprout matrix [34].

Differences in thermal degradation of chemically identical GLs have been described between five *Brassica* vegetables. The indolic GL glucobrassicin was six times more stable in red cabbage compared to Brussels sprouts and two times more stable in red cabbage compared to broccoli. Twentyfold difference in thermal stability was observed for the aliphatic GL gluconapin between broccoli and Brussels sprouts. Fifteen minutes of heating at 100°C hardly reduced the glucobrassicin concentration in red cabbage, but in Brussels sprouts only 40% of the initial glucobrassicin concentration remained [35]. By kinetic modelling Oerlemans *et al.* [32] estimated GL losses caused by thermal degradation in red cabbage for cooking and canning. Cooking of red cabbage at 100°C for 40 min was predicted to lead to moderate GL losses (38% loss of indolic, 8% loss of aliphatic GLs), whereas for canning at 120°C for 40 min severe GL losses (78% loss of total GLs) were predicted.

Studies in watercress seeds and broccoli sprouts identified simple nitriles as thermal degradation products [14, 36] that don't show biological activity at dietary levels [16]. These studies indicate that thermal degradation is an undesired process during heating since GLs are lost and no bioactive ITCs seem to be formed; however, studies only exist on seeds and sprouts and so far nitrile formation has not been shown in fully grown vegetables.

1.5 Factors influencing initial glucosinolate concentration

GL composition and content vary widely between *Brassica* vegetables, with genotype being the most important factor determining GL content and concentration. However, environmental factors and cultivation practices, such as soil, climate and fertilization, influence GL concentrations as well [6]. For instance, total GL concentrations were higher in spring than in autumn season in ten *Brassica oleracea* cultivars [37]. Furthermore, GL profiles change during plant development and vary significantly between different organs. The highest GL concentrations are reported in reproductive tissues and organs,

like seeds, siliques, flowers and developing inflorescences [38]. For example, the glucoraphanin concentration was between 7- and 12-fold higher in florets compared to the stem at the heading stage of broccoli. Glucoraphanin concentrations in broccoli decreased from seeds and seedlings till flowering stages between 80- and 1400-fold depending on the cultivar and its initial glucoraphanin concentration in the seeds [39]. Moreover, during a seven-day sprouting period glucobrassicin concentrations increased whereas the total aliphatic GL concentrations decreased in a study with broccoli, cauliflower, white, red and savoy cabbages [40].

1.6 Glucosinolate biosynthesis pathway

Amino acids are precursors of GLs; based on the structure of the amino acid, GLs are divided into three groups: a) aliphatic GLs are derived from alanine, leucine, isoleucine, methionine and valine, b) aromatic GLs are derived from phenylalanine and tryptophan and c) indolic GLs are derived from tryptophan. As shown in Figure 1.3, three main stages of the GLs biosynthesis can be defined: a) chain elongation of the amino acid precursor (only methionine, tryptophan), b) formation of the core GL structure and c) secondary modifications of the amino acid side chain [38].

In the first stage, chain elongation is catalysed by methylthioalkylmalate (MAM) synthases by condensation of acetyl-CoA, which is the methyl group donor of the reaction. Further chain elongations can occur by repeated cycles of acetyl-CoA condensation [6].

Five steps are required for the formation of the core GL structure, the second stage in GL biosynthesis. First, the amino acid precursor is oxidised to aldoximes by side chain-specific cytochrome P450 monooxygenases (cytochromes P450) of the CYP79 gene family. The aldoximes are further oxidised by cytochromes P450 of the CYP83 gene family to aci-nitro compounds or nitrile oxides, which are strong electrophiles that react spontaneously with thiols to form S-alkylthiohydroximates. Subsequently, a C-S lyase causes the cleavage of the S-alkylthiohydroximate conjugates into thiohydroximates. Thiohydroximates are unstable and reactive compounds, which are glycosylated and sulphated to form the GL core structure [38].

Secondary modifications of the side chain, oxidations, eliminations, alkylations and esterifications, are the final stages of GL biosynthesis. Methionine-derived GLs are modified by α-ketoglutarate-dependent dioxygenases, which are encoded by AOP genes, to produce alkenyl and hydroxyalkenyl GLs [38]. Furthermore, a flavin monooxygenase

that localises within the GS-OX locus is potentially responsible for the S-oxygenation in the GL side chain of aliphatic GLs. The hydroxylation of the indolic GL glucobrassicin caused by the gene CYP81F2 [41].

Next to the MAM, CYP79, CYP83 and AOP gene families, MYB genes are involved in GL biosynthesis. MYB28, MYB29 and MYB76 genes are transcription factors that control expression of the structural genes for production of aliphatic GLs [41]. Additionally, MYB34, MYB51 and MYB122 have been described as transcription factors for indolic GL synthesis, affecting CYP79 and CYP83 genes [42-44].
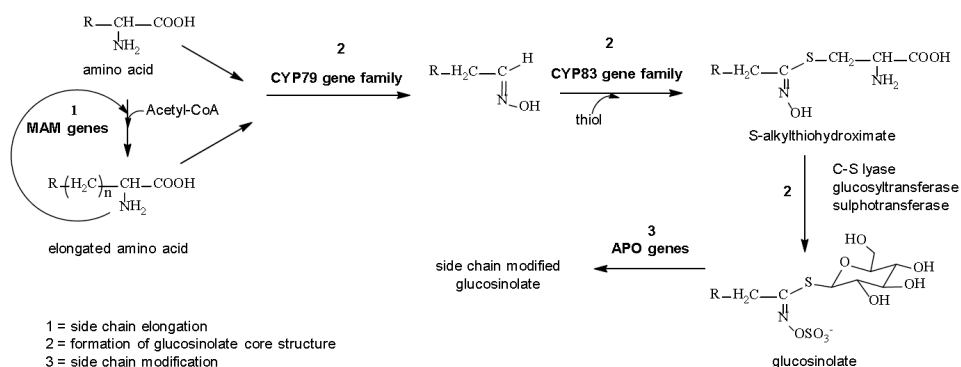

Figure 1.3: Main steps of the glucosinolate biosynthesis, modified from Verkerk *et al.* [6]

1.7 Breeding for increased initial glucosinolate content

Selection for resistance to herbivores and selection for milder flavours have likely indirectly influenced the levels and types of GLs present in *Brassica* vegetables. Certain GLs, namely sinigrin and gluconapin are responsible for the strong flavour of certain cabbages and Brussels sprouts. The selection for milder flavours has either led to a decreased total GL concentration or GL profiles are altered towards GLs that influence the flavour only slightly, like glucoraphanin and glucoiberin [6]. Direct breeding for enhanced levels of specific GLs has been performed in broccoli to increase the concentration of glucoraphanin and glucoiberin; their ITCs are potent inducers of phase II detoxification enzymes. This cultivar is marketed under the name *Beneforte* [45-47].

1

1.8 Phytochemicals in the food chain

Despite convincing beneficial effects on health, the fruit and vegetable consumption is below the recommended intake (400 g per capita per day) in Northern Europe and North America; only a small and negligible minority of the world's population reaches the recommended intake [1]. The development of new products with increased phytochemical concentrations might partly compensate for the low consumption of fruit and vegetables and provides new market opportunities for several actors in the food chain.

In Figure 1.4 the relations between the different actors in the food chain are illustrated using the example of GLs as one type of phytochemicals. Food technologists study the effects of processing methods on the level of GLs [6, 24]. Food technologists and horticulturists have common research areas in post-harvesting processes, such as ripening, deterioration and changes of GLs during storage [48]. Furthermore, horticulturists study the effect of cultivation practices and climate [49], which is a common research area with plant breeders who study genotype × environment interaction. Additionally, breeders investigate the genetic regulation of GL metabolism and apply this knowledge to develop vegetables with high GL content, such as a specific cultivar of broccoli that is rich in glucoraphanin [47].
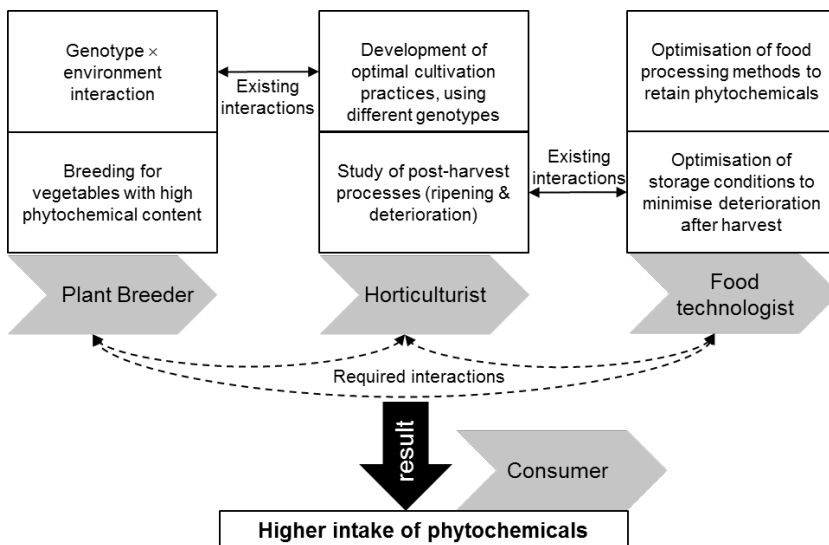

Figure 1.4: Existing and required interactions and overlapping research topics of glucosinolate research among different actors in the food chain.

Interdisciplinary research combining plant breeding and food technology, by using food processing parameters as breeding traits, appears to be a promising approach in the effort to retain high GL levels throughout the supply chain until the stage of consumption. Combined with breeding efforts to produce vegetables with increased initial phytochemical contents, this approach may lead to an increased intake of GLs.

1.9 Introduction to marker-assisted breeding

Many agriculturally important traits such as yield, disease resistance, sugar or vitamin content are controlled by many genes and are known as quantitative traits. The regions within genomes that are associated with a particular quantitative trait are known as quantitative trait loci (QTL) [50]. Phenotypic variation for the trait of interest is the first step to find the molecular nature of genetic differences [51].

Genetic differences between individual organisms or species can be revealed by using molecular markers. Mostly, molecular markers don't target genes themselves but are located in the proximity of genes. Three different groups of molecular markers can be identified, based on their method of detection: a) hybridisation, b) polymerase chain reaction and c) DNA-melting curves. The availability of DNA sequence information greatly facilitated the generation of molecular markers that belong to the groups b and c. Genetic linkage maps indicate positions and relative genetic distances of molecular markers. One important application of linkage maps is to identify QTLs associated with a trait of interest. For the construction of linkage maps, segregating plant populations are required that are derived from the cross of two homozygous parents with phenotypic variation in the trait of interest, as illustrated in Figure 1.5. From the F_1-generation, segregating populations can be developed by backcrossing with one parental line (backcross population), by regenerating plants from microspore mother cells (immature pollen grains) that spontaneously undergo chromosome duplication (doubled haploid population) and by self-pollination of the F_2-generation for six to eight generations (recombinant inbred line population). The genotypes of the individual plants of the segregating population differ in their composition of parental DNA, due to recombination between the homologous parental chromosomes. Furthermore, molecular marker alleles and parental alleles of genes segregate throughout the plant population which is caused by chromosome recombination (called crossing-over) during meiosis. The chromosome recombination can be exploited for genetic map

construction, as between genetic markers/genes that are physically located nearby the recombination frequency will be lower than between markers/genes that are located far apart. Segregating populations contain typically between 50 and 250 individuals (lines). Phenotypic information of the individual lines of the segregating population are combined with information of the molecular markers on the linkage map and statistical methods are applied to detect associations between phenotype and molecular markers, which are called QTLs [50].
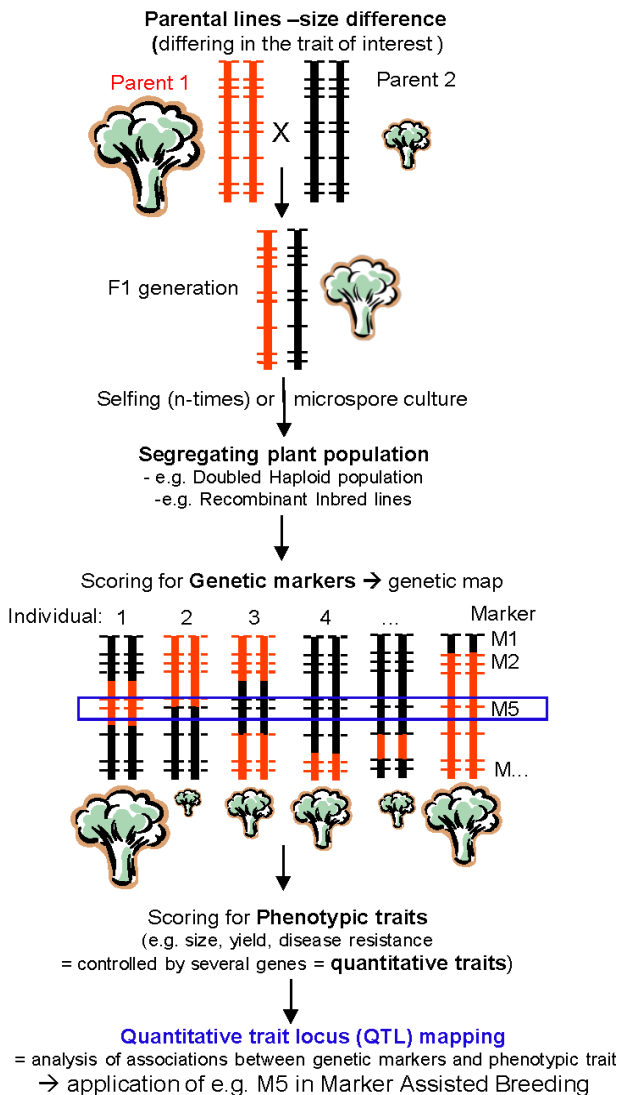

Figure 1.5: Introduction to quantitative genetics based on Collard, *et al.* [50] on the example of the quantitative trait *broccoli size* and the application of marker-phenotype-relations in marker-assisted breeding.

In addition, quantitative traits are not only affected by several genes, but also by the environmental conditions during growth. Plant breeders dissect these genetic and environmental effects, with the aim to cross QTLs with beneficial alleles into commercial cultivars and obtain varieties with improved genetic architecture of the target traits. The presence/absence of a molecular marker underlying the QTL of the trait of interest can be used in marker-assisted selection (MAS), which is a selection method based on the genotype at a specific position. The selection of a genotype for further breeding has several advantages a) that genotypes can be selected at the seedling stage before the phenotype is detectable, b) traits with low heritability can be selected and c) specific traits where phenotypic evaluation is laborious, costly or not possible can be tested. These advantages make MAS more efficient, reliable and cost-effective compared to phenotypic selection used in conventional breeding [50].

1.10 Objective and outline of the thesis

The objective of this thesis was to investigate genetic effects related to GL degradation during food processing to breed for vegetables with improved GL retention. This approach may complement the existing breeding efforts for vegetables with increased initial GL concentrations. Furthermore, genetic regions associated with GL degradation provide a starting point to investigate molecular mechanisms causing GL losses during food processing. This interdisciplinary approach can lead to higher GL concentrations at the stage of consumption and subsequently to a higher intake of health promoting compounds, as illustrated in Figure 1.4. Variation in the trait of interest is required to identify genetic loci associated with it [51]. Variation of GL thermal degradation of chemically identical GLs has been shown among different vegetables [35], which provides the basis to study genetic effects. Therefore, the focus of this thesis is on GL thermal degradation as an important mechanism of GL losses during food processing.

Genetic studies require the investigation of segregating populations or collections with high numbers of different genotypes. Studying food processing in a systematic way using a kinetic modelling approach involves several processing time points. The combination of the two methodologies requires high-throughput sample preparation to investigate genetic effects of GL degradation during food processing. GLs are most commonly analysed as desulpho-GLs, which necessitates an enzymatic desulphation step during sample preparation; however, desulphation protocols described in literature differ

1

widely. In **chapter 2** we investigated whether this desulphation could be optimized in microtiter plates to obtain reliable results with a high-throughput method.

Differences in thermal degradation of chemically identical GLs among different vegetables might be caused by differences in the vegetables matrix [35], which is affected by the genotype but also growing seasons, age of the plants, storage time and conditions after harvest. The first aim of the experiments described in **chapter 3** was to study the effect of the genotype and season on GL thermal degradation in well-characterised plant material. The second aim was to develop a rapid method to quantitatively determine degradation rate constants using kinetic modelling, which will provide the methodology to study the genetic and environmental effects on thermal GL degradation in a more efficient way.

In **chapter 4** the kinetic modelling approach was applied to a segregating plant population developed by crossing broccoli and Chinese kale to test the hypothesis if GL thermal degradation is genetically regulated.

Metabolites form, among other compounds, the matrix of a vegetable which may influence GL thermal degradation. Untargeted metabolomics approaches are used to measure as many metabolites as possible to obtain patterns or fingerprints of processed food samples [52, 53]. Since there is yet no knowledge on which metabolites can influence GL thermal degradation, an untargeted metabolomics approach is a promising tool to associate differences in metabolite composition to differential GL thermal degradation rates. In **chapter 5** this approach was applied to identify metabolites associated with GL thermal degradation using random forest regression and QTL mapping.

The final Chapter (**chapter 6**) summarises the main finding of this thesis, followed by a critical discussion and implications for further research. Furthermore, the possibilities and benefits of this interdisciplinary approach, breeding for food processing traits, are discussed.

CHAPTER 2

PITFALLS IN DESULPHATION OF GLUCOSINOLATES IN A HIGH-THROUGHPUT ASSAY

2

Abstract

Glucosinolates are phytochemicals with health promoting properties. Determination as desulpho-glucosinolates is widely used and a desulphation in microtiter plates has been applied to reach high-throughput. The use of various sulphatase concentrations and volumes throughout literature necessitates the identification of an appropriate desulphation procedure in microtiter plates. High sulphatase concentrations (≥ 15 mg/mL) decreased the concentration of the internal standard glucotropaeolin, whereas the other glucosinolates were less affected. Due to the calculation based on the recovery of glucotropaeolin, this leads to an overestimation of glucosinolate concentrations after desulphation with high sulphatase concentrations. A glucosidase side-activity, present in the crude sulphatase powder, is likely causing this phenomenon. At lower sulphatase concentrations (1 mg/mL) glucoiberin and glucoraphanin were insufficiently desulphated. Combining these effects, results in a small range of applicable sulphatase concentrations. A purified sulphatase preparation resulted in good recoveries for a diversity of samples and is hence recommended for high-throughput desulphation in microtiter plates.

2.1 Introduction

Glucosinolates (GLs) are a group of secondary plant metabolites, which are almost exclusively found in the *Brassicaceae* family. Chemically, GLs are *β*-thioglucoside-*N*-hydroxysulphates with a side chain (R) and a sulphur linked *β*-D-glucopyranose moiety [8]. Depending on the structure of the side chain, GLs are divided into different classes: aliphatic, aromatic and indolic GLs [6]. GLs are not directly bioactive, but after conversion into breakdown products by the enzyme myrosinase (*β*-thioglucosidase, E.C. 3.2.1.147) several biological activities are ascribed to them [8]. Breakdown products (BDP) are involved in the defence system of the plant [11], are responsible for the typical flavour of *Brassica* vegetables and some breakdown products exhibit health promoting properties (reviewed by Verkerk *et al.* [6]). GLs have been linked with the strongly inverse association of *Brassica* consumption and several cancer risks [6]. Due to the health promoting properties of GLs and their BDP, research has been performed to optimise the GL content of *Brassica* vegetables, specifically breeding for vegetables with altered GL content and reduction of GL losses during industrial and domestic processing. The effects caused by breeding, different post-harvest treatments, storage and processing are reviewed by Verkerk *et al.* [6].

In order to determine changes in GL content due to breeding and food processing, an accurate and high-throughput determination method is needed. Different methods for GL determination have been developed, but the HPLC analysis of desulpho-GL is most generally used [54]. In this method, GL extracts are applied to ion-exchange columns, where GLs are bound due to the negative charge of their sulphate group. The columns are incubated with the enzyme sulphatase, removing the sulphate group and subsequently the GLs can be eluted with water. The publication of a protocol to determine GLs in rapeseeds by the International Standard Organization harmonized the GL analysis in 1992 (ISO 9167-1). Since then, pitfalls in the method were identified [55] and improvements were suggested [56]. Furthermore, the method has been modified and applied for different samples, which generated confusion about the reliability of GL determinations [54]. In order to reach a high sample throughput, desulphation has been performed in 96-well microtiter plates [57]. Compared to the ISO 9167-1 method, a reduced amount of ion-exchange material is used in microtiter plates, which can lead to new pitfalls. Consequently, it is important to use an appropriate amount of sulphatase, because desulphation is a crucial step in the sample preparation of desulpho-GLs.

2

Desulphation protocols found in literature differ widely. According to ISO 9167-1, sulphatase has to be purified prior to application on the ion-exchange column. Simple purification procedures have been suggested by Wathelet *et al*. [56] and applied by Van Eylen *et al*. [58]. In some laboratories sulphatase is purified prior to application [13, 58-63]. In other laboratories purification is omitted and sulphatase is used as a crude preparation [32, 57, 64-67]. Furthermore, sulphatase concentrations range from 1 mg/mL [65] to 25 mg/mL [32] and volumes of enzyme solutions range from 10 μL [57] to 500 μL [68, 69]. In several publications the desulphation procedure is not explicitly described; instead references to different articles are given, which adds more confusion. A review discussing several aspects of the GL analysis has been published recently by Clarke [70]; however, desulphation procedures were not discussed. Due to the wide range of applied sulphatase concentrations, the aim of this study is to determine an appropriate desulphation procedure for the preparation of desulpho-GLs in microtiter plates. Here we report on our findings and we recommend a simple purification procedure of the sulphatase for accurate analysis of desulpho-GLs with high-throughput and in a smaller scale.

2.2 Materials and Methods

2.2.1 Plant material

Red cabbage cultivar "Buscaro F1" (Bejo, The Netherlands) was sown in late April 2009, grown for 6 weeks in the greenhouse, then transplanted into the field (Wageningen, The Netherlands) and heads were harvested in the middle of September 2009. A plant line (AG2186) of a doubled haploid population (AGDH) developed by crossing *B. oleracea* spp. *alboglabra* (A12DHd) and a *B. oleracea* ssp. *italica* (GDDH33) [71] was sown in March 2009 and leaves were harvested 8 weeks after sowing. Plants were grown in the greenhouse in Wageningen, The Netherlands.

After harvesting, red cabbage heads and leaves of AG2186 were brought to the laboratory and immediately chopped (into approximately 3 x 3 cm pieces), frozen with liquid nitrogen, freeze-dried (GRInstruments, Model GRI 20-85 MP 1996, The Netherlands) and ground into a fine powder. Dried plant powder was stored in closed containers at -20°C in the dark.

2.2.2 Chemicals

Solvents used for extraction and chromatography were of HPLC grade and bought from Biosolve (Valkenswaard, The Netherlands). The DEAE Sephadex-A25 and sinigrin (SIN) were purchased from Sigma Aldrich (Zwijndrecht, The Netherlands). The internal standard glucotropaeolin was purchased from the Laboratory of Biochemistry, Plant Breeding and Acclimatization Institute at Radzikow, Błonie, Poland. Four different batches of sulphatase from *Helix pomatia* (Sigma Aldrich, cat. no. S9626) were used for the experiments with the lot no.: 105K3796 (sulphatase A), 058K3805 (sulphatase B), 029K3782 (sulphatase C) and 089K3797 (sulphatase D). Names in parentheses will be used in the following sections to refer to the different batches. The sulphatase activities were specified by the supplier as follows: sulphatase A: 14 000 U/g, sulphatase B: 24 190 U/g, sulphatase C: 23 500 U/g and sulphatase D: 26 290 U/g. The β-glucuronidase activities were given as: sulphatase A: 498 800 U/g, sulphatase B: 1 521 800 U/g, sulphatase C: 1 774 000 U/g and sulphatase D: 2 921 000 U/g.

2.2.3 Glucosinolate extraction

The extraction procedure was performed according to ISO 9167-1:1992 with slight modifications. Freeze dried *Brassica* powder (20 mg) was weighed into 2 mL tubes, extracted with 1 mL hot methanol (70%, 75°C) for 10 min and centrifuged for 5 min at room temperature at 16 000g. Glucotropaeolin was added as an internal standard (3 mM, 50 µL to red cabbage, 25 µL to AG2186) during the first extraction. The pellet was re-extracted in the same way and the combined supernatants were used for desulphation. A sinigrin-glucotropaeolin-mixture was prepared by dissolving different volumes of sinigrin and glucotropaeolin standard solutions (3 mM) in 70% methanol.

2.2.4 Desulphation of glucosinolates

Desulphation in microtiter plates

Desulphation of GLs was performed according to ISO 9167 1:1992 and Kliebenstein *et al.* [57] with modifications. Ion-exchange columns were built with DEAE Sephadex-A25 in 96-well microtiter-filter plates (Multiscreen® Deep Well Solvinert Filter Plate, Millipore, cat. No. MDRL N04 10). In order to pack the columns, to remove rinsing liquids, to load and to elute the samples, the microtiter plates were centrifuged for 5 min at 900g. Sephadex (10 g) was suspended in 80 mL 2 M acetic acid, swollen overnight and 80 mL of 2 M acetic acid were added the following day. Ion-exchange columns were

built by pipetting 500 μL of swollen Sephadex suspension (equates to about 30 mg dry Sephadex) into each well of the microtiter plate and the columns were rinsed twice with 300 μL water. The GL extracts were loaded in two aliquots of 1 mL on the ion-exchange columns and subsequently rinsed twice with 500 μL sodium acetate buffer (20 mM, pH = 4.0). For desulphation of GLs, 100 μL freshly prepared sulphatase solution in varying concentrations (1 – 25 mg/mL, purified) were added to the ion-exchange columns and incubated for 16 h at 25°C. Desulpho-GLs were eluted after incubation in microtiter plates by adding 100 μL water twice.

Purification of sulphatase

Purification of sulphatase was performed by a simplification of the procedure of Van Eylen *et al.*, [58]: 70 mg sulphatase were dissolved in 3 mL water and precipitated with 12 mL ethanol. After centrifugation (1500g, room temperature), the pellet was dissolved in 2 mL water and passed through a 0.5 mL DEAE Sephadex column built in syringes. Before sulphatase was passed through the Sephadex column, the column was rinsed twice with 0.5 mL water and 0.5 mL sodium acetate buffer (20 mM, pH = 4.0). The sulphatase was purified directly prior to the application on the ion-exchange column in the microtiter plates to avoid changes in activity due to storage.

Desulphation in larger scale

Desulphation in microtiter plates was compared to desulphation at larger scale in 2 mL-syringes as described by Verkerk *et al.* [67]. Ion-exchange columns were built with DEAE Sephadex in syringes with 3 mL swollen Sephadex suspension (equates to about 190 mg dry Sephadex) and rinsed twice with 1 mL water. Samples were loaded in two portions of 1 mL and the columns were rinsed twice with 1 mL sodium acetate buffer (20 mM, pH=4.0). The ion-exchange columns were incubated with 75 μL sulphatase (25 mg/mL) for 16 h at 25°C. The scale of this method is comparable to the scale of the ISO 9167-1 method. The desulphation in microtiter plates is at a 6-fold smaller scale.

The data presented in the results section refer to the desulphation in microtiter plates, if not mentioned otherwise.

2.2.5 HPLC analysis

Desulpho-GLs were analysed by high-performance liquid chromatography (HPLC) as described by [67]. The HPLC analysis was conducted on a gradient system HPLC (Spectra Physics) using a GraceSmart RP C18 column (150 mm x 4.6 mm) with a flow rate of

1 mL/min, column temperature of 20°C and an injection volume of 20 µL. Solvent A consisted of 0.05% tetramethylammoniumchloride (Merck 822156) dissolved in water and solvent B consisted of 0.05% tetramethylammoniumchloride dissolved in water/acetonitrile (60:40, v/v). The total running time amounted to 31 min and the gradient was as followed: 100% A and 0% B for 1 min, then in 20 min to 0% A and 100% B, in 5 min back to 100% A and 0% B, after which the column was equilibrated for 5 min with solvent A. Detection was performed with a DAD detector (Spectra System UV 6000 LP) at 229 nm and UV spectra were recorded from 220-550 nm. Identification of GLs was based on retention times compared to standard GLs (sinigrin, glucotropaeolin), GLs present in the certified rapeseeds (Colza, European Reference material ERM®-BC367), GLs present in reference vegetables and UV-spectra of GLs [72]. GLs were quantified against the internal standard glucotropaeolin using relative response factors as given in ISO 9167-1:1992, recalculated to glucotropaeolin.

2.2.6 Statistics
Averages, standard deviations were calculated and t-test (p < 0.05) was performed using Microsoft Excel.

2.3 Results and Discussion

2.3.1 Effect of sulphatase concentration on standard solutions
The glucosinolates (GLs) sinigrin and glucotropaeolin are recommended as internal standards in the ISO 9167-1 method [73]. Thus, sinigrin and glucotropaeolin were used as standards to check the performance of the GL determination with desulphation in microtiter plates. The desulphation in microtiter plates provides a high-throughput and less Sephadex material is used to build the ion-exchange columns. This sinigrin-glucotropaeolin-mixture was loaded on ion-exchange columns and incubated with sulphatase concentrations of 25 mg/mL and 5 mg/mL (both sulphatase batch C). sinigrin is present in rape seeds, mustard and some *Brassica oleracea*, such as red cabbage, white cabbage, savoy and cauliflower [40], and thus cannot be used as internal standard in these samples. For that reason, glucotropaeolin is used as internal standard to calculate the sinigrin concentration in the sinigrin-glucotropaeolin-mixture.

In Figure 2.1, the chromatogram overlay of the desulphated sinigrin-glucotropaeolin-mixture shows a 33% lower glucotropaeolin area after desulphation with 25 mg/mL sulphatase compared to the 5 mg/mL sulphatase treatment, whereas sinigrin areas differed only by 0.8%. Taking into account that sinigrin has a relative response factor (RRF) relative to glucotropaeolin of 1.05, and the same concentration of the two GLs was used for desulphation, the peak areas of sinigrin and glucotropaeolin were expected to be similar within one sample. The glucotropaeolin area was 34% smaller than the area after desulphation with 25 mg/mL, while the areas were the same after desulphation with 5 mg/mL sulphatase (Figure 2.1). Consequently, using glucotropaeolin as an internal standard for the calculation of the sinigrin concentration resulted in a higher sinigrin concentration after desulphation with 25 mg/mL than after desulphation with 5 mg/mL. As a result, desulphation with 25 mg/mL sulphatase caused a substantial overestimation of the sinigrin concentration. The impact of different sulphatase concentrations on calculated sinigrin recoveries is shown in Table 2.1.
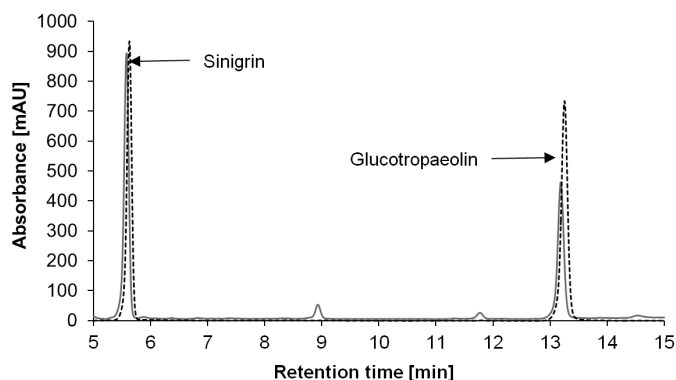

Figure 2.1: Chromatogram overlay of a sinigrin-glucotropaeolin-mixture, treated with two different sulphatase concentrations. (———) 25mg/mL sulphatase, (-----) 5 mg/mL sulphatase

Table 2.1: Calculated sinigrin recovery in a sinigrin-glucotropaeolin-mixture after desulphation with sulphatase concentrations ranging from 25 mg/mL to 1 mg/mL and a purified preparation (35 mg/mL without accounting for losses during the purification). The sinigrin concentration on the ion-exchange column amounted to 0.075 µmol and the glucotropaeolin concentration amounted to 0.076 µmol. n = number of independent repeats in duplicate.

Sulphatase concentration	Calculated recovery desulpho-sinigrin:	
	Average (%)	Relative standard deviation (%)
25 mg/mL (n = 4)	212	50
15 mg/mL (n = 3)	149	26
5 mg/mL (n = 4)	116	4
1 mg/mL (n = 2)	104	2
purified (n = 2)	106	11

The best calculated sinigrin recoveries were obtained after desulphation with 1 mg/mL sulphatase (104%) and with the purified sulphatase (106%). With increasing sulphatase concentration, the calculated recoveries increased; the maximum was observed after desulphation with 25 mg/mL (212%). Furthermore, Table 2.1 shows that the repeatability decreased at higher sulphatase concentrations. The high standard deviation after desulphation with 25 mg/mL can be explained by the observed instability of glucotropaeolin in the sample solution during the time in the auto-sampler until HPLC analysis. When the time until analysis amounted to 26 h, the measured glucotropaeolin area was lower than after 9 h in the auto-sampler. Consequently, 26 h in the auto-sampler led to lower glucotropaeolin concentrations. This effect was not observed for desulphation with 5 mg/mL, 1 mg/mL and purified sulphatase.

The small-scale desulphation in microtiter plates was compared to a larger scale desulphation in syringes, using six times more ion-exchange material according to Verkerk *et al.* [67] and the highest sulphatase concentration (25 mg/mL). Experiments at this larger scale, performed with the sinigrin-glucotropaeolin-mixture of the same concentration as in the microtiter plates, resulted in a calculated sinigrin recovery of 116% ± 5%. A possible reason for the differences between syringes and microtiter plates is the different amount of ion-exchange material used and consequently, a different efficiency of the enzymatic reactions.

From these results we can conclude that for a reliable GL analysis in microtiter plates a reduction in the sulphatase concentration or a purification is required. A too low sulphatase concentration, however, could result in incomplete desulphation. A study performed by Fiebig [55] showed that a low sulphatase activity (0.05 U/mL) is not sufficient to desulphate solutions of 20 mmol/L GLs on ion-exchange columns after 23 h of incubation. This study determined a minimum sulphatase concentration that is needed for the complete desulphation of GLs. An upper limit was not suggested, because after complete desulphation with a higher sulphatase activity (0.5 U/mL), peak areas did not change with longer incubation times. Based on the results of Fiebig [55] our observations of breakdown of glucotropaeolin at high sulphatase concentrations were not expected; however, the samples in the study of Fiebig [55] were prepared according to the ISO 9167-1 method, with a purified sulphatase and a higher amount of ion-exchange material than in the microtiter plates.

In summary, two phenomena were observed when high sulphatase concentrations were used for desulphation in microtiter plates: 1) different peak areas of sinigrin and glucotropaeolin, although the applied concentration and response factor were similar and 2) an instability of the desulpho-glucotropaeolin in the sample solution until HPLC analysis. Both observations indicate a side activity of the sulphatase, which will be discussed in the following section.

2.3.2 Effect of sulphatase concentration on red cabbage samples

The effect of sulphatase concentration on the internal standard glucotropaeolin in red cabbage samples was similar to what was shown for the standard solution (sinigrin-glucotropaeolin-mixture). When the same amount of glucotropaeolin was added to a red cabbage sample as was to the sinigrin-glucotropaeolin-mixture, the glucotropaeolin areas were not equal after desulphation with a high sulphatase concentration (25 mg/mL). In contrast with a low sulphatase concentration (5 mg/mL) the areas were equal and higher than after desulphation with high sulphatase concentration. Apparently, desulphation is differently affected in sinigrin-glucotropaeolin-mixture and red cabbage due to the different type of GLs and the higher total GL concentration in red cabbage extract.

In order to study the effects of various sulphatase concentrations (25 mg/mL – 1 mg/mL, purified) on all GLs in red cabbage, peak areas are presented in Figure 2.2. Different sulphatase concentrations did not affect the peak areas of most of the aliphatic GLs, except for glucoiberin and glucoraphanin. After desulphation with 1 mg/mL sulphatase, the peak areas of glucoiberin and glucoraphanin were 60% smaller than after desulphation with the purified sulphatase. The indolic GLs glucobrassicin, neoglucobrassicin and the benzyl GL glucotropaeolin were affected by different sulphatase concentrations. Peak areas were decreasing with increasing sulphatase concentrations and highest peak areas were measured after desulphation with the purified sulphatase. The peak area of glucotropaeolin, which is, as mentioned above, widely used as internal standard, was 24% smaller after desulphation with 25 mg/mL sulphatase compared to the purified sulphatase. Glucobrassicin was affected to a lesser extent, though the peak area was 10% smaller after desulphation with 25 mg/mL compared to the purified sulphatase.

Also, an additional twofold dilution of purified sulphatase resulted in a significant decrease in peak areas for the two aliphatic GLs glucoiberin and glucoraphanin ($p < 0.05$; data not shown). Summing up, desulphation of GLs in red cabbage with purified

sulphatase resulted in the highest peak areas; however, peak areas of all GLs were not significantly different between desulphation with 5 mg/mL and purified sulphatase (p < 0.05). Further dilution of the purified sulphatase was not possible, because the amount of glucoiberin and glucoraphanin would be underestimated.
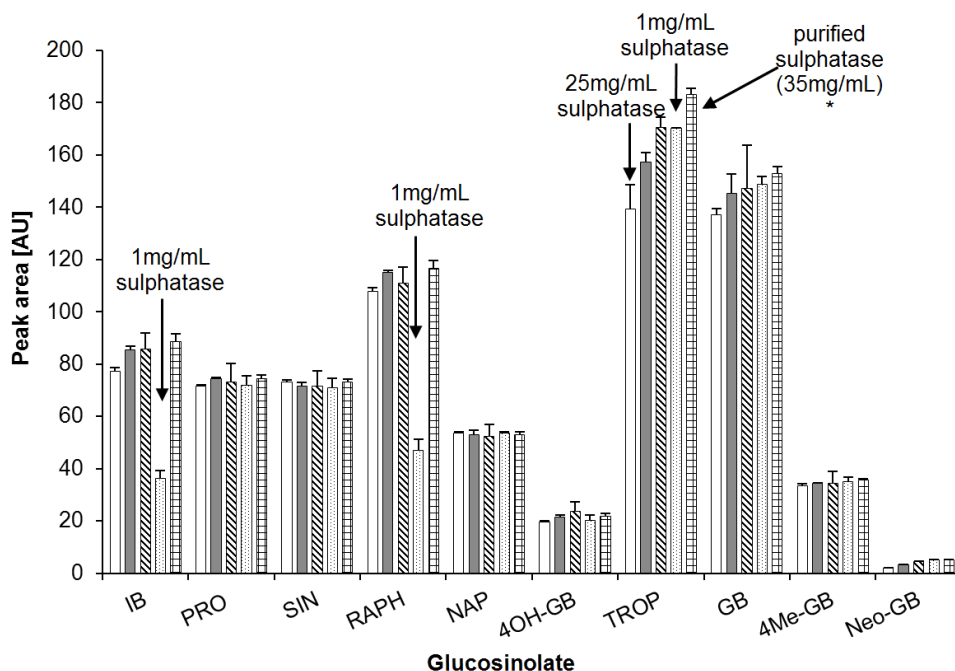

Figure 2.2: Peak areas of glucosinolates present in red cabbage after desulphation with different sulphatase concentrations: ☐ 25 mg/mL, ▨ 15 mg/mL, ▨ 5 mg/mL, ▨ 1 mg/mL, ▨ purified (35 mg/mL*). Glucotropaeolin was added as internal standard. The bars show average values of duplicates and error bars represent the standard deviations. (IB) glucoiberin, (PRO) progoitrin, (SIN) sinigrin, (RAPH) glucoraphanin, (NAP) gluconapin, (4OH-GB) 4-hydroxy-glucobrassicin, (TROP) glucotropaeolin, (GB) glucobrassicin, (4-Me-GB) 4-methoxy-glucobrassicin, (Neo-GB) neoglucobrassicin.
* Sulphatase concentration, not accounting for losses during purification

The differences in peak areas after desulphation with different sulphatase concentrations reflect the differences in total GL concentration calculated with glucotropaeolin as internal standard. Based on the calculated total GL concentration (the sum of the calculated individual GLs), it would be concluded, falsely, that 25 mg/mL sulphatase resulted in the highest GL concentration and is consequently the best sample preparation (Figure 2.3). However, the decrease in glucotropaeolin peak area with increasing sulphatase concentration caused an overestimation of the individual and total GL concentration. Desulphation with purified sulphatase resulted in the highest peak areas for all GLs,

including the internal standard glucotropaeolin, hence gave the most accurate result.

Combining the results of red cabbage samples and the sinigrin-glucotropaeolin-mixture, the purified sulphatase gave best results for sinigrin-glucotropaeolin-mixture and red cabbage, whereas 1 mg/mL would be appropriate for sinigrin-glucotropaeolin-mixture and 5 mg/mL for red cabbage. Using 5 mg/mL sulphatase could overestimate the GL concentration in a sample with a low number of GL and of a low concentration, as it was shown for the sinigrin-glucotropaeolin-mixture. However, desulphation with 1 mg/mL appeared not complete for the aliphatic GLs glucoiberin and glucoraphanin in a sample with more GLs present and a higher total GL concentration. Taking these aspects into account, the purified sulphatase is applicable for a broader range of samples and is recommended for the desulphation of GLs with high-throughput in microtiter plates.
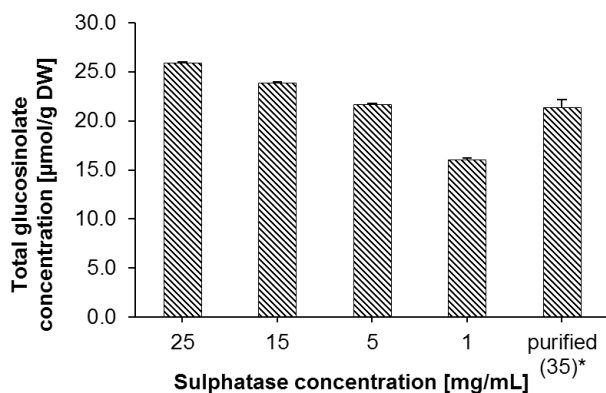

Figure 2.3: Calculated total GL content (sum of individual GLs) in red cabbage after desulphation with sulphatase concentrations ranging from 25 mg/mL to 1 mg/mL and purified sulphatase preparation (35 mg/mL*). The bars show one repeat in duplicate, error bars show standard deviations.
* Sulphatase concentration, not accounting for losses during purification

In accordance with our findings, Wathelet *et al*. [54] stated that the sulphatase-GL-ratio is crucial for the analytical result. Although the reason was not known, the authors suggested differences in desulphation velocity and feedback inhibition, depending on the type of GL, as possible reasons for this phenomenon. In our study, desulphation of only glucoiberin and glucoraphanin, which consist both of a sulfinyl side chain and differ only by one C-atom, was insufficient at the lowest sulphatase concentration (1 mg/mL). A lower affinity of the sulphatase to glucoiberin and glucoraphanin and consequently a lower desulphation velocity could explain this observation. Wathelet *et al*. [54] suggested a feedback inhibition for samples containing glucoiberin, because an

elution and second incubation step were needed to complete the desulphation. Due to the setup of our experiments the exact mechanism leading to lower glucoiberin and glucoraphanin concentrations remains speculative.

In contrast to glucoraphanin and glucoiberin, peak areas of glucotropaeolin were decreasing with increasing sulphatase concentrations, whereas the sinigrin peak areas remained constant. Since Fiebig [55] reported a higher desulphation velocity of glucotropaeolin than sinigrin, our observation indicates a side reaction present in the crude enzyme, breaking down glucotropaeolin at high sulphatase concentrations (15 and 25 mg/mL). An inhibition of the sulphatase activity at high enzyme concentrations seems unlikely, because inhibition should be observed for more GLs than glucotropaeolin, especially since glucotropaeolin is preferably hydrolysed by the sulphatase. Furthermore, the instability of desulpho-glucotropaeolin in the sample solution during storage in the auto-sampler until HPLC analysis pointed to a possible side reaction. The sulphatase powder used shows a β glucuronidase side activity that is specified by the supplier Sigma-Aldrich. If the β-glucuronidase reacts with a glucose residue as well, a myrosinase-like breakdown of GLs could occur. In studies of Kwon *et al*. [74] the same sulphatase powder was used to remove a glucose moiety from the glycoside-isoflavon genistin, because it contains β-glycosidase side activity. The side-chain of glucotropaeolin consists of a benzyl-moiety, as does genistin that consists of several benzyl groups. Due to the similar structure, the β-glycosidase side activity reported for genistin, could lead to a breakdown of glucotropaeolin at high sulphatase concentrations. Reacting preferably with molecules with benzyl rings, the glucosidase side activity explains that glucotropaeolin was mostly adversely affected at high sulphatase concentrations, since it was the only benzyl GL present in the red cabbage sample.

2.3.3 Effect of sulphatase batch

In addition to the effect of different sulphatase concentrations on the GL analysis, the performance of desulphation in microtiter plates was tested for three different enzyme batches (sulphatase A, B and C). Solutions of the three sulphatase batches were used in varying concentrations for desulphation of a sinigrin-glucotropaeolin-mixture. The sinigrin recovery was calculated based upon the glucotropaeolin area as internal standard, as described in section "Effect of sulphatase concentration on standard solutions". Lower glucotropaeolin areas, due to a potential side activity of the sulphatase, led to a too high calculated sinigrin recovery.

As shown in Figure 2.4, the sulphatase batch had a large impact on the calculated sinigrin recovery. For sulphatase batch A and C, the too high recoveries became lower with decreasing sulphatase concentrations; however, for sulphatase batch B, recoveries of nearly 100% were obtained for all sulphatase concentrations. This experiment demonstrates that different enzyme batches from the same supplier affects the accuracy of GL analysis.
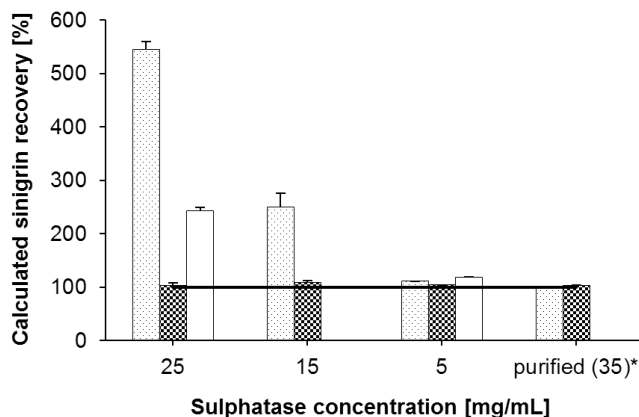

Figure 2.4: Calculated sinigrin recovery after desulphation with different sulphatase batches: ▨ sulphatase A, ▨ sulphatase B, ☐ sulphatase C. Sulphatase concentrations ranged from 25 mg/mL to 1 mg/mL and a purified sulphatase preparation (35 mg/mL*). The sinigrin concentration on the ion-exchange column amounted to 0.15 μmol and the glucotropaeolin concentration amounted to 0.14 μmol. The bars represent averages of one repeat in duplicate and error bars represent standard deviations. * Sulphatase concentration, not accounting for losses during purification

The specification of the sulphatase and β-glucuronidase activity by the supplier cannot explain the obtained sinigrin recoveries. Sulphatase A had the lowest sulphatase activity (14 000 U/g solid) and resulted in too high recoveries. Keeping in mind that glucotropaeolin is faster desulphated than sinigrin [55] a calculated sinigrin recovery below 100% could be explained for sulphatase A, but not a too high recovery. For sulphatase B the highest sulphatase activity of the three batches (24 190 U/g solid) was given, and desulphation resulted in recoveries of nearly 100% for all sulphatase concentrations. Due to that fact, an inhibition of the desulphation at high enzyme concentrations is unlikely, although the discussion is limited by the fact that sulphatase activity is not determined on GLs by the supplier. The glucuronidase side activities are also not correlating with the observed sinigrin recovery pattern. A slightly higher glucuronidase activity in sulphatase C could

explain the too high results after desulphation with sulphatase C compared to sulphatase B, but the three times lower glucuronidase activity of sulphatase A is conflicting.

As a result, no conclusion about the expected accuracy of the desulphation can be drawn from the specifications given by the supplier. Purification of the sulphatase leads to accurate recoveries, independent from the sulphatase batch and is hence recommended prior to desulphation.

2.3.4 Repeatability of GL determination with purified sulphatase

The repeatability of the GL determination with desulphation in microtiter plates was assessed by repeating the analyses (extraction, desulphation and measurement) ten times in duplicate for a high GL sample (red cabbage) and a low GL sample (plant line AG2186 of a doubled haploid population). The two samples were stored at -20°C between the analyses. A new batch of sulphatase (batch D) was used and purified directly before desulphation. Both samples contained the same individual GLs: the aliphatic GLs glucoiberin, sinigrin, progoitrin, glucoraphanin, gluconapin and the indolic GLs glucobrassicin, 4-hydroxy-glucobrassicin, 4-methoxy-glucobrassicin and neoglucobrassicin, but in different concentrations. The total GL content, calculated as the sum of the individual GL concentrations (Figure 2.5), differed by 3-fold.
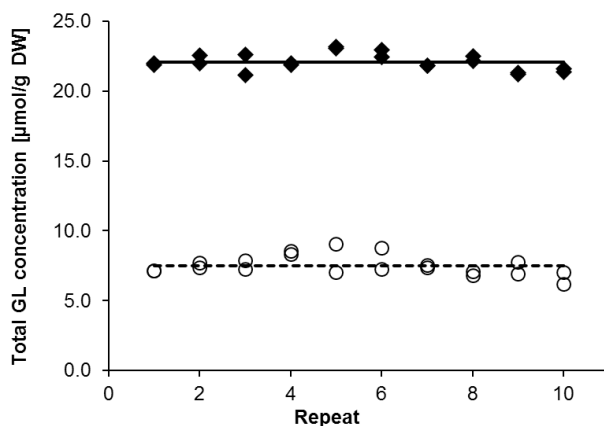
Figure 2.5: Variation of the calculated total glucosinolate concentration in red cabbage and AG2186 during ten repeats in duplicate. Desulphation was performed with purified sulphatase (35 mg/mL without accounting losses during the purification) and the total GL concentration was calculated as the sum of individual GLs. Symbols represent single values: ○ AG2186, ◆ red cabbage. Lines represent the average of all repeats: ▬ red cabbage, ... AG2186.

In the ISO 9167-1 method [73] the repeatability of the GL determination is specified to a maximum difference between two independent single test results of 2 μmol/g for samples containing less than 20 μmol GLs/g and to a maximum difference of 4 μmol/g for samples containing 20 μmol/g to 35 μmol GLs/g. The average total GL concentration for red cabbage amounted to 22.1 μmol/g DW with a maximum difference measured between duplicates of 1.5 μmol/g DW. The average total GL concentration AG2186 amounted to 7.5 μmol/g DW with a maximum difference between duplicates of 2.1 μmol/g DW. For both samples the maximum differences have met the ISO requirement. Regarding the overall range of the 10 repeats in duplicate the measured range for red cabbage amounted to 2.0 μmol/g DW and 2.9 μmol/g DW for AG2186. For red cabbage the ISO requirements are met again, whereas the range of AG2186 slightly exceeds the ISO specifications.

The relative standard deviation of all the repeats amounted to 2.7% for red cabbage and 9.5% for sample AG2186. The relative standard deviations show a good repeatability of the GL determination for two samples, measured with purified sulphatase over a period of three months.

The minimum total GL concentration measured in red cabbage in the ten repeats amounts to 21.2 μmol/g DW and the maximum amounts to 23.2 μmol/g DW. Comparing this range to the GL concentrations determined with different sulphatase concentrations during desulphation (Figure 2.3) shows that the range of the ten repeats is much smaller than the differences determined with varying sulphatase concentrations. The GL concentrations determined with 25 mg/mL, 15 mg/mL and 1 mg/mL sulphatase lie outside the measured range of the ten repeats, whereas the GL concentrations determined with 5 mg/mL and purified sulphatase lie within the measured range. Hence, the desulphation with 25 mg/mL, 15 mg/mL and 1 mg/mL results in differences in total GL concentration far exceeding the variation observed during the ten repeats in duplicate and can be seen as true differences.

2.4 Conclusion

The effect of different sulphatase concentrations and a purified sulphatase preparation on the accuracy of the GL determination in microtiter plates was tested. The presented data demonstrate that the applied sulphatase concentration is crucial for the analytical result, which is in accordance with previous reports [54, 55]. However, our data show for the first time an inverse effect of sulphatase concentration on the peak area of a GL. Tests, performed with a sinigrin-glucotropaeolin-standard and red cabbage, resulted in decreased peak areas of glucotropaeolin at high sulphatase concentrations (15 mg/mL and 25 mg/mL), whereas most other GLs were hardly affected. Hence, the calculation of the GL concentration using glucotropaeolin as an internal standard leads to a substantial overestimation of GLs after desulphation with high sulphatase concentrations. A glucosidase side activity, reacting preferentially on glucotropaeolin, is suggested as a possible reason for the observed phenomenon. Surprisingly, the glucotropaeolin degradation was not detected in one of the three applied sulphatase batches.

On the one hand, the degradation of glucotropaeolin at high sulphatase concentration determines an upper limit of the sulphatase concentration, but on the other hand a minimum concentration is needed for the complete desulphation of all GLs. The lowest sulphatase concentration (1 mg/mL) led to an insufficient desulphation of the aliphatic GLs glucoiberin and glucoraphanin and therefore to an underestimation of their concentration. As a result, a small range of crude sulphatase concentrations can be applied for GL determination. Furthermore, the optimal sulphatase concentration depends on the amount and the number of GL present in the sample. The application of a purified sulphatase preparation resulted in good recoveries for the standard solutions and the red cabbage samples. In conclusion, the use of purified sulphatase is recommended for determination of GL in microtiter plates for a broad range of samples and independently from the sulphatase batch.

2.5 Acknowledgements

CHAPTER 3

RAPID ESTIMATION OF GLUCOSINOLATE THERMAL DEGRADATION RATE CONSTANTS IN LEAVES OF CHINESE KALE AND BROCCOLI (*B. OLERACEA*) IN TWO SEASONS

Abstract

Kinetic modelling was used as a tool to quantitatively estimate glucosinolate thermal degradation rate constants. Literature shows that thermal degradation rates differ in different vegetables. Well-characterised plant material, leaves of broccoli and Chinese kale plants grown in two seasons, was used in the study. It was shown that a first order reaction is appropriate to model glucosinolate degradation independent from the season. No difference in degradation rate constants of structurally identical glucosinolates was found between broccoli and Chinese kale leaves when grown in the same season. However, glucosinolate degradation rate constants were highly affected by the season (20% to 80% increase in spring compared to autumn). These results suggest that differences in glucosinolate degradation rate constants can be due to variation in environmental as well as in genetic factors. Furthermore, a methodology to estimate rate constants rapidly is provided to enable the analysis of high sample numbers for future studies.

3.1 Introduction

Kinetic modelling is a mathematical tool with which chemical changes in foods can be described in a quantitative way [75]. Modelling has been used to study the effect of food processing on the concentration of health promoting compounds, as the awareness on the relation of food and health has increased over the last decades [23]. One group of bioactive compounds comprises the glucosinolates (GLs), which occur in high concentrations in plants of the *Brassicaceae* family. The basic chemical structure of GLs is a *β*-thioglucoside *N*-hydroxysulphates with a sulphur linked *β*-D-glucopyranose moiety and a side chain (R), which determines the class of GL (aromatic, indolic and aliphatic) [6]. The biological activity of GLs is ascribed to their breakdown products, which are obtained by the enzymatic conversion of GLs by the endogenous plant enzyme myrosinase (*β*-thioglucosidase, E.C. 3.2.1.147) [8]. Depending on the reaction conditions, several types of breakdown products can be formed, such as isothiocyanates, thiocyanates, epithionitriles and nitriles, from which isothiocyanates are most biological active [14]. Breakdown products of specific GLs contribute to the typical flavour of *Brassica* vegetables and isothiocyanates in particular have been associated with a lower risk of cancer [6].

The majority of *Brassica* vegetables is consumed after domestic cooking, which leads to inactivation of myrosinase and leaves intact GLs [15]. Intact GLs can be hydrolysed by the gut microbiota into breakdown products, but the conversion rate is less than after ingestion of vegetables with intact myrosinase [19, 76]. Several mechanisms of GL losses during food processing have been described: a) enzymatic breakdown by myrosinase upon cell lysis; b) leaching of GLs into the cooking water and c) thermal degradation [23]. The term thermal degradation refers to the degradation solely induced by heat, without any enzymatic reactions. Next to leaching of GLs into the cooking water, thermal degradation is one of the major mechanisms leading to losses of GLs during food processing, since myrosinase is inactivated quickly in most processes [23]. Thermal degradation leads to the formation of nitriles, which are potentially toxic [36] and thus undesired during food processing. To quantitatively study thermal degradation in isolation, the plant material should be treated to inactivate myrosinase and heated in the absence of a separate water phase.

Thermal degradation has been studied in this way in red cabbage and broccoli sprouts, showing that the GL degradation is dependent on the structure of the GL side chain

[31, 32]. The degradation kinetics were described by a first order model [32]. A study comparing the thermal degradation of GLs for five different vegetables showed that the stability of structurally identical GLs differed in different vegetables (broccoli, Brussels sprouts, red cabbage, Chinese cabbage and pak choi). The five studied vegetables have different matrices and hence provide different reaction environments for thermal degradation [35].

In order to get more insight into the causes of differences in thermal degradation, well-characterised plant material is required to exclude differences in degradation due to different growing locations, seasons, age of the plants, storage time and storage conditions. All the mentioned factors possibly influence the food matrix and hence the thermal degradation. It has been shown that the initial GL concentrations in vegetables are affected by the season [77]. Consequently, the question arises if the thermal degradation is affected by the season as well.

The aims of this investigation were to study the effect of the genotype and season on thermal GL degradation in well-characterised plant material and to develop a rapid method to quantitatively determine the thermal degradation rate constants. This information will provide the methodology for future experiments to study in detail the genetic and environmental effects on thermal GL degradation in a more efficient way. Knowledge of the genetic factors influencing GL degradation will enable plant breeders to develop vegetables with low GL degradation, resulting in higher GL concentrations at the point of consumption to increase health benefits for the consumer.

3.2 Materials and methods

3.2.1 Plant material

A rapid-cycling Chinese kale doubled haploid (DH) line, *B. oleracea* var. *alboglabra* (A12DHd), and a broccoli DH line, *B. oleracea* var. *italica* (GDDH33), are the parental lines of a DH population developed by Bohuon *et al.* [71]. For the sake of convenience the genotype A12DHd will be referred to as Chinese kale and the genotype GDDH33 will be referred to as broccoli. Five plants of each genotype were grown in autumn 2010 and in spring 2011. Seeds were sown into soil, transplanted into 19 cm diameter pots after two weeks and grown randomised for six weeks after transplanting in an air-conditioned greenhouse in Wageningen (The Netherlands). Fertilizer was given two to three times per week (electric conductivity 2.1). In autumn 2010 (middle of October until middle of

December) temperatures ranged from 17°C to 22°C (night/ day) and a photoperiod of 16 h was applied. In spring 2011 (end of March till end of May), temperatures were set to 17°C/ 22°C (night/ day); however, on sunny and warm days the temperature could rise to 30°C during the afternoon, because greenhouses are not cooled. Artificial light was applied if the natural photoperiod was shorter than 16 h.

Six weeks after transplanting, all the leaves without petioles from five plants per genotype were harvested in the morning and transported on ice to the laboratory for further sample preparation. All leaves of the five plants per genotype were mixed to prepare one homogenous sample. Previous studies showed [32, 78] that a microwave treatment at high power for a short time was effective in inactivating the enzyme myrosinase while retaining a high level of GLs. This allowed studying the thermal degradation as the sole mechanism of GL degradation. Leaves were cut into pieces of about 3 x 3 cm, 75 g were collected into a plastic beaker and held on ice until microwave treatment. In total, five plastic beakers, each containing 75 g of leaves, were placed at the same time in a microwave at 900 W for 6 min. Samples were immediately cooled on ice, weight loss was recorded and the samples were subsequently freeze dried. Dried samples were weighed to record the water loss, ground into a fine power and stored at -20 °C until further treatment.

3.2.2 Chemicals

Solvents used for extraction and chromatography were of HPLC grade and bought from Biosolve (Valkenswaard, The Netherlands). The DEAE Sephadex-A25 and sinigrin (prop-2-enyl-GL) were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands). The internal standard glucotropaeolin (benzyl-GL) was purchased from the Laboratory of Biochemistry, Plant Breeding and Acclimatization Institute (Radzikow, Błonie, Poland). Sulphatase from *Helix pomatia* (Sigma-Aldrich, Zwijndrecht, The Netherlands) was used for the experiments.

3.2.3 Thermal degradation

Dried plant powder was weighed into 2 mL plastic tubes with screw caps and reconstituted to the water content present in the plant before microwave treatment and drying. Varying amounts of dried plant powder were weighed into tubes and 180 μL of water were added to each tube and mixed. Tubes with reconstituted samples were heated in a heating block at 100°C for 0, 15, 30, 45, 60, 75, 90, 105 and 120 min. All

heat treatments were performed in duplicate. Temperatures were monitored inside the sample with a thermocouple, which was placed through the cap of the tube. The average time to reach 100°C in the samples amounted to 4 min. Time recording started when the tubes were placed into the heating block. After heating, samples were cooled on ice and stored at -20°C till GL analysis.

3.2.4 Glucosinolate analysis

Samples were prepared by heating for nine different time points (in duplicate) per genotype and season. Each sample was prepared once for GL analysis. The GLs were analysed as desulpho-GLs in microtiter plates (chapter 2, [79]), which is based on ISO 9167-1:1992. The reconstituted and heated samples (about 200 mg) were extracted with 1 mL hot methanol (70%, 75°C) for 10 min and centrifuged for 5 min at room temperature at 16 000g. Glucotropaeolin was added as an internal standard (3 mM, 25 µL) during the first extraction. The pellet was re-extracted in the same way and the combined supernatants were used for desulphation. DEAE Sephadex-A25 columns were built in 96-well microtiter-filter plates (Millipore, Amsterdam, The Netherlands). In order to pack the columns, to remove rinsing liquids, to load and to elute the samples, the microtiter plates were centrifuged for 5 min at 900 g. The GL extracts were loaded in two aliquots of 1 mL on the ion-exchange columns and subsequently rinsed twice with 500 µL sodium acetate buffer (20 mM, pH=4.0). For desulphation of GLs, 100 µL freshly purified sulphatase solution was added to the ion-exchange columns and incubated for 16 h at 25°C. Desulpho-GLs were eluted after incubation in microtiter plates by adding 100 µL water twice.

Desulpho-GLs were analysed by high-performance liquid chromatography (HPLC) as described by Verkerk *et al.* [67]. Identification of GLs was based on retention times compared to standard GLs (sinigrin, glucotropaeolin), GLs present in the European Reference Material Rapeseeds (Sigma-Aldrich, Zwijndrecht, The Netherlands), GLs present in reference vegetables and UV-spectra of GLs [72]. GLs were quantitated against the internal standard glucotropaeolin using relative response factors as given in ISO 9167-1:1992, recalculated to glucotropaeolin. The initial GLs concentrations prior to heating were determined after the myrosinase inactivation step.

3.2.5 Modelling

Modelling of GL degradation has been performed using all the data for each genotype and season separately. The molecular mechanism of thermal GL degradation is not known. The general rate law for a single reactant at a concentration c is described in Equation 3.1, with the rate of the reaction *r*, the concentration *c*, the reaction rate constant *k* and the order of the reaction, *n* [80]. The order of a reaction describes the dependence of the reaction rate on the reactant concentration. In contrast to elementary reactions *n* can be fractional and does not necessarily correspond to the stoichiometry of the reaction [80].

Equation 3.1: $$r = -\frac{dc}{dt} = kc^n$$

The integral with respect to time of the general rate law allows to follow the GL degradation over time (Equation 3.2 and Equation 3.3) [80] with c the GL concentration at a certain time point, c_0 the GL concentration at time zero (prior to heating), k_d the degradation rate constant, the time *t* and *n* the order of the reaction.

Equation 3.2: $$c = [c_0^{(1-n)} + (n-1) \cdot k_d \cdot t]^{\frac{1}{1-n}} \quad \text{for } n \neq 1$$

Equation 3.3: $$c = c_0 \cdot e^{-k_d \cdot t} \quad \text{for } n = 1$$

Kinetic parameters (k_d, c_0 for the first order reaction and k_d, c_0, *n* for the nth order reaction) were obtained by fitting the model through the individual data points by minimizing the sum of squared residuals [75] using the solver function in Microsoft Excel. The macro Solver Aid was used to obtain the standard deviation of the parameters and the linear correlation coefficients of the parameters [81, 82]. The corrected Akaike information criterion (AIC_c) was used to discriminate between the first and the nth order model (Equation 3.4) [80], with *m* the number of data points, SS_r the sum of squared residuals and *p* the number of parameters. For model discrimination only the differences between the AIC_c for the different models is relevant (Equation 3.5) [80]. The probability P_{AIC} to estimate which model is more likely was calculated using Equation 3.6. If P_{AIC} is smaller than 0.5 the first order model is more likely, whereas if P_{AIC} is greater than 0.5 the nth order model is more likely [80].

Equation 3.4: $$AIC_c = m\ln\left(\frac{SS_r}{m}\right) + 2(p+1) + 2(p+1)\left(\frac{p+2}{m-p}\right)$$

Equation 3.5: $$\Delta AIC_c = AIC_{c,\,n^{th}} - AIC_{c,1^{st}}$$

Equation 3.6: $$P_{AIC} = \frac{e^{(-0.5\Delta_{AIC})}}{1 + e^{(-0.5\Delta_{AIC})}}$$

3.2.6 Statistics

Means of the initial GL concentration were compared using one-way Analysis of Variance (significance level $p < 0.05$), followed by a Tukey's Honestly Significant Different test (HSD) to test for differences between individual groups using the software IBM SPSS statistics 19. To test for significances of the k_d values between two genotypes and the two seasons, t-tests have been performed. The standard deviations of the k_d values were obtained by the macro SolverAid and the degrees of freedom were calculated from the number of data points reduced by the number of parameters. Furthermore, a paired t-test was applied to test for significances of the k_d values of all GLs between both seasons (Microsoft Excel 2010).

3.3 Results and Discussion

3.3.1 Variation of the initial glucosinolate concentration in different seasons

Nine different GLs were identified in leaves of Chinese kale and broccoli used in this study (Figure 3.1). The broccoli leaves contained glucoraphanin (4-methylsulfinylbutyl-GL, **1**), glucobrassicin (indol-3-ylmethyl-GL, **6**), 4-hydroxy-glucobrassicin (4-hydroxyindol-3-ylmethyl-GL, **7**), 4-methoxy-glucobrassicin (4-methoxyindol-3-ylmethyl-GL, **8**) and neoglucobrassicin (*N*-methoxyindol-3 ylmethyl-GL, **9**). Glucoiberin (3-methylsulfinylpropyl-GL, **2**) was detected in broccoli in spring 2011 only. In Chinese kale leaves contained the same GLs and additionally sinigrin (prop-2-enyl-GL, **3**), gluconapin (but-3-enyl-GL, **4**) and progoitrin ((2R)-2-hydroxybut-3-enyl-GL, **5**) (concentrations shown in Table 3.1). The total GL concentration, calculated as the sum of the individual GLs, differs significantly between the broccoli and Chinese kale in the same season and also between the two seasons for each genotype. In broccoli the total GL concentration was increased in spring 2011 compared to autumn 2010 , which is in accordance with results from Charron *et al.* [37], who found increased GL concentrations in spring compared to autumn season. The total GL concentration decreased in Chinese kale in spring 2011 compared to autumn 2010 that may be explained by the different developmental stage in the two seasons. The Chinese kale plants were already bolting at the time of harvest in spring 2011 and not in autumn 2010, and GL concentrations are changing during plant development [83].
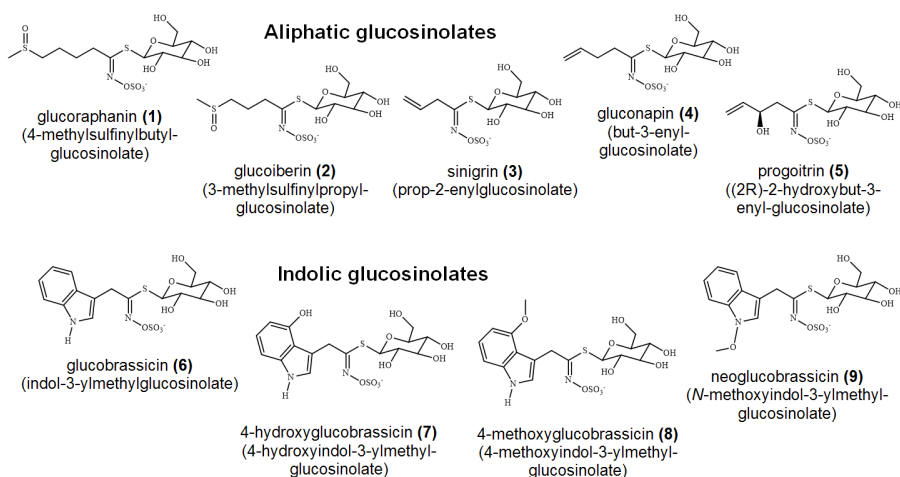

Figure 3.1: Chemical structures and trivial names of the glucosinolates identified in leaves of a broccoli and a Chinese kale genotype used in this study.

Table 3.1: Total and individual glucosinolate concentrations in broccoli and Chinese kale in two different seasons. Different letters indicate significant differences between the different genotypes and seasons for each glucosinolate (p < 0.05, Tukey's HSD test and t-test if the glucosinolate was present in Chinese kale only).

Glucosinolate	Broccoli average ± standard deviation [µmol/100 g FW]		Chinese kale average ± standard deviation [µmol/100 g FW]	
	autumn 2010	spring 2011	autumn 2010	spring 2011
Glucoiberin		13.13 ± 0.70 [a]	2.67 ± 0.03 [b]	2.99 ± 0.05 [b]
Progoitrin			1.19 ± 0.00 [a]	0.35 ± 0.05 [b]
Sinigrin			11.24 ± 0.19 [a]	7.64 ± 0.10 [b]
Glucoraphanin	15.89 ± 0.51 [a]	22.22 ± 1.12 [b]	4.07 ± 0.05 [c]	6.66 ± 0.54 [c]
Gluconapin			27.90 ± 0.64 [a]	23.63 ± 0.08 [b]
4-hydroxy-glucobrassiscin	0.46 ± 0.01 [a]	1.97 ± 0.08 [b]	0.32 ± 0.07 [a]	0.42 ± 0.01 [a]
Glucobrassicin	22.20 ± 1.05 [a]	10.50 ± 0.37 [b]	7.07 ±0.09 [c]	4.63 ± 0.08 [d]
4-methoxy-glucobrassicin	1.65 ± 0.03 [a]	2.94 ± 0.05 [b]	0.72 ± 0.01 [c]	1.16 ± 0.00 [d]
Neo-glucobrassicin	4.44 ± 0.04 [a]	2.57 ± 0.19 [b]	0.53 ± 0.02 [c]	0.39 ± 0.03 [c]
Total glucosinolates	44.65 ± 1.61 [a]	53.33 ± 2.50 [b]	55.72 ± 1.05 [b]	47.88 ± 0.58 [a]

3.3.2 Model discrimination for glucosinolate degradation

Modelling of GL degradation provides a good tool to quantitate the overall degradation progression over time. The obtained parameters facilitate the comparison of data to study genetic and environmental effects of GL degradation. Parameters can be used by the food industry to predict GL concentrations at other heating time points. The aim of the study is to compare thermal degradation of structurally identical GLs in leaves of a broccoli and a Chinese kale genotype grown in autumn and spring season. Glucoraphanin (**1**) and glucobrassicin (**6**), the most abundant GLs in both genotypes, were used to determine the best fitting model. For model discrimination experimental GL concentrations were fitted to the nth and first order model (equation 2 & 3). Table 3.2 presents the parameters of most interest for both models and both GLs. The reaction order n was estimated close to one for glucoraphanin in broccoli and Chinese kale in spring 2011 and for glucobrassicin in Chinese kale in both seasons, which is in accordance with previous results [32, 35]. However the order deviated from one for the other GL-season combinations. The AIC_c confirms these results (Table 3.2). The graphs in Figure 3.2 demonstrate the difference between the models for glucoraphanin.

Table 3.2: Estimated parameters by fitting different models through the measured glucosinolate concentrations after various heating times (0-120 min). B = broccoli, CK = Chinese kale. k_d = rate constant, n = reaction order, SD = standard deviation, AIC_c = corrected Akaike information criterion, $\Delta AIC = AIC\ n^{nt} - AIC_{1st}$ and P_{AIC} = the probability of the AIC ($P_{AIC} < 0.5$, first order model is more likely, $P_{AIC} > 0.5$ n^{nt} order model is more likely).

Genotype	Season	n	SD (n)	$k_d \cdot 10^{-2}$ [μmol $^{(n-1)}$ min^{-1}]	SD (k_d) [%]	$k_d \cdot 10^{-2}$ [min^{-1}]	SD (k_d) [%]	Δ_{AIC}	P_{AIC}
				n^{th} order model		1st order model			
				Glucoraphanin					
B	A 2010	1.70	0.24	8.4	59.6	1.4	6.7	-7.2	0.97 [1]
CK	A 2010	1.48	0.18	9.5	67.6	1.5	5.0	-6.1	0.96 [1]
B	S 2011	1.09	0.11	3.2	24.4	2.6	1.4	2.2	0.25 [2]
CK	S 2011	1.06	0.11	2.9	37.7	2.4	3.9	2.8	0.20 [2]
				Glucobrassicin					
B	A 2010	1.18	0.07	4.5	16.6	3.0	3.5	-4.1	0.89 [2]
CK	A 2010	0.99	0.07	3.0	23.0	3.1	2.8	2.9	0.19 [1]
B	S 2011	0.87	0.06	2.5	19.5	3.7	3.1	-1.3	0.66 [2]
CK	S 2011	0.88	0.08	2.4	32.3	3.8	3.8	1.1	0.36 [1]

[1] n^{nt} order is more likely, [2] first order is more likely

Comparing the resulting degradation rate constants k_d of the first and the nth order model (Table 3.2), shows that the k_d values differ considerably between the two models. This is as expected, when the order differs the concentration term in the rate equation 1 is changing and therefore very different rate constants are estimated at the same reaction rates. To compare the k_d values among different plants and seasons, one common order for all the plants has to be chosen. The introduction of one more parameter to a model increases the uncertainty of the parameters and they could become less meaningful; hence, the uncertainty is an important factor to discuss in model discrimination. The standard deviations of the k_d values were much higher in the models with the nth order equation, ranging from 17% to 60% of the estimated k_d values. Contrarily, the standard deviations of the k_d values estimated with a first order reaction were all below 7%. Hence, the k_d values estimated with the first order equation were more precise and more meaningful than the k_d values estimated with the nth order equation. This fact shows that the nth order model fitted better, but the parameters were less meaningful for the purpose of this study.
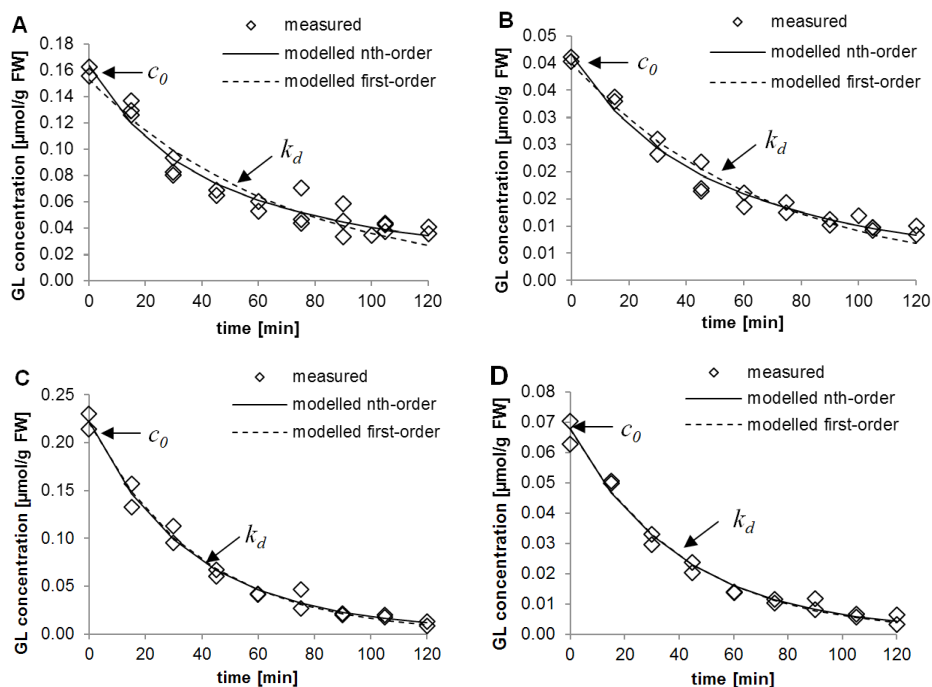
Figure 3.2: First order and nth order fit through the measured data for glucoraphanin in broccoli autumn 2010 (A), Chinese kale autumn 2010 (B), broccoli spring 2011 (C) and Chinese kale spring 2011 (D).

3.3.3 Seasonal and genetic effects on glucosinolate thermal degradation

The k_d values of all GLs at 100°C are presented in Figure 3.3. The GLs 4-hydroxy-glucobrassicin (**7**) and progoitrin (**5**) in spring 2011 in Chinese kale could not be modelled, because the initial concentrations were too low to quantitate their concentration after heating. Figure 3.3 shows clearly higher degradation rate constants for the indolic GLs (**6-9**) compared to the aliphatic GLs (**1-5**), which is in accordance with previous results [32, 35]. The order of k_d values of all GLs was the same in the two genotypes Chinese kale and broccoli grown in the two seasons (from lowest to highest k_d value): glucoiberin < progoitrin < sinigrin < glucoraphanin < gluconapin < neoglucobrassicin < glucobrassicin < 4-methoxy-glucobrassicin. Experiments with five different *Brassica* vegetables showed the same order of degradation at 100°C [35].

Additionally our experiments point out that the GL degradation is significantly higher in spring 2011 compared to autumn 2010 for all GLs in both genotypes (p < 0.01, paired t-test). In spring 2011 the k_d values for all GLs analysed in Chinese kale were

between 20% and 80% higher and for broccoli between 16% and 80% higher than in autumn 2010. The difference between the seasons was lower for indolic GLs than for aliphatic GLs in both genotypes (Figure 3.3). The observed differences in k_d for the same GLs in the two genotypes are generally small and only a few are significantly different. This result is in contrast to previous results that showed that the degradation of chemical identical GLs differed in five different *Brassica* vegetables [35]. Dekker *et al.* [35] suggested differences in vegetable matrix and hence differences in the reaction environment for the thermal degradation as reason for the differences. Other authors showed that the pH influenced thermal degradation of GLs, being lower in neutral and slightly acidic medium, whereas degradation was increased in basic medium [31]. Water, as a mediator of chemical reactions, influences the degradation of GLs as well. It has been shown that thermal GL degradation is increasing with decreasing water content, except when water contents are very low (13%) [84]. In the study of Dekker *et al.* [35], the GL degradation was determined in the edible parts of five different vegetables, red cabbage, Brussels sprouts, broccoli, Chinese cabbage and pak choi, at their commercial maturity. Hence different plant organs were studied at different developmental stages of the plants. Furthermore the vegetables were bought in a supermarket, thus the growing and storage conditions were not known. In this study we can separate the genetic and environmental effects on degradation, since we used leaves of broccoli and Chinese kale grown under two defined conditions (spring 2010, autumn 2011), which were analysed immediately after harvest, without storage. This resulted in the same degradation speed for several GLs in broccoli leaves and Chinese kale leaves. Our data showed also clearly that the season has a big impact on both GL concentration and the GL degradation suggesting that the differences among different vegetables in the study of Dekker *et al.* [35] can be caused by seasonal, environmental effects as well as by genetic effects. Studying the GL degradation in a segregating plant population, like the AGDH doubled haploid population developed from a F1 of a cross between the broccoli and the Chinese kale used in this study [71], could answer the question regarding the effects of environment and genetics on GL degradation. With knowledge of genetic factors on GL degradation, vegetables with a higher retention of GLs during food processing can be bred. With knowledge of the environmental influence on GL degradation, the optimal genotype-environment combination can be selected. This strategy provides a novel opportunity to the food industry to improve the level of health promoting compounds at the stage of consumption.
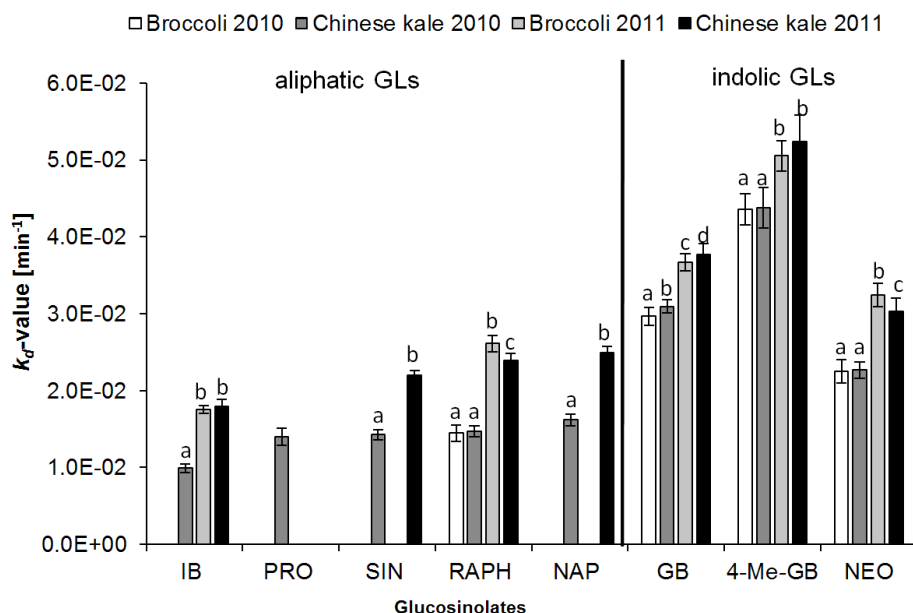
3

Figure 3.3: Rate constants (k_d values) of glucosinolate thermal degradation at 100°C in broccoli and Chinese kale in autumn 2010 and spring 2011. The bars show the modelled k_d values and error bars represent the standard deviations of the parameters. (IB) glucoiberin, (PRO) progoitrin, (SIN) sinigrin, (RAPH) glucoraphanin, (NAP) gluconapin, (GB) glucobrassicin, (4-Me-GB) 4-methoxy-glucobrassicin, (NEO) neoglucobrassicin. Missing bars indicate the absence of the glucosinolate or the levels were too low for modelling. The letters indicate significant differences ($p < 0.05$).

3.3.4 Improving the methodology to estimate rate constants rapidly

In a follow-up study, we want to quantitate the genetic effect of GL thermal degradation, for which a large segregating population is needed, the progeny from the parental genotypes described in this study. These large data (> 100 progeny) require high-throughput GL degradation studies; one step towards this is the use of less heating time points for GL determination. To investigate this possibility, the dataset of the two investigated genotypes grown in both seasons was modelled using the first order reaction (Equation 3.3) in two ways: with all nine different heating times and with a subset of four heating times (0, 15, 30 and 60 min). In total 24 datasets of GLs were modelled. The obtained kinetic parameters (k_d and c_0) estimated with nine heating times were plotted against the parameters estimated with four heating times (Figure 3 4). The slope amounted to 1.08 for the k_d value (Figure 3.4 A) and to 0.98 for the c_0 value (Figure 3.4 B), which shows that almost equal parameters are obtained by modelling with four data points (k_d value: $R^2 = 0.993$, c_0 value: $R^2 = 0.999$). The average difference

between the parameters estimated with nine and four heating times amounted to 2.1% for the k_d value and to 0.8% for the c_0 value, which is lower than the average standard deviations of the estimated parameters with nine heating times (4.7% for k_d and 2.5% for c_0) and hence confirm the results of Figure 3.4. As expected, the standard deviations of the parameters estimated with four heating times were higher than after estimation with nine heating times. An average standard deviation from the 24 models of 6.8% for the k_d value and 2.8% for the c_0 after modelling with four heating time points were considered to be acceptable.

All the data show that four heating times (0, 15, 30 and 60 min) are sufficient to estimate the k_d value, which will be used for determination of thermal GL degradation in a high number of samples in future studies to study the genetic regulation of GL composition and thermal degradation.
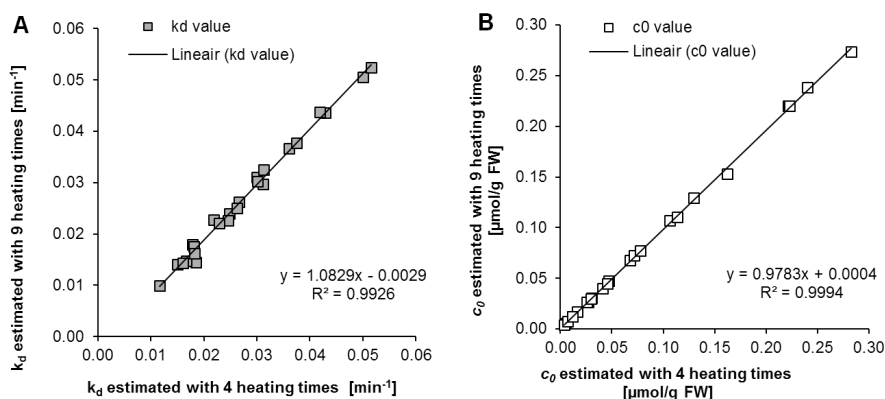
Figure 3.4: Plots of estimated parameters using 9 heating times and 4 heating times for the first order modelling. Data of all glucosinolates present in Broccoli and Chinese kale in two different seasons were modelled. A: degradation speed (k_d values), B: initial glucosinolate concentration (c_0 value).

3.4 Acknowledgements

CHAPTER 4

Quantitative Trait Loci Analysis of Non-enzymatic Glucosinolate Degradation Rates in *Brassica oleracea* During Food Processing

Abstract

Epidemiological and mechanistic studies show health promoting effects of GLs and their breakdown products. In literature differences in non-enzymatic glucosinolate degradation rates during food processing among different vegetabl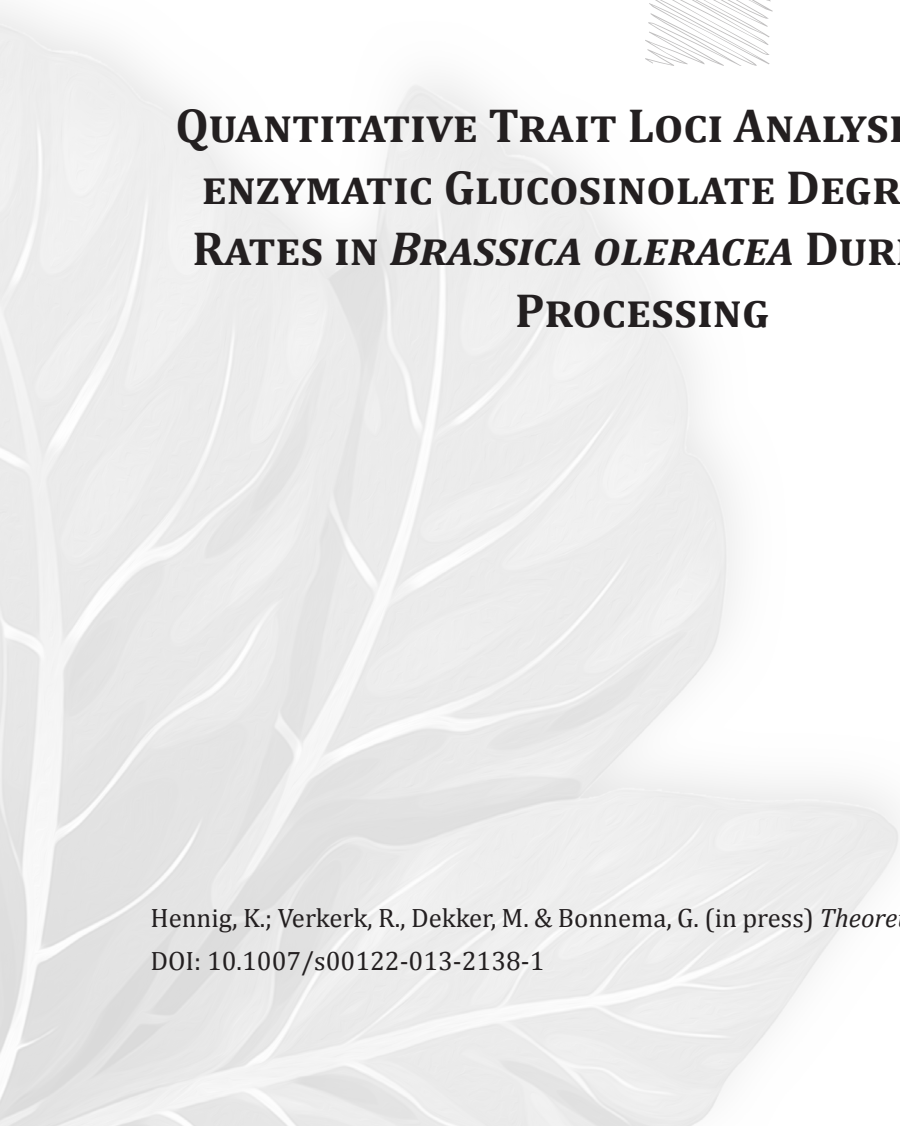es are described, which provides the basis for studying genetic effects of this trait and breeding for vegetables with high glucosinolate retention during food processing. Non-enzymatic glucosinolate degradation, induced by heat, was studied in a publically available *Brassica* oleracea doubled haploid population, data were modelled to obtain degradation rate constants that were used as phenotypic traits to perform quantitative trait loci (QTL) mapping. Glucosinolate degradation rate constants were determined for five aliphatic and two indolic GLs. Degradation rates were independent from the initial glucosinolate concentration. Two QTL were identified for the degradation rate of the indolic glucobrassicin and one QTL for the degradation of the aliphatic glucoraphanin, which co-localises with one of the QTL for glucobrassicin. Factors within the plant matrix might influence the degradation of different GLs in different genotypes. In addition to genotypic effects, we demonstrated that growing conditions influence glucosinolate degradation as well. The study identified QTL for glucosinolate degradation, giving the opportunity to breed for vegetables with a high retention of GLs during food processing, although the underlying mechanisms remain unknown.

Keywords

4.1 Introduction

Glucosinolates (GLs) are an important group of secondary plant metabolites involved in the plant defence system and, from a human point of view, GLs exhibit interesting health promoting properties. In epidemiological studies the intake of GL containing vegetables is related to a reduced incidence of certain cancers. Epidemiological data associate the cancer protective effects of *Brassica* vegetables to GLs; however, no biological activity of intact GLs could be demonstrated in in-vitro studies [15]. Upon cell damage, GLs are hydrolysed by the endogenous enzyme myrosinase (*β*-thioglucosidase, EC 3.2.1.147), which results in enzymatic degradation and the formation of a variety of breakdown products. One type of breakdown products, isothiocyanates, are bioactive by affecting the process of carcinogenesis in several organs, such as lung, stomach, colorectal and breast [15]. In the absence of active myrosinase in the food ingested, GLs can also be converted into breakdown products by the human gut flora, although this conversion is less efficient [19, 76].

The large variation in the content and composition of GLs in *Brassica* vegetables is reviewed extensively by Verkerk *et al.* [6]. Although this variation is caused by several factors such as environmental factors, including soil, climate and fertilisation, the most important factor determining the GL content is genetic variation [6]. Many genes which are involved in the GL pathway have been studied using the natural variation combined with molecular marker techniques and subsequent cloning of the identified quantitative trait loci (QTL). Characterization of QTL for GL biosynthesis has been conducted in Arabidopsis [85] and *Brassicaceae* [86, 87]. Genes involved in the GL pathway are summarized by Sønderby *et al.* [41] and Wang *et al.* [88].

The levels and types of GLs in *Brassica* vegetables show natural genetic variation and have been altered further by breeding, indirectly through the selection of taste, flavour and possibly through the selection for resistance to herbivores [6], and directly by breeding for enhanced levels of GLs expected to be associated with health benefits. Breeding for a higher GLs content in broccoli, in particular glucoraphanin (4-methylsulfinylbutyl-GL), was achieved by introgression of three small segments of the genome of a wild *Brassica* species, *Brassica villosa*, into a broccoli background [45, 47]. Furthermore, it was shown that the high glucoraphanin levels are retained during cooking for short times. As a result of this breeding strategy, a broccoli with a high content of glucoraphanin was launched into the UK market in autumn 2011 under the name *Beneforte* [89].

4

Breeding for higher GL concentrations is one way to increase the intake of GLs. However, from a food science perspective, optimizing food processing methods to retain the GLs is equally important, since enhanced levels in primary production can easily be lost during processing and preparation [23]. Three main mechanisms of GL losses during food processing have been described as: a) enzymatic breakdown by myrosinase, b) leaching of GLs and enzymatic breakdown products into the cooking water and c) thermal degradation of GLs [23]. The term "thermal degradation", as one sub-process of GL degradation, refers to the GL degradation solely induced by heat without myrosinase activity. Myrosinase is a thermolable enzyme, the activity of which is considerably reduced at moderately high temperatures. A treatment for 3 min at 60°C reduced the myrosinase activity by 90% in broccoli [27], a microwave treatment for 120 s with a final temperature of 88°C resulted in almost complete loss of myrosinase activity in cabbage [29] and steaming for 6 min led to almost complete loss of myrosinase activity in broccoli reaching a temperature of 80°C [30]. These studies demonstrate that the temperatures applied during industrial processing and domestic preparation of *Brassica* vegetables quickly inactivate myrosinase, leaving leaching and thermal degradation as the main mechanisms for GL losses in processing and preparation.

A study, comparing the thermal degradation of GLs for five different vegetables, showed that the stability of structurally identical GLs differed in different *Brassica* vegetables (broccoli, Brussels sprouts, red cabbage, Chinese cabbage and pak choi). The five studied vegetables differ in many traits, like metabolic composition, and hence provide different reaction environments for thermal degradation [35]. These differences in metabolic composition are determined by biochemical traits and will be (partly) genetically regulated. Therefore, it is hypothesized that thermal degradation of GLs is (partly) genetically regulated and therefore genomic regions influencing the thermal degradation rates of GLs can be identified.

To test this hypothesis, GL thermal degradation was studied in a *Brassica oleracea* doubled haploid population (AGDH, broccoli x Chinese kale, [71]) using a kinetic modelling approach as described in chapter 3, [90]. The values of the obtained degradation rate constants were used as traits to perform QTL analysis. The described QTL analysis of GL degradation provides novel insights in the role of the genetic background during food processing. Observed QTL could facilitate breeding for vegetables with a higher retention of GLs during processing. Furthermore, the future identification of biochemical traits underlying the observed QTL and subsequently genes influencing the processing

of vegetables could lead to new breeding strategies to breed for vegetables that retain GLs better than others during food processing. The proposed strategy complements breeding strategies for high GL content, in order to retain high GL concentrations throughout the entire food production chain and ensure a high GL concentration at the stage of consumption.

4.2 Materials and methods

4.2.1 Plant material

The doubled haploid (DH) population AGDH, developed by Bohuon *et al.* [71], was used to study GL thermal degradation. This population was developed by crossing two DH parents, a rapid-cycling Chinese kale line, *B. oleracea* var. *alboglabra* (A12DHd), and a calabrese broccoli line, *B. oleracea* var. *italica* (GDDH33), through microspore culture of the F_1. Seeds were sown into soil, plants were transplanted into 19 cm diameter pots after two weeks, randomised and grown for 6 weeks after transplanting in a greenhouse in Wageningen (The Netherlands). A total of 100 DH lines of the population was grown in five replicates in spring 2009 (end of March till end of May) under natural light and temperature conditions. Due to the lack of seeds, the parental lines were not included in the study. Temperatures ranged from 5°C to 16°C during the night and from 13°C to 30°C during the day. Fertiliser was given two to three times per week (electric conductivity 2.1). A subset of ten DH lines of the population, selected based on their GL degradation in 2009, was grown again in spring 2011 (end of March till end of May) in a greenhouse in Wageningen. Temperatures ranged from 17°C to 22°C (night/day); however, on sunny and warm days the temperature could rise to 30°C during the afternoon. Artificial light was applied if the natural photoperiod was shorter than 16 h.

Eight weeks after sowing, leaves without petioles were harvested in the morning and transported on ice to the laboratory for further sample preparation. Harvesting of the 100 DH lines in spring 2009 was performed on four consecutive mornings, but all five biological replicates of the same DH line were harvested at the same day. Harvesting in spring 2011 was performed in one morning. All leaves from the five plants per DH line were pooled to prepare one homogenous sample per DH line. This approach was chosen to prepare an average sample accounting for the variation during growth.

In order to study the GL thermal degradation as a sole mechanism, without enzymatic GL degradation, the myrosinase was inactivated by microwave treatment. Microwave

treatment was performed at high power for a short time, which has shown to fully inactivate myrosinase with high retention of GLs [32, 78]. Leaves were cut into pieces of about 3 x 3 cm, of which 75 g was placed into a plastic beaker and held on ice until microwave treatment. In total, five plastic beakers, each containing 75 g of leaves, were placed at the same time in a microwave at 900 W for 6 min. After the microwave treatment, samples were immediately cooled on ice, weight loss was recorded and the samples were subsequently freeze dried. Dried samples were weighed to record the water loss, ground into a fine power and stored at -20°C until further treatment.

4.2.2 Chemicals

Solvents used for extraction and chromatography were of HPLC grade and bought from Biosolve (Valkenswaard, The Netherlands). The DEAE Sephadex-A25 and sinigrin (prop-2-enyl-GL) were purchased from Sigma-Aldrich (Zwijndrecht, The Netherlands). The internal standard glucotropaeolin (benzyl-GL) was purchased from the Laboratory of Biochemistry, Plant Breeding and Acclimatization Institute at Radzikow, Błonie, Poland. Two different batches of sulphatase from *Helix pomatia* (Sigma-Aldrich, Zwijndrecht, The Netherlands) were used for the experiments (lot nr. 089K3797 and 029K3782).

4.2.3 Thermal degradation

Dried plant powder was weighed into 2 mL plastic tubes with screw caps and reconstituted with MilliQ water to obtain 200-220 mg plant material having the same water content as before the microwave treatment and drying. Tubes with reconstituted samples were heated in a heating block at 100°C for 0, 15, 30 and 60 min. All heating experiments were performed in duplicate. Temperatures were monitored inside the sample with a thermocouple, which was placed through the cap of the tube. The average time to reach 100°C in the samples amounted to 4 min. After the predefined heating times, samples were cooled on ice and stored at -20°C till GL analysis.

4.2.4 Glucosinolate analysis

GLs were analysed as desulpho-GLs by HPLC. The reconstituted and heated samples were extracted with hot methanol (70%) and subsequently desulphated in microtiter plates with a purified sulphatase as described in chapter 2, [79].

4.2.5 Modelling

Mathematical modelling was used as a tool to describe the thermal degradation speed of GLs in a quantitative way. Since the molecular mechanism of GL degradation is not known, the reaction can be described by empirical models. Equation 4.1 displays the integral of the general rate law with respect to time for the case that the order of the reaction equals 1 with c the GL concentration at a certain time point, c_0 the GL concentration at time zero (without heating), k_d the degradation rate constant and the heating time t at a constant temperature. For this first order mechanism GLs are in one limiting reaction step degraded into other products, whereby the rate is proportional to the concentration of the remaining GLs. In several studies the first order equation has been applied to model GL thermal degradation [32, 35, 84] and it has been shown for the parental lines of the DH population used in this study that GL degradation follows a first order reaction independently from the growing season (chapter 3, [90]).

Equation 4.1: $\qquad c = c_o \cdot e^{-k_d \cdot t}$

4.2.6 Statistics

The Shapiro-Wilk normality test was performed using the program R to test for the normal distribution of the data (significance level $p < 0.05$). The Pearson correlation coefficients and the significance of the correlations were determined using the software IBM SPSS Statistics 19. The Bonferoni adjustment was used to correct the significance value for multiple testing by dividing the significance levels $\alpha = 0.05$ and $\alpha = 0.01$ by the number of paired comparisons made. Analysis of Variance (ANOVA) was performed to test for an influence of the sampling on the degradation rate constants (k_d) using the software IBM SPSS Statistics 19. No significant influence of the harvesting day on the k_d values was found for glucobrassicin, sinigrin, progoitrin, gluconapin and neoglucobrassicin. Small differences were detected for glucoiberin and glucoraphanin between harvesting day 1 and day 3.

4.2.7 QTL mapping

The linkage map used for QTL analysis has been constructed with a subset of 107 DH lines of the AGDH population. The map was constructed based on restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP) and microsatellite markers published by Sebastian *et al.* [91]. The final map was constructed with the subset of 107 DH lines using the software JoinMap 4.0. The genetic map constructed with 210 DH lines of the population is publically available on www.brassicainfo.com. The total length of the two maps are comparable, the length of the individual linkage groups differ, due to a different number of DH lines used for the map construction in this study.

QTL mapping was performed using the Software MapQTL 6.0 [92] using the GL thermal degradation rate constants (k_d values) and the initial GL concentrations (c_0 values) as phenotypic traits. The analysis started with interval mapping (IM) and a permutation test was applied to each dataset (1000 repetitions) to decide the LOD (logarithm of odds) threshold ($p \leq 0.05$). In this study a genome-wide LOD value of 2.7 was used as significant threshold for the c_0 value and the k_d value of the GLs glucobrassicin (indol-3-ylmethyl-GL) and glucoraphanin (4-methylsulfinyl-GL); a genome-wide LOD of 2.8 was used as the significant threshold for the c_0 value of glucoiberin (3-methylsulfinylpropyl-GL) and 2.6 for the c_0 value of gluconapin (but-3-enyl-GL). The results of the interval mapping were used to identify putative QTLs, which were selected as cofactors if their LOD value exceeded the chromosome-wide threshold. Cofactors were verified by using the automatic cofactor selection tool and only significant markers ($p < 0.02$) were used as cofactors for the subsequent restricted multiple-QTL model mapping (RMQM). A map interval of 5 cM was used for both IM and RMQM analyses.

The data were analysed for epistatic interaction by regression analysis using the software IBM SPSS Statistics 19. The peak markers of each QTL which had low missing values were used to estimate the contributions of individual markers and their interaction to the phenotypic variance. The individual effects and the interactions were fitted into the regression model for the interaction of two QTL (Equation 4.2) and a p-value < 0.05 was applied to determine the significance of the interaction.

Equation 4.2: $\quad y = \beta_0 + \beta_1 Q_1 + \beta_2 Q_2 + \beta_k Q_1 \cdot Q_2$

4.3 Results

4.3.1 Initial glucosinolate concentrations in the Doubled Haploid population

The GL profile and concentrations varied substantially throughout the DH population and only few GLs could be identified in all DH lines. The most abundant GLs identified in all of the 100 DH lines of the DH population were the aliphatic GL glucoraphanin (4-methylsulfinylbutyl-GL) and the indolic GL glucobrassicin (indol-3-ylmethyl-GL). Table 4.1 gives an overview of the range of initial GL concentrations (c_0 values) determined in the DH population and shows the structures of the GLs including the chemical and trivial names. Other aliphatic GLs present in most of the DH lines are glucoiberin (3-methylsulfinylpropyl-GL), progoitrin ((2R)-2-hydroxybut-3-enyl-GL), sinigrin (prop-2-enyl-GL) and gluconapin (but-3-enyl-GL). Furthermore, the indolic GL neoglucobrassicin (*N*-methoxyindol-3-ylmethyl-GL) could be identified in 91 DH lines with the lowest GL concentrations. The histogram of the initial GL concentrations (c_0 values) of glucoraphanin and glucobrassicin are presented Figure 4.1 a & b. The Shapiro test for normal distribution shows that both datasets are not normally distributed (p < 0.05). QTL mapping was conducted for the initial GL concentrations (c_0 values) and the results are presented in Figure 4.2 a-d. Two QTL were identified for the aliphatic GL glucoraphanin, one on linkage group C09 (30% explained variation) and on one linkage group C01 (9.4% explained variation). The QTL for glucoiberin co-localises with one of the QTL for glucoraphanin on linkage group C09 (28% explained variation), both GLs are methylsulfinyl-GL differing by only one C-atom in their side chain. On linkage group C07 a small QTL for another aliphatic GL, gluconapin, is located explaining 16% of the variation. Two QTL for the initial GL concentration of the indolic GL glucobrassicin are located on linkage group C03 (15% and 12% explained variation).
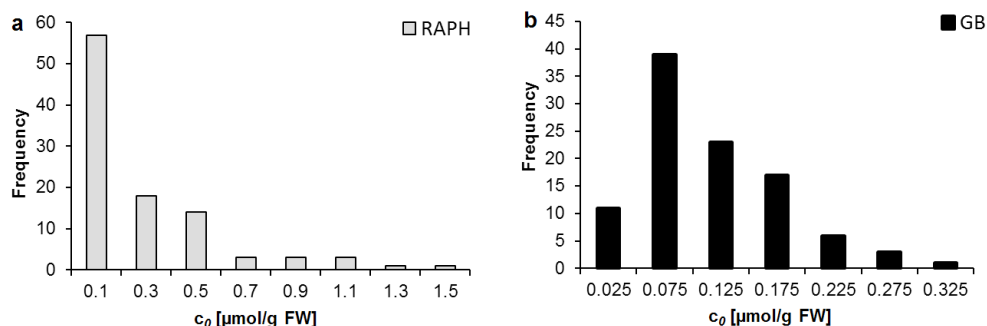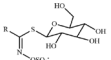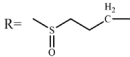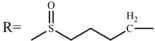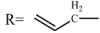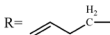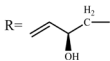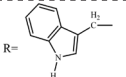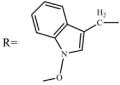

Figure 4.1: Histogram of the initial glucoraphanin (a, RAPH) and glucobrassicin (b, GB) concentration (c_0 value) in the DH population in spring 2009, bars represent the observed frequencies

Table 4.1: Chemical structure of identified glucosinolates and range of modelled initial glucosinolate concentrations (c_0 values) in the DH population in spring 2009.

General glucosinolate structure				
Glucosinolate	**Structure side-chain**	**DH lines with detected glucosinolate (N)**	**Concentration range [µmol/g FW]**	**Average range [µmol/g FW]**
Glucoiberin [a] 3-Methylsulfinylpropyl-	R=	65	0.011 – 0.321	0.096
Glucoraphanin [a] 4-Methylsulfinylbutyl-	R=	100	0.006 – 1.568	0.277
Sinigrin [a] Prop-2-enyl-	R=	40	0.017 – 0.675	0.169
Gluconapin [a] But-3-enyl-	R=	60	0.028 – 1.180	0.306
Progoitrin [a] (2R)-2-Hydroxybut-3-enyl-	R=	45	0.010 – 1.134	0.360
Glucobrassicin [i] Indol-3-ylmethyl-	R=	100	0.008 – 0.328	0.114
Neoglucobrassicin [i] N-methoxyindol-3-ylmethyl-	R=	91	0.004 – 0.088	0.020

[a] Classified as aliphatic glucosinolate according to the structure of the side chain,
[i] Classified as indolic glucosinolate according to the structure of the side chain

Figure 4.2: QTLs for the initial glucosinolate concentrations (c_0 values) of glucobrassicin, glucoraphanin, glucoiberin and gluconapin in spring 2009. The black horizontal line represents the genome-wide LOD value, determined with the permutation test. ⸺ glucoraphanin; ...... gluconapin, ‑ ‑ ‑ glucoiberin ⸺ glucobrassicin, ‑ ·· LOD
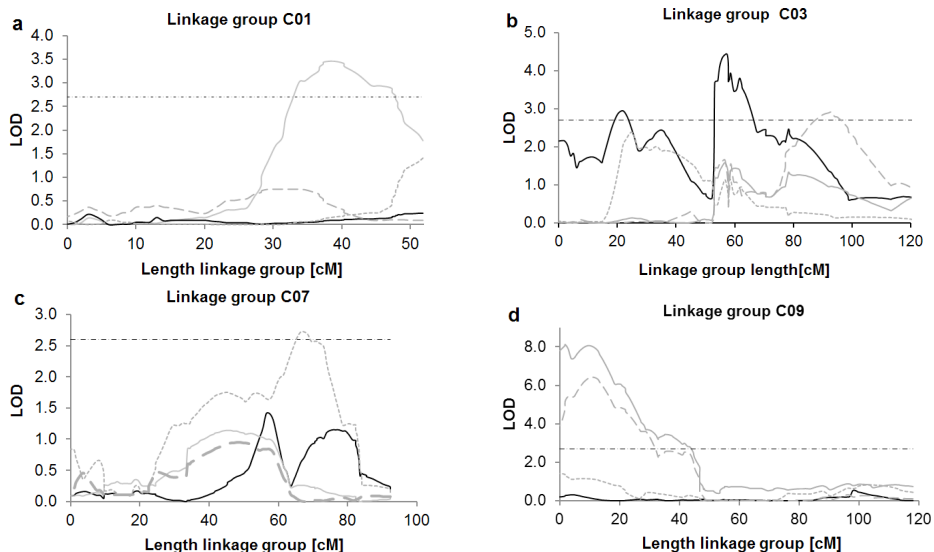
4.3.2 Glucosinolate thermal degradation rate constants

The value of the thermal degradation rate constant k_d describes the steepness of the degradation curve and is hence a quantitative measure for the rate of the GL degradation (chapter 3, [90]). The distribution of GL degradation rate constants (k_d values) of the two most abundant GLs in the DH population is presented in Figure 4.3. The average k_d value of glucobrassicin amounted to $3.21 \bullet 10^{-2}$ min^{-1} and the average k_d value of glucoraphanin amounted to $1.85 \bullet 10^{-2}$ min^{-1}, hence the degradation of glucobrassicin is almost twice as fast as the degradation of glucoraphanin. For both GLs the k_d values show a normal distribution throughout the DH population (Shapiro-Wilk normality test, $p > 0.05$). Table 4.2 gives an overview of the variation of the k_d values for all GLs identified in the DH population. The highest variation was observed for glucoraphanin (8.7-fold difference), whereas the variation of the k_d values for the other GLs was lower (around 3-fold). The average k_d values are lower for the GLs with an aliphatic side chain (glucoiberin, sinigrin, progoitrin, glucoraphanin, gluconapin) compared to the GLs with an indolic side chain (glucobrassicin, neoglucobrassicin). The data show that GL degradation is dependent on two factors: a) the structure of the GL side chain and b) on the plant matrix, as illustrated by the distribution of the GL degradation throughout the DH population. Furthermore, the low correlation coefficients between the k_d and the c_0 values (Table 4.2) show that the GL degradation is not dependent on the initial GL concentration.

Table 4.2: Variation of the glucosinolate thermal degradation (k_d values) for all the GLs identified in the DH population and the correlation with the initial glucosinolate concentrations (c_0 values) in spring 2009.

Glucosinolate (present in nr. of DH lines)	Mean k_d value $k_d \cdot 10^{-2}$ [min^{-1}]	Range k_d $k_d \cdot 10^{-2}$ [min^{-1}]	Max. difference (-fold)	Correlation k_d - c_0 values [ns]
Glucoiberin [a] (65)	1.26	0.66 – 2.04	3.1	0.038
Progoitrin [a] (45)	1.44	0.63 – 2.46	3.9	-0.174
Sinigrin [a] (40)	1.65	1.00 – 2.30	2.3	-0.191
Glucoraphanin [a] (100)	1.85	0.34 – 3.00	8.7	-0.044
Gluconapin [a] (60)	1.91	0.80 – 2.74	3.4	-0.117
Glucobrassicin [i] (100)	3.21	1.55 – 4.63	3.0	-0.027
Neoglucobrassicin [i] (91)	2.52	1.64 – 4.21	2.9	0.193

[ns] None of the correlations was significant ($p > 0.05$)

[a] Classified as aliphatic GL according to the structure of the side chain,

[i] Classified as indolic GL according to the structure of the side chain
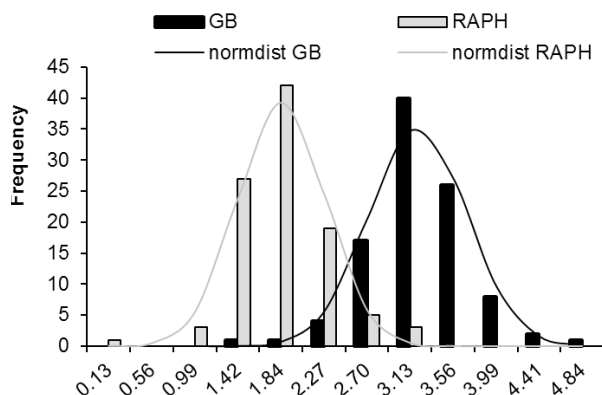
Figure 4.3:

Histogram of thermal degradation rate constant (k_d value) of glucoraphanin (RAPH) and glucobrassicin (GB) in the DH population in spring 2009, bars represent the observed frequencies and the lines represent the normal distribution curve.

4.3.3 Correlation of thermal degradation of different glucosinolates

As shown in Table 4.1, the GL concentrations varied throughout the DH population and not all GLs could be identified in all the DH lines. In order to determine if the degradation of different GLs in the same DH lines are correlated, the k_d values of two different GLs present in the same DH lines were plotted against each other. As an example, the plot of the k_d values of the aliphatic GL glucoraphanin against the k_d values of the indolic GL glucobrassicin is shown in Figure 4.4, which shows a positive linear relationship. To estimate the strength of the relationship between the k_d values of all different GLs, the Pearson correlation coefficients and significances were calculated (Table 4.3). For all the possible combinations of the different GLs positive correlations were found. All was observed were strong ($r > 0.5$); except for neoglucobrassicin ($r < 0.5$) with glucoiberin and sinigrin, but all the correlations are significant ($p < 0.05$) and most of the correlations are highly significant ($p < 0.01$). The strongest correlations ($r > 0.9$) were observed for the aliphatic GLs gluconapin with sinigrin and gluconapin with progoitrin. Gluconapin and sinigrin are both alkenyl-GLs, which differ in their side chain only by one CH_2-group. The side chain of progoitrin consists also of an alkenyl-group with the only modification compared to gluconapin being a hydroxyl-group at the second C-atom. The k_d values of glucoraphanin and glucoiberin are also highly correlated ($r = 0.886$), which are both sulfinyl-GLs with one CH_2-group difference in their side chains. The correlations for GLs with structurally different side chains are less strong, for example the aliphatic GL glucoraphanin versus the indolic GL glucobrassicin ($r = 0.771$). The linear correlation of the k_d values of different GLs in the same DH lines indicates that similar factors influence the thermal degradation of structurally different GLs. The influencing factors for the degradation of aliphatic GLs may differ from the influencing factors for the degradation of indolic GLs, since the correlation between k_d values of aliphatic and indolic GLs is less strong.
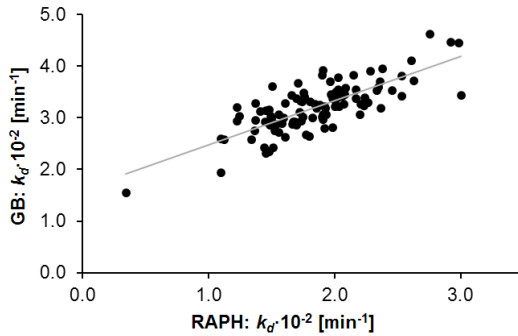
Figure 4.4:

Plot of the degradation rate constants (k_d values) of glucoraphanin (RAPH) against the k_d values of glucobrassicin (GB) in the DH population in spring 2009

Table 4.3: Correlation coefficients of the k_d values for glucosinolate thermal degradation of the DH population in spring 2009 for the different glucosinolates, when they were present in the same DH lines (Abbreviations: aliphatic GLs: glucoiberin (IB), progoitrin (PRO), sinigrin (SIN), glucoraphanin (RAPH), gluconapin (NAP); indolic GLs: glucobrassicin (GB), neoglucobrassicin (NEO)).

k_d 2009	IB	PRO	SIN	RAPH	NAP	GB	NEO
IB	1	0.577*	0.586**	0.886**	0.514*	0.658**	0.410*
PRO	0.577*	1	0.871**	0.799**	0.928**	0.604**	0.557**
SIN	0.586**	0.871**	1	0.734**	0.940**	0.677**	0.489*
RAPH	0.886**	0.799**	0.734**	1	0.832**	0.771**	0.543**
NAP	0.514*	0.928**	0.940**	0.832**	1	0.750**	0.530**
GB	0.658**	0.604**	0.677**	0.771**	0.750**	1	0.584**
NEO	0.410**	0.557**	0.489*	0.543**	0.530**	0.584**	1

** Correlation is significant at the level 0.01 (2-tailed, corrected for multiple testing: $\alpha < 0.00048$);

* Correlation is significant at the level 0.05 (2-tailed, corrected for multiple testing: $\alpha < 0.00238$)

4.3.4 Quantitative trait loci analysis of glucosinolate thermal degradation

To test the hypothesis if the variation in GL degradation can be explained genetically, QTL mapping was performed for GL thermal degradation using the degradation rate constants (k_d values) obtained from modelling the measured GL concentrations after several heating times using the first order equation (Equation 4.1). Two significant QTLs could be identified for GL degradation (Figure 4.5). The major QTL was identified on linkage group C09, which explains 13.6% of phenotypic variation for glucobrassicin (LOD = 3.7) and 12.2% of the phenotypic variation for glucoraphanin (LOD = 3.5). Furthermore, for the degradation of glucobrassicin a second QTL (LOD = 3.24) was identified on linkage group C07, explaining another 11.4% of the phenotypic variation. QTL mapping for the GLs, which were present in less than 100 DH lines, did not result in

any significant results, probably due to the lower number of data. The co-localization of one QTL for the degradation of glucobrassicin and glucoraphanin suggests that there are common genetic factors involved in the degradation of the two GLs. No epistatic effects were found for the two QTLs identified for the degradation of glucobrassicin using regression analysis (p > 0.05). As shown in Table 4.2, the GL degradation rate constants (k_d values) are independent from the initial GL concentration. This result was confirmed by QTL mapping of the initial GL concentrations of glucoraphanin and glucobrassicin, which did not co-localise with the QTLs for GL degradation (Figure 4.2).
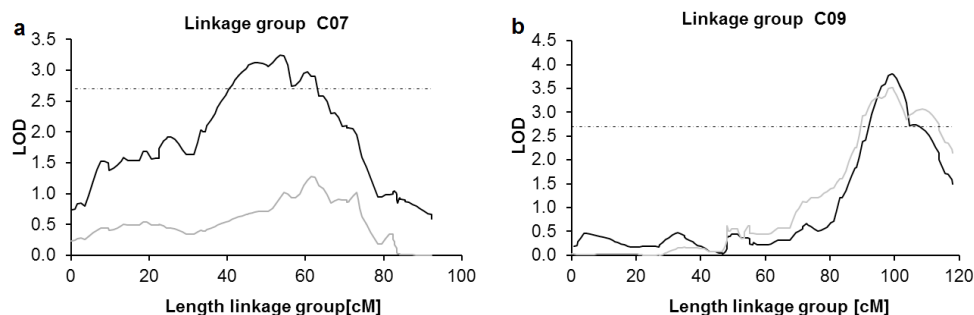
Figure 4.5: QTLs for glucosinolate thermal degradation of glucobrassicin and glucoraphanin in spring 2009, the black horizontal line represents the genome-wide LOD value, determined with the permutation test. ⎯⎯ glucoraphanin; ⎯⎯ glucobrassicin, ⎯··⎯ LOD

4.3.5 Influence of the season on the glucosinolate thermal degradation

To test if the GL degradation is reproducible, ten DH lines selected to cover the observed range of GL degradation rate constants for glucobrassicin and glucoraphanin in spring 2009 were grown in 2011 in the same season and treated in the same way as in spring 2009. The obtained first order degradation rate constants (k_d values) were compared (Figure 4.6). The k_d values for glucoraphanin were between 23% lower and 120% higher in spring 2011 compared to spring 2009 (Figure 4.6a) and the k_d values for glucobrassicin were between 22% lower and 72% higher in spring 2011 compared to spring 2009 (Figure 4.6b). The observed range of k_d values was smaller in spring 2011 compared to spring 2009 for glucobrassicin and glucoraphanin. Furthermore, Figure 4.6 illustrates that in most of the DH lines a change in the k_d value in 2011 compared to 2009 for glucobrassicin was associated with a change in the same direction in the k_d value for glucoraphanin. For that reason the Pearson correlation coefficients for the k_d values of different GLs in spring 2011 were calculated (Table 4.4). The k_d values of the different GLs are all strongly correlated (r > 0.5), although only some correlations are

significant. A strong correlation of the degradation of different GLs in the same DH lines was also found in spring 2009. However, the k_d values of the same GLs in spring 2009 and spring 2011 are not correlated (r < 0.5, not significant, Table 4.5). The correlation of the k_d values of different GLs within each season indicates again, that there are common factors in the DH lines influencing the GL thermal degradation which are affected by the environment. The initial GLs concentrations were found to be independent from the k_d values in spring 2011 for most of the GLs (correlations not significant), similar as it was shown for the DH population in spring 2009. In contrast to the k_d values, the initial GL concentrations (c_0 values) of the same GLs are significantly correlated between the seasons, except for glucobrassicin and neoglucobrassicin, which were present at low levels compared to the aliphatic GLs.
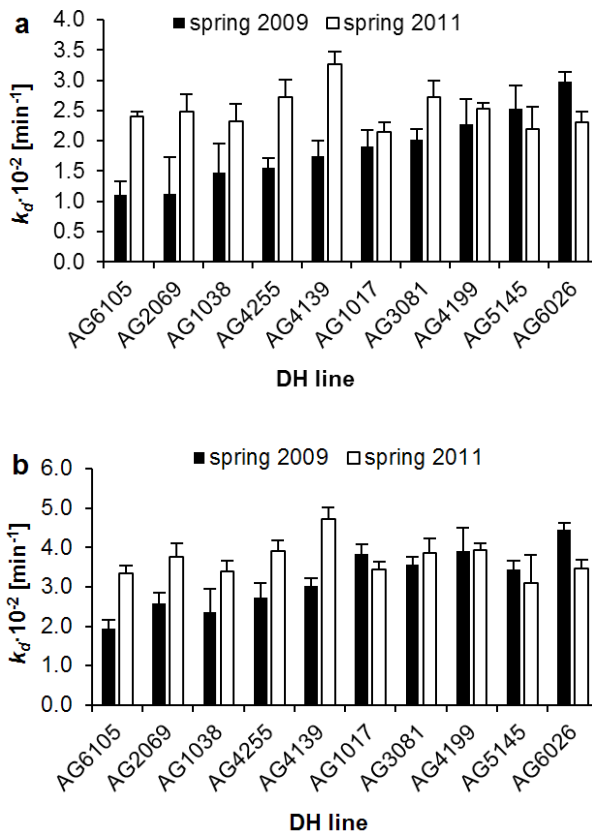

Figure 4.6: Comparison of the rate constants (k_d values) in a subset of the DH population grown in spring 2009 and grown in spring 2011 for glucoraphanin (a) and glucobrassicin (b). The bars represent obtained k_d values from a first order model and the error bars represent the standard deviations

4

Table 4.4: Correlation coefficients for glucosinolate thermal degradation (k_d values) of a subset of the DH population (10 lines) in 2011, when present in the same DH lines (Abbreviations: aliphatic GLs: glucoiberin (IB), progoitrin (PRO), sinigrin (SIN), glucoraphanin (RAPH), gluconapin (NAP); indolic GLs: glucobrassicin (GB), neoglucobrassicin (NEO)).

k_d 2011	IB	PRO	SIN	RAPH	NAP	GB	NEO
IB	1	0.963	0.862	0.929*	0.894	0.874	0.973*
PRO	0.963	1	0.881	0.976	0.963	0.961	0.920
SIN	0.862	0.881	1	0.871	0.919	0.809	0.785
RAPH	0.929*	0.976	0.871	1	0.867	0.944*	0.898*
NAP	0.894	0.963	0.919	0.867	1	0.878	0.971
GB	0.874	0.961	0.809	0.944*	0.878	1	0.835
NEO	0.973*	0.920	0.785	0.898*	0.791	0.835	1

* Correlation is significant at the level 0.05 (2-tailed, corrected for multiple testing: $\alpha < 0.00238$)

Table 4.5: Correlation coefficients for glucosinolate thermal degradation (k_d values) and the initial glucosinolate concentrations (c_0 values) of a subset of the DH population (10 lines) in 2011 and correlation with values determined in the same DH lines in 2009, when present in the same DH lines.

Glucosinolate	Correlation k_d - c_0 values 2011	Correlation k_d 2009 - 2011	Correlation c_0 2009 - 2011
Glucoiberin	0.160	-0.257	0.926**
Progoitrin	0.963**	0.063	0.998**
Sinigrin	0.599	0.281	0.992**
Glucoraphanin	0.356	-0.217	0.944**
Gluconapin	-0.261	0.309	0.997**
Glucobrassicin	-0.195	0.005	0.541
Neoglucobrassicin	0.386	0.353	-0.91

** Correlation is significant at the level 0.01 (2-tailed), * correlation is significant at the level 0.05 (2-tailed)

4.4 Discussion

A publically available DH population, obtained by crossing a broccoli genotype with a Chinese kale genotype [71], was used to study a novel agronomic trait, GL degradation rate during food processing. A mechanistic approach was applied by studying the GL degradation solely induced by heat without degradation caused by the enzyme myrosinase. Previous research has shown that enzymatic degradation during processing of *Brassica* vegetables is far less important for GL loss than thermal degradation. Experimental data on GL thermal degradation have been modelled to obtain first order rate constants (k_d values) to compare the degradation speed of the GLs in different DH lines. A rapid method for quantitatively determining the rate constants was used (chapter 3,[90]). Threefold difference in GL degradation rate constants throughout the DH population were observed for most of the GLs, except for glucoraphanin which differed by 9-fold. GL degradation in the parental lines was investigated in autumn 2010 and spring 2011 (chapter 3, [90]). The degradation rates of structurally identical GLs did not differ significantly between the parental lines, a Chinese kale and a broccoli, within each season; however, a large variation for GLs degradation rate constants in the DH progeny was observed, a phenomenon called transgressive segregation, which occurs especially if multiple genetic locus controls the trait and when both parents have alleles with opposite effects at the different loci [93].

Previously, thermal degradation of the GLs gluconapin, glucobrassicin and 4-methoxyglucobrassicin was studied in five different *Brassica* vegetables by Dekker *et al.* [35]. The highest difference in GL degradation rates, 21-fold, was observed for gluconapin between broccoli and Brussels sprouts, whereas 6-fold difference were found for the degradation rates of glucobrassicin and 4-fold difference for 4-methoxyglucobrassicin between red cabbage and Brussels sprouts. Dekker *et al.* [35] suggested differences in the reaction environment caused by the different plant matrices of the vegetables as a reason for the differences in GL degradation. The term "plant matrix" refers here to the metabolite composition which is the biochemical basis of a vegetable since fully ground tissue was used to study GL thermal degradation and hence the cell and tissue structure does not influence GL degradation. In our study, the differences in GL degradation rates throughout the DH population were smaller than in the study of Dekker *et al.* [35], who investigated very different edible parts of each studied vegetable, which are the inflorescences of Broccoli, the auxiliary buds of Brussels sprouts, the leaves of the

cabbages and pak choi at commercial maturity bought in a supermarket. The vegetables are likely produced in different geographical regions and had different transport and storage histories. So the observed differences in that case were the combined effect of growing conditions, genetic variation, developmental variation and plant age. The DH lines of the DH population in this study were grown in the same conditions and the same organs were investigated, which is likely the reason for smaller differences in the plant matrix and hence a smaller differences in the GL degradation rates. Additionally, the DH population reflects only the genetic variation of two parents which is only part of the variation within the species. Identified factors in the plant matrix that influence GL degradation rates are pH and water content [31, 84].

In our study we could identify two QTLs for the degradation of glucobrassicin, explaining in total 25% of the explained variation. One QTL was identified for the degradation of glucoraphanin, explaining 12% of the phenotypic variation, which co-localised with one of the QTL identified for the glucobrassicin degradation. With these results the hypothesis, that there is a (partly) genetic regulation of GL degradation, can be accepted. The genetic regulation of GL degradation is most likely indirectly via factors in the plant matrix influencing the stability of GLs during heat treatment. The two QTLs identified for the degradation of glucobrassicin and the transgressive segregation show that GL degradation is a complex trait with several loci involved. Furthermore, the GL degradation appears to be strongly affected by the season and year. In the parental lines, thermal degradation was between 20% and 80% higher in spring 2012 compared to autumn 2011 depending on the type of GL (chapter 3, [90]). In contrast to the GL degradation in the parental lines, the degradation of the GLs in subset of the DH population (10 DH lines) was in some DH lines increased and in some DH lines decreased between the two seasons, which indicates that the factors that form the plant matrix were differently affected by the season in different DH lines. The correlation of the degradation of different GLs within both seasons, suggests that the same factors in one DH line influence the degradation rate of different GLs. These factors, which may be specific metabolites, are assumed to be not only affected by the genotype of the DH line, but also by the environment, since the GL degradation of the same GL was not correlated between the two years.

The presented correlation data demonstrate that the speed of the GL degradation is independent from the initial GL concentration in the DH population. Furthermore, the identified QTLs for GL degradation do not co-localise with the identified QTLs

for the initial GL concentration. Several studies have been conducted to elucidate the genes involved in GL biosynthesis, such as the AOP-family (2-oxoglutarate-dependent dioxygenase), MAM-family (methylthioalkylmalate synthase) and MYB-family [41, 57, 85]., Lou *et al.* [87] predicted the AOP genes on two homologous loci, one on linkage group A03 and another one on linkage group A09, in *Brassica rapa*. The QTLs for the initial GL concentration in this *Brassica oleracea* population mapped on linkage group C03 and on top of linkage group C09; the genes underlying these QTLs are possibly AOP due to the synteny between the *B. oleracea* and *B. rapa* genome.

The precise factors underlying the differences in GL degradation between the different DH lines remain unknown. If oxygen played a role in GL thermal degradation, antioxidants are possible substances to prevent GL degradation. Possibly a combination of several factors determines the GL degradation speed in a DH line. The identification of GL thermal degradation products in broccoli sprouts showed that the corresponding nitriles are the dominant breakdown products and the corresponding isothiocyanates are formed in minor quantities [36]. Since the thermal breakdown products are the same as the enzymatic breakdown products formed by myrosinase (only the profile is different), it can be speculated that during heating the glucose is hydrolysed from the GL leaving the same unstable intermediate as in the enzymatic breakdown, which rearranges into nitriles and isothiocyanates. However, the formed breakdown products do not provide information about the factors influencing GL degradation. The identified QTL regions are too large to scan for orthologous genes with an assigned function in Arabidopsis, because there is no indication which type of genes could cause the differences in GL degradation. Fine mapping of the DH population is an option to narrow down the number of genes. Further analysis of the metabolic composition of the DH lines will shed light on the factors retarding or increasing GL degradation.

QTL studies for food processing traits have not been performed frequently, only two examples were found in literature: One on the colour retention during canning of black beans [94] and one on the pasting properties of barley [95]. Both authors identified multiple QTL on several chromosomes which were affected by the environmental and growth conditions. These studies, similar to the presented study on GL degradation, show that food processing traits are complex traits and the genes underlying these traits are not known or only partly known. QTL studies could not only facilitate marker-assisted breeding to select for varieties that perform better during food processing, but also represent the first steps toward identification of the genes responsible for

4

these QTLs. Identifying these genomic regions is the first step towards understanding the mechanism of GL degradation and identification of factors that influence this degradation.

In conclusion, this study shows that QTL mapping for GL degradation rate constants during food processing is possible and gives new opportunities to develop vegetables with an increased retention of their health promoting properties during industrial processing or home cooking. Breeding for vegetables with a high retention of GLs during processing can complement the development of vegetables with high initial GL content to obtain vegetables with a high concentration of health promoting compounds at the stage of consumption.

4.5 Acknowledgements

4

CHAPTER 5

A Metabolomics Approach to Identify Factors Influencing Glucosinolate Thermal Degradation Rates in *Brassica* Vegetables

K. Hennig; R. de Vos; C. Maliepaard, M. Dekker, R. Verkerk; G. Bonnema; *submitted for publication*

Abstract

Thermal processing of *Brassica* vegetables can lead to substantial loss of potential health-promoting glucosinolates. The extent of thermal degradation of a specific glucosinolate varies in different vegetables, possibly due to differences in the composition of other metabolites within the plant matrices. An untargeted metabolomics approach followed by random forest regression was applied to identify metabolites associated to glucosinolate thermal degradation in a segregating *Brassica oleracea* population. Out of 413 metabolites, 15 were significantly associated with the degradation rate constant of glucobrassicin, 6 with the rate constant of glucoraphanin and 2 with both GL degradation rate constants. Among these twenty-three different metabolites three were identified as flavonols (one kaempferol- and two quercetin-derivatives) and two as glucosinolates (4-methoxyglucobrassicin, gluconasturtiin). Twenty quantitative trait loci (QTL) for these metabolites that were significantly associated to glucoraphanin and glucobrassicin degradation were identified on linkage groups C01, C07 and C09. Two flavonols mapped on linkage groups C07 and C09 and co-localise with the QTLs for glucosinolate degradation determined previously.

Keywords

5.1 Introduction

Brassica vegetables contain a specific group of phytochemicals, glucosinolates (GLs) that are almost exclusively found in this plant family. GLs coexist with an endogenous enzyme, myrosinase (E.C. 3.2.1.147), in the plant tissue but physically separated to avoid hydrolysis until tissue damage. Epidemiological and mechanistic studies have shown health protective effects of GLs and their enzymatic hydrolysis products against several types of cancer, for example colon, colorectal, breast, bladder and prostate cancer [15]. However, as a result of food processing of *Brassica* vegetables the amount of GLs can be lowered substantially and hence there is growing interest to minimize losses during vegetable processing. Several mechanisms lead to losses of GLs during food processing: a) enzymatic breakdown of GLs, b) leaching of GLs and breakdown products into the cooking water and c) thermal degradation [23]. The term thermal degradation refers to the degradation solely induced by heat. Losses of 78% of the total GLs caused by thermal degradation were estimated during canning of red cabbage [32]. Thermal degradation in Brussels sprouts reduced one type of GL, the indolic glucobrassicin by 60% after heating for 15 min at 100°C [35]. It has been shown that the thermal degradation is dependent on the chemical structure of the GL [31, 32], but also on the reaction environment, i.e. vegetable matrix, since the thermal degradation of chemically identical GLs differs in different vegetables [35]. It was hypothesised that the different matrices in these vegetables caused the differences in GL degradation. In this chapter, the term "matrix" refers to metabolites and other components present in the vegetables, since fully ground tissue was used to determine the thermal degradation and influences of the cell walls and tissue structure could be excluded [35]. The metabolite composition of plants is (partly) genetically determined [96], hence we hypothesised that thermal degradation of GLs can be (partly) genetically regulated. Quantitative trait loci (QTL) mapping is the association of quantitative traits with molecular markers and is one possibility to identify genetic regions [50]. If GL degradation is genetically regulated, QTL may be identified for this trait and may help plant breeders in developing *Brassica* vegetables with an improved matrix for thermal stability of GLs. The genetic effects of thermal degradation have been investigated in chapter 4 [97], where QTLs explaining part of the variation in thermal degradation were identified. Furthermore, it has been shown in another study that environmental factors like the growing season influence thermal degradation (chapter 3, [90]). From these results the question arises which mechanism underlies the identified QTLs.

Metabolomics is an important comparative tool to study global metabolite levels of e.g. plant materials treated with various conditions. Untargeted metabolomics approaches are used to measure as many metabolites as possible to obtain patterns or fingerprints of processed food samples [52, 53]. Since there is yet no knowledge on which metabolites can influence GL thermal degradation, an untargeted metabolomics approach is a promising tool to associate differences in metabolite composition to differential GL thermal degradation. In the present study an untargeted LC-MS-based metabolomics approach was applied to test if GL thermal degradation during food processing can be associated to metabolites. If certain metabolites influence the degradation rate of GLs, breeding for changing the concentrations of these metabolites can result in improved stability of GLs during food processing. The screening and selection for genotypes by analysing their concentration of certain metabolites is less laborious than testing the GL degradation during food processing in different genotypes. Hence the knowledge of metabolites influencing GL thermal degradation will facilitate breeding for vegetables with a high retention of GLs during food processing and will help to maximize the health promoting effects of GLs in *Brassica* vegetables at the stage of consumption.

5.2 Materials and methods

5.2.1 Plant material

A Doubled haploid (DH) population, developed by Bohuon *et al.* [71], was used to study GL thermal degradation. This population was developed by crossing two DH parents, a rapid-cycling Chinese kale line, *B. oleracea* var. *alboglabra* (A12DHd), and a Calabrese broccoli line, *B. oleracea* var. *italica* (GDDH33), through microspore culture of the F_1. A total of 100 DH lines of the DH population were grown in spring 2009 and a subset of 10 DH lines was grown again in spring 2011. Growing, harvesting and subsequent microwave treatments to inactivate the myrosinase were performed as described in the materials and methods of chapter 4.

5.2.2 Sample treatment

A preliminary study was performed to investigate the effect of heating on the metabolite composition of three DH lines of the DH population grown in 2009 and 2011 (AG6105, AG1017 and AG6026). The lines were selected based on previously determined thermal degradation rate constants (k_d) of glucoraphanin (4-methylsulfinylbutyl-GL, compound 1 in Figure 5.1) and glucobrassicin (Indol-3-ylmethyl-GL, compound 2 in Figure 5.1) which

were the GLs present in all lines of the DH population (chapter 4). Thermal degradation rate constants were calculated by modelling the measured GL concentrations over the heating time at 100°C (0, 15, 30 & 60 min) using a first order kinetics as described in chapter 3, [90]. The degradation rate constant (k_d) describes the steepness of the degradation curve. Plant material that was not microwaved was analysed without heating, whereas the microwaved material (with inactivated myrosinase) was used for the heating study. Dried plant powder was reconstituted with MilliQ water to obtain 0.5 g plant material having the same water content as before the microwave treatment and drying. Closed tubes containing this plant material were then placed in a heating block at 100°C for 0, 15, 30 and 60 min. The average time for samples to reach boiling point (100°C) was 4 min. After heating, samples were quickly cooled on ice and stored at -20°C till metabolite analysis.

The preliminary study showed a main effect of the genotype and a smaller effect of the heating step on the metabolite profiles. Hence the metabolites of 85 DH lines were analysed in the main study in the microwaved plant material, without further heating.
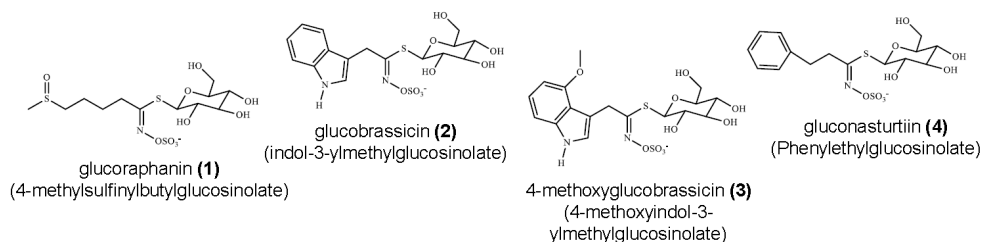
Figure 5.1: Chemical structures of some glucosinolates identified in the DH population

5.2.3 Metabolite profiling

Extraction of semi-polar metabolites was performed according to De Vos *et al.* [98]. For the preliminary experiment the reconstituted samples (corresponding to 0.5 g fresh weight) were extracted with 1.5 mL methanol (99.867%) acidified with formic acid (0.133%) and shaken with a vortex immediately. Based on 95% water in the plant tissue the final solvent concentration is 75% methanol with 0.1% formic acid. The dried plant material corresponding to the amount present in 0.5 g fresh weight of the 85 lines from the main study was directly extracted with 1.5 mL methanol (75%) acidified with 0.1% formic acid. Metabolites were extracted by 15 min sonication at 40 kHz and subsequently

centrifuged at 20 000g at room temperature. The supernatants were filtered through 0.2 µm PTFE filters, transferred into vials and analysed directly. All samples were extracted and analysed in a randomised order. Chromatographic separation was performed on a Luna C18 column (2.0 x 150 mm, 3 µm pore size, Phenomenex) on a Waters Alliance HPLC and detected by photodiode array detector (Waters 2997) followed by an accurate mass QTOF-MS (Waters QTOF Ultima) operating in negative ionization mode. LC gradient and MS settings were as described in De Vos *et al*. [98].

5.2.4 Data processing

The software MetAlign (www.metalign.nl; [99]) was used for peak picking at a minimal signal to noise ratio of 3 and for ion-wise mass spectral alignment. Subsequently peaks present in at least 3 samples at an intensity of at least 3 times the noise were selected, leaving 1310 reproducible and reliable mass features. Absent values were replaced by randomised values between 2.5 and 3 times the noise. Afterwards, 413 metabolites were reconstructed by using the tool MSClust, in which an unsupervised clustering method based on similarities between mass features of both chromatographic retention time and relative abundance over samples is applied [100]. The annotation of the reconstructed metabolites was based on in-house metabolite databases of accurate masses combined with experimentally obtained retention times, as well as based on accurate masses of known broccoli metabolites as published by Vallejo *et al*. [101]. The metabolite annotations were divided into 4 different identification levels, according to Sumner *et al*. [102]. The elemental composition calculator of the MassLynx acquisition software (Waters) of the LC-QTOF MS was used to calculate elemental compositions based on accurate masses and isotopic fit for the metabolites that could not be annotated with other databases.

5.2.5 Statistical analyses

From each reconstructed metabolite (mass cluster) the ion with the highest intensity within the mass cluster (SIM) was selected and subjected to statistical analysis. Data were log transformed and mean centred prior to principal components analysis (PCA), which was performed using the software Latentix™ 2.11 (http://www.latentix. com). The software Multiexperiment viewer (http://www.tm4.org/mev/about, [103]) was used to perform hierarchical cluster analysis after log transformation of the data using Euclidean distance and weighed average linkage clustering algorithm. Pearson correlation coefficients were calculated with previously determined thermal

degradation rate constants (k_d values) and log intensities of the selected mass signals from 413 metabolites using Microsoft Excel 2010.

Random forest regression [104] was performed using the Random Forest library in R statistical software (http://www.r-project.org/) to relate the previously determined GL thermal degradation rate constants (k_d values) of glucobrassicin and glucoraphanin, separately, to the log transformed intensity values of the 413 metabolites measured with LC-MS. Random forest importance values (i.e. the increase in mean squared error (MSE) after permutation of each individual metabolite) were used to quantify the relevance of individual metabolites (as ranked in Table 5.2). Significance of the Random Forest models and of the individual metabolites was assessed using a second permutation approach: here the response was permuted 100 times and the random forest model was re-run with the permuted data, and importance values were stored for all metabolites. Then the 95%-percentiles of R^2 values (for significance of the model) and the importance values were used as threshold levels for significance. For glucoraphanin the 95% percentile of the R^2 value was 3.3% and for glucobrassicin 3.4%. The random forest prediction values R^2 represent the explained variance for the samples which were not used to build the random forest model (out of the bag samples) and were predicted using the built model over all metabolites.

5.2.6 QTL mapping

The linkage map, used for QTL analysis, has been constructed with a subset of 107 DH lines of the AGDH population. The map was constructed based on restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP) and microsatellite markers published by Sebastian *et al.* [91]. The final map was constructed with the subset of 107 lines using the software JoinMap 4.0. The genetic map constructed with 210 lines of the plant population is publically available on www.brassicainfo.com. The total length of the two maps are comparable, the length of the individual linkage groups differ, due to a different number of plant lines used for the map construction in this study.

QTL mapping was performed using the Software MapQTL 6.0 [92] using the GL degradation rate constants (k_d values) of glucobrassicin and glucoraphanin (compounds 1 & 2 in Figure 5.1) determined in chapter 4, and from each metabolite the single ion fragment with the highest intensity (SIM) as phenotypic traits. The analysis started with interval mapping (IM) and a permutation test was applied to each dataset (1000

5

repetitions) to decide the LOD (logarithm of odds) threshold (p ≤ 0.05). In this study an average genome-wide LOD value of 2.8 was used as significant threshold for the degradation rate constants and the metabolites. The results of the interval mapping were used to identify QTLs, which were selected as cofactors if their LOD value exceeded the chromosome-wide threshold. Cofactors were verified by using the automatic cofactor selection tool and only significant markers (P < 0.02) were used as cofactors for the subsequent restricted multiple-QTL model mapping (RMQM). A map interval of 5 cM was used for both IM and RMQM analyses.

5.3 Results

5.3.1 Metabolite profiling during processing of three DH lines grown in two different years

In order to estimate differences in semi-polar metabolite profiles between genotypes and growing year, three doubled haploid (DH) lines of the DH population were selected for LC-MS analysis in a preliminary experiment. Since the aim of this study was to test whether metabolites related to GL thermal degradation rates can be identified, these three lines were selected based on their differences in degradation rate constants (k_d) determined in a previous study. The k_d value of DH line AG 6105 increased in 2011 compared to 2009, for DH line AG 1017 it remained almost constant and for DH line AG 6026 it decreased (Table 5.1) & (chapter 4). To investigate the effect of heating on metabolite profiles, leaves of three DH lines grown in two different years were analysed fresh (without prior microwave treatment), after microwave treatment (time point 0) and after heating at 100°C up to 60 min. From these LC-MS data 227 metabolites were extracted in all the samples with varying concentrations. Hierarchical cluster analysis (HCA) showed three main clusters each representing a different genotype. Each genotype cluster is separated again based on the growing year and the smallest clusters are formed by changes in metabolite composition during heating (Figure 5.2 A). Similar effects were visible in the principle component analysis (PCA, Figure 5.2 B, C): the first PC (x-axis) separates all genotypes from each other and in the second PC (y-axis) effects of both the genotype and the growing years are visible. The first two components account for 57% of the variance, hence it is concluded that the main factors influencing metabolite composition was firstly the genotype and secondly the growing year. The effect of processing is clearly expressed in PC 4 only (11% explained variance): all the fresh samples are located in the positive part of PC 4 and processing

times are increasing in the negative direction of PC4. Genotype AG6105 showed the lowest GL degradation rate constant in 2009 and genotype AG6026 showed the highest rate constant in 2009, whereas the rate constants of the genotypes AG6105, AG6026 in 2011 and AG1017 in both years are very similar (Table 5.1).

Figure 5.2: Dendogram of hierarchical cluster analysis (HCA) and scores plots of principle component analysis (PCA) of metabolite profiles (SIM) of three doubled haploid lines, grown in two years and processed for several time points. Sample codes first represent the growing year, followed by the processing time: F = fresh sample, 0 = microwaved sample, 15 = microwaved & 15 min heated, 30 = microwaved & 30min heated, 60 = microwaved and 60 min heated at 100°C. One colour represents one genotype: green = AG6026, pink = AG1017, blue = AG6105. Figure A: Dendogram from HCA, B: x-axis = PC1, y-axis = PC2; figure C: x-axis = PC1, y-axis = PC4.

Table 5.1: Glucosinolate thermal degradation rate constants (k_d value) in three different DH lines of the DH population for two different types of glucosinolates in two growing years determined in chapter 4. Rate constants were determined by modelling glucosinolate concentrations over the heating time using a first order kinetics. The rate constants describe the steepness of the degradation curve.

DH line	k_d value glucobrassicin [min⁻¹]		k_d value glucoraphanin [min⁻¹]	
	2009	2011	2009	2011
AG 6105	0.019	0.033	0.011	0.024
AG 1017	0.038	0.034	0.019	0.023
AG 6026	0.045	0.035	0.030	0.023

No separation in the PCA according to this pattern could be observed. In summary, HCA and PCA show that the overall LC-MS profile of the selected DH lines was most affected by genotype while thermal treatment had much smaller effects. For the main study of this chapter it was therefore sufficient to only analyse the metabolite profiles after microwave blanching to inactivate myrosinase, without further heating.

5.3.2 Metabolite profiling of the segregating *Brassica* population

A more comprehensive study was performed by analysing semi-polar metabolites in leaves of 85 lines from the DH population (all grown within the same year), after microwave treatment, without further heating (i.e. time point 0 of thermal treatment). From the LC-MS data, a total of 413 reconstructed metabolites was detected and HCA based on the relative intensities of these metabolites revealed four main clusters of genotypes (Figure 5.3A). One of these clusters was separated by PC1 (22% of the explained variance), whereas the other clusters did not show clear groupings (Figure 5.3B). The genotypes in the scores plot are coloured according to the GL degradation rate constants of glucobrassicin (compound 2 in Figure 5.1). The genotypes with higher glucobrassicin degradation rate constants tend to be in positive direction of PC2 (15.5% explained variance) and the genotypes with lower degradation rate constants tend to be in the negative direction of PC2, but there is no clear separation. The loadings plot (Figure 5.3C) of the first two PCs did not show a separation of the metabolites, likely because the analysed population is a segregating population with a continuous variation of traits throughout the population. PCA could not reveal a relation between GL degradation rate constants and metabolite variation.
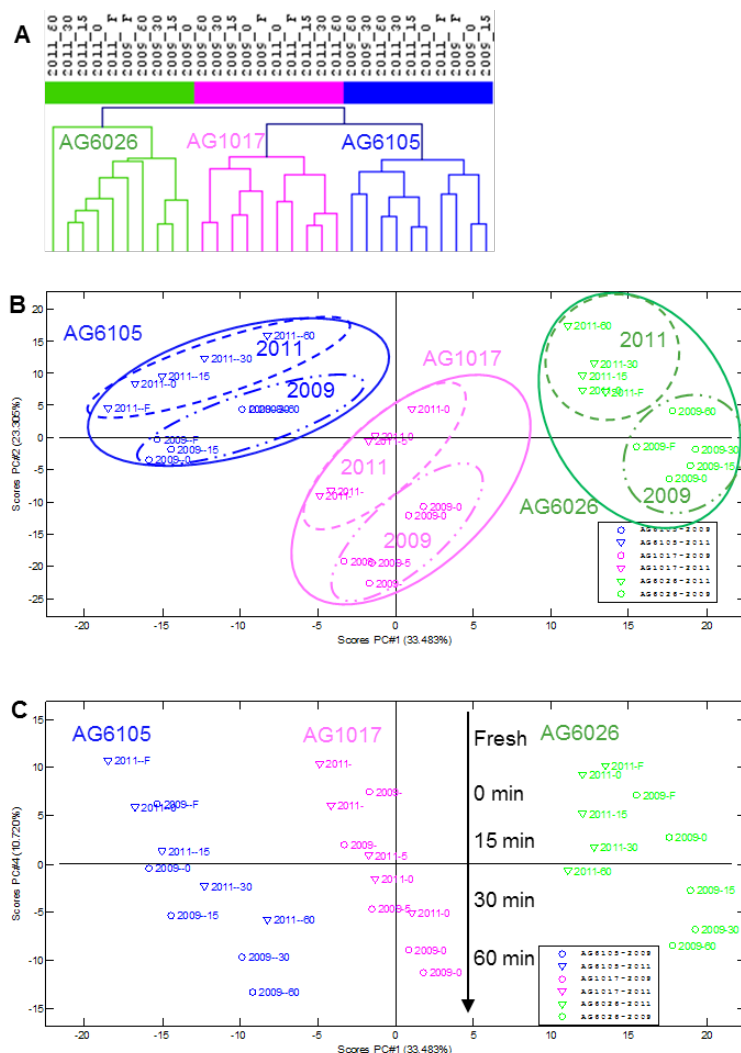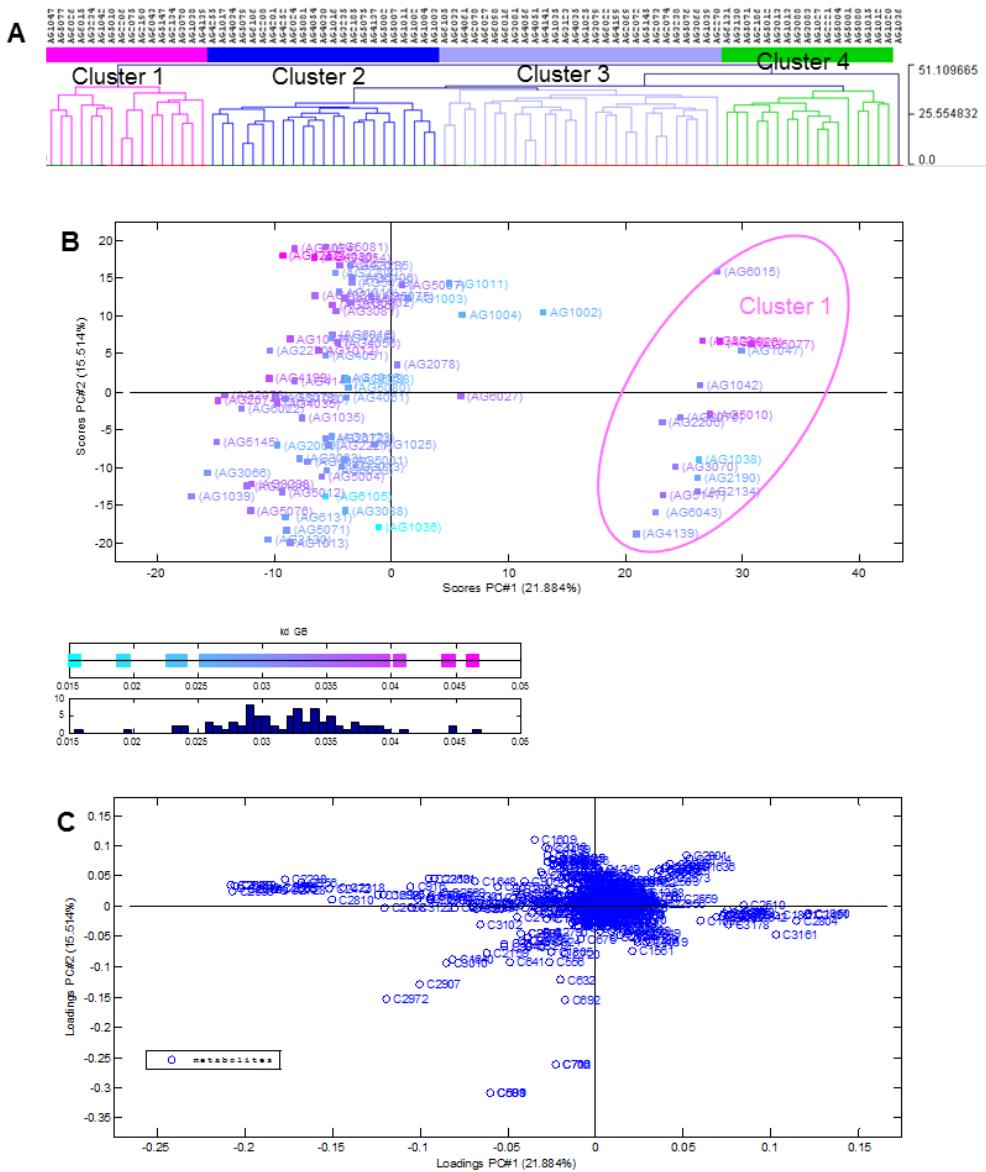
Figure 5.3: Dendogram of hierarchical cluster analysis (HCA) and scores and loadings plots of principle component (PC) analysis of metabolite profiles (SIM) of 85 doubled haploid lines (grown it the same year) after microwave treatment. Figure A: Dendogram from HCA, B: scores plot: x-axis = PC1, y-axis = PC2, the colours in the scores plot represent the glucosinolate degradation rate constant of this genotype determined previously according to the colour scale; figure C: loadings plot: x-axis = PC1, y-axis = PC2. The pink circle in the scores plot (Figure B) represents the genotype in cluster 1 retrieved from HCA.

5.3.3 Relating metabolites to glucosinolate degradation rate constants

Subsequently, random forest regression [104] was used to relate the GL degradation rate constants of the two GLs present in all the lines of the DH population (compound 1 & 2 in Figure 5.1) to the relative intensities of the 413 LC-MS metabolites. The random forest prediction values (R^2) are a measure of the performance of the model and represents the explained variance for the samples which were not used to build the random forest model over all metabolites. The R^2 value of the model was 27.0% for glucobrassicin and 8.3% for glucoraphanin, which are both above the significance threshold (3.4%). These results show that the GL degradation rate constants could be partly predicted using a set of semi-polar metabolites. Seventeen metabolites significantly contributed to the explained variance of the degradation rate constant of glucobrassicin and eight metabolites contributed to that of glucoraphanin (Table 5.2). The metabolites are ranked based on the contribution of each compound to the prediction of the degradation rate constant of both GLs, while the correlation coefficient indicates the direction of the effect. Ten metabolites were positively correlated to the degradation of glucobrassicin, while seven metabolites were negatively correlated. With regard to degradation of glucoraphanin, four metabolites were positively correlated and four metabolites were negatively correlated. The metabolite information and their putative identities are given in Table 5.2. Two GLs, i.e. 4-methoxyglucobrassicin (GB 16 & GR 4, compound 3 in Figure 5.1) and gluconasturtiin (GR 7, compound 4 in Figure 5.1), are negatively correlated to the degradation of glucoraphanin, while 4-methoxyglucobrassicin was also negatively correlated to the degradation of glucobrassicin. These negative correlations suggest a lower degradation with increasing amounts of these two compounds. Two quercetin-derivatives (GB 10 & GB 14 in Table 5.2) were positively correlated to the degradation rate constants of glucobrassicin, while eight other yet unknown metabolites were also positively correlated to the degradation rate constant of glucobrassicin. Another flavonol-derivative (GB 6) showed no correlation (r = -0.03) with the degradation rate constants of glucobrassicin.

As shown in Figure 5.2, heating alters the overall metabolite composition, although the effect of the genotype was much stronger in the three DH lines tested. To check the thermal stability of those identified metabolites that were significantly correlated to the GL degradation rate constants in the entire DH population, the preliminary dataset with the thermal treatment was re-evaluated for these compounds. Observed peak intensities of both, the two GLs and the three identified flavonol-derivatives over different heating

times are displayed in Figure 5.4. In all three DH lines the degradation of these GLs varied between 0% and 60% after 15 min of heating, whereas the three flavonol-derivatives were not degraded. Sixty min of heating degraded these GLs completely. In contrast, the three flavonol-derivatives were degraded by only 20% after 60 min of heating.
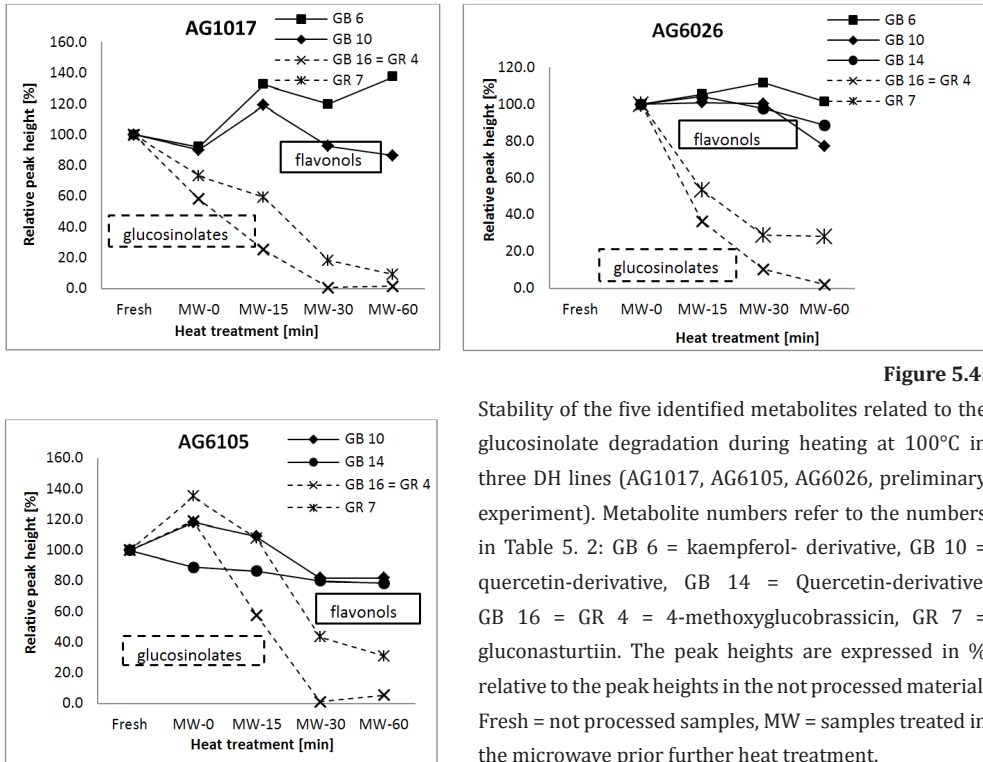
Figure 5.4:

Stability of the five identified metabolites related to the glucosinolate degradation during heating at 100°C in three DH lines (AG1017, AG6105, AG6026, preliminary experiment). Metabolite numbers refer to the numbers in Table 5. 2: GB 6 = kaempferol- derivative, GB 10 = quercetin-derivative, GB 14 = Quercetin-derivative, GB 16 = GR 4 = 4-methoxyglucobrassicin, GR 7 = gluconasturtiin. The peak heights are expressed in % relative to the peak heights in the not processed material. Fresh = not processed samples, MW = samples treated in the microwave prior further heat treatment.

5.3.4 Quantitative trait loci mapping of the metabolites

QTL mapping was performed for those metabolites which were significantly associated with GL thermal degradation, in order to check whether the QTLs for these metabolites and the QTLs for GL degradation rate constants overlap. The QTLs on linkage groups C01, C07 and C09 (where most of the identified QTLs are located) are displayed in Figure 5.5 together with the QTLs of the GL degradation rate constants determined in our previous study on the same plant material (chapter 4).

QTL of metabolites significantly associated to the degradation rate of glucobrassicin

The QTL of two metabolites (GB 3, GB 10 in Table 5.2) associated to the degradation rate constant of glucobrassicin co-localised with the QTL for the degradation rate constant for glucobrassicin on linkage groups C09 and C07. The metabolite GB 10 was identified as quercetin-3-sophoroside-7-glucoside-2-caffeoyl whereas metabolite GB 3 could not be identified yet. The QTL of two other metabolites (GB 12, GB 14; Table 5.2) co-localised with the locus for the degradation rate constant of glucobrassicin, which was present on linkage group C09, while the QTL of another five metabolites (GB 1, GB 2, GB 4, GB 7, GB 11; Table 5 2) co-localised with the locus for the degradation rate constant of glucobrassicin on linkage group C07. Metabolite GB 14 was identified as quercetin-3-sophoroside-7-glycoside, whereas the other metabolites could not be identified yet. Two of these nine metabolites mapping on linkage groups C07 and C09 were identified as flavonol-derivatives and show a co-localization with the yet unknown five metabolites. The intensities of these nine metabolites are positively correlated with the degradation rate constant of glucobrassicin. No QTL was identified for the third identified flavonol-derivative. The QTL of three metabolites (GB 5, GB 9, GB 15 in Table 5.2) that were significantly associated to the degradation of glucobrassicin were located on linkage group C01, where no QTL for degradation rate constants have been identified. These three metabolites, which could not be identified yet, were negatively correlated with the degradation rate constant of glucobrassicin.

QTL of metabolites significantly associated to the degradation rate of glucoraphanin

For the degradation rate constant of glucoraphanin one QTL was identified previously on linkage group C09 co-localizing with one QTL for the degradation rate constant of glucobrassicin (chapter 4). No metabolite QTL related to the degradation rate constant of glucoraphanin were co-localizing on linkage group C09. On linkage group C07, co-localizing with metabolite QTL related to the degradation of glucobrassicin, four QTL for metabolites related to the degradation of glucoraphanin are located (GR 1, GR 2, GR 7, GR 8 in Table 5.2). The intensities of three of the metabolites (GR 2, GR 7, GR 8) are positively correlated with the degradation rate constants of glucoraphanin and a negative correlation was found for the intensities of one other metabolites (GR 1) mapping on this linkage group. Metabolite GR 7 was identified as gluconasturtiin (compound 4 in Figure 5.1). Another GL, 4-methoxyglucobassicin (compound 3 in Figure 5.1), was also significantly related to the degradation of glucoraphanin, but no QTL could be identified for this GL.

Table 5.2: Metabolites and identities with significant relation to degradation rate constants of glucoraphanin and glucobrassicin determined with random forest regression.

Metabolite number	Rank to the RF model[1]	Corr. coeff.[2]	Retention time [min]	Observed mass [M-H]⁻ (m/z)	Mass difference [ppm][3]	Elemental formula	Putative identity	Level of annotation[4]	Linkage group of QTL
						Glucobrassicin			
GB 1	1	0.51	15.4	1095.28332	0.4	$C_{48}H_{56}O_{29}$	(flavonol-phenylpropanoid-glucoside)[A]	4	C07
GB 2	2	0.49	17.9	1079.29077	2.1	$C_{48}H_{56}O_{28}$	(flavonol-phenylpropanoid-glucoside)[A]	4	C07
GB 3	3	0.54	25.0	1317.34143	3.9	$C_{59}H_{66}O_{34}$	(flavonol-phenylpropanoid-glucoside)[A]	4	C07 & C09
GB 4	4	0.50	19.1	1095.28833	5.0	$C_{48}H_{56}O_{29}$	(flavonol-phenylpropanoid-glucoside)[A]	4	C07
GB 5	5	-0.24	35.0	587.25220	4.1	$C_{31}H_{40}O_{11}$		4	C01
GB 6	6	-0.03	13.0	771.19788	1.4	$C_{33}H_{40}O_{21}$	Kaempferol-3-sophoroside-7-glucoside[C]	2	/
GB 7	7	0.51	27.0	1271.33887	2.2	$C_{51}H_{68}O_{37}$		4	C07
GB 8	8	-0.05	12.8	861.19525	1.8	$C_{35}H_{42}O_{25}$		4	/
GB 9	9	-0.29	18.6					4	C01
GB 10	10	0.45	13.8	949.22638	1.45	$C_{42}H_{46}O_{25}$	Quercetin-3-sophoroside-7-glucoside-2-caffeoyl[B]	3	C07 & C09
GB 11	11	0.46	27.7	1447.4149		$C_{53+76}N_8O_{39}$			C07
GB 12	12	0.40	24.3	1479.3986	6.4	$C_{65}H_{76}O_{39}$	(flavonol-phenylpropanoid-glucoside)[A]	4	C09
GB 13	13	0.34	36.4	531.15118	0.7	$C_{26}H_{28}O_{12}$		4	/
GB 14	14	0.41	15.1	625.14117	0.2	$C_{27}H_{30}O_{17}$	Quercetin-3-glucoside-7-glucoside[D]	2	C09
GB 15	15	-0.24	15.7	285.12512	2.4	$C_2H_{10}N_{18}$		4	C01
GB 16 (=GR 4)	16	-0.30	24.9	477.06320	1.2	$C_{17}H_{22}N_2O_{10}S_2$	4-methoxy-glucobrassicin	1	/
GB 17 (=GR 3)	17	-0.30	24.8					4	/

[1] Ranking of importance to the Random forest (RF) model is based on increase in means square error after permutation of each individual metabolite
[2] Pearson correlation coefficient between glucosinolate degradation rate constant and metabolite intensities (log transformed)
[3] Mass accuracy of the instrument 5 ppm
[4] According to Sumner et al. [102]; 1 = identified compounds, 2 = putatively annotated compounds, 3 = putatively characterised compound classes, 4 = unknown compound
[A] Mass is matching with phenylpropanoids described by Vallejo, Tomás-Barberán & Ferreres [101] and unpublished results in broccoli, however there are no further indications towards specific compounds
[B] Described by Vallejo, Tomás-Barberán & Ferreres [101] in broccoli
[C,D] Described in *Brassica oleracea* and in *Brassica muricata*, respectively, in KNApSAcK database [105]

5

Table 5.2: continued

Metabolite number	Rank to the RF model[1]	Corr. coeff.[2]	Retention time [min]	Observed mass [M-H]⁻ (m/z)	Mass difference [ppm][3]	Elemental formula	Putative identity	Level of annotation[4]	Linkage group of QTL
						Glucobrassicin			
GR 1	1	-0.25	4.4					4	C07
GR 2	2	0.29	39.1	821.324707	3.5	$C_{37}H_{46}N_{10}O_{12}$		4	C07
GR 3 (=GR 17)	3	-0.27	24.8						/
GR 4 (=GB 16)	4	-0.27	24.9	477.063202	1.3	$C_{17}H_{22}N_2O_{10}S_2$	4-methoxy-glucobrassicin	1	/
GR 5	5	0.30	36.1					4	C01
GR 6	6	0.36	45.1	707.295044	0.5	$C_{22}H_{36}N_{20}O_8$		4	C01
GR 7	7	-0.47	23.8	422.059021	2.4	$C_{15}H_{21}NO_9S_2$	Gluconasturtiin	1	C07
GR 8	8	0.36	2.2	488.165588	6.8	$C_{12}H_{28}N_9O_{10}P$		4	C07

[1] Ranking of importance to the Random forest (RF) model is based on increase in means square error after permutation of each individual metabolite

[2] Pearson correlation coefficient between glucosinolate degradation rate constant and metabolite intensities (log transformed)

[3] Mass accuracy of the instrument 5 ppm

[4] According to Sumner et al. [102]: 1 = identified compounds, 2 = putatively annotated compounds, 3 = putatively characterised compound classes, 4 = unknown compound

[A] Mass is matching with phenylpropanoids described by Vallejo, Tomás-Barberán & Ferreres [101] and unpublished results in broccoli, however there are no further indications towards specific compounds

[B] Described by Vallejo, Tomás-Barberán & Ferreres [101] in broccoli

[C,D] Described in *Brassica oleracea* and in *Brassica muricata* , respectively, in KNApSAcK database [105]

Figure 5.5: Quantitative trait loci (QTL) of the metabolites significantly related to the glucosinolate degradation rate constants on linkage groups C01 (red), C07 (green), C09 (purple).The legend on the left side describes to which glucosinolate degradation rate constant the metabolites were significantly related. For comparison the QTL of the glucosinolate degradation rate constants (k_d RAPH = glucoraphanin, k_d GB = glucobrassicin) determined in chapter 4 are displayed. The white colour in the heatmap indicates a logarithm of odd (LOD) score < 2.8 (average genome wide LOD score determined with the permutation test) and from red to yellow the LOD scores are increasing.

5.4 Discussion

Untargeted metabolite profiling of semi-polar metabolites was performed to identify metabolites in the plant matrix influencing the thermal degradation of GLs, as occurring during *Brassica* vegetable processing. PCA revealed a major influence of the genotype on the metabolite profile, followed by the growing season. The effect of heating was more than three times lower than the genotypic effect. To our knowledge, large-scale untargeted metabolomics studies in relation to heating of *Brassica* vegetables have not been reported before, although differences in metabolite profiles have been assessed in relation to differential accessions, cultivars and growing locations [106, 107].

To relate extensive metabolite profiles to GL degradation rate constants, leaves from 85 lines of a segregating DH population, with varying thermal stability of their GLs glucobrassicin and glucoraphanin (chapter 4), were analysed by accurate mass LC-MS. Random forest regression is an approach which has good predictive power even when most regressor variables are not related to the response, it is a suitable tool in case of more variables than observations, as is mostly the case for -omics data, and it is protected against over fitting [104]. Because of these attributes, it is a potentially suitable approach to relate the variation in abundance of the many metabolites present in *Brassica* leaves to the degradation rate constants of GLs, in order to pinpoint candidate compounds influencing GL stability. The explained variance of the random forest model over all metabolites, 27% and 8% for the degradation rate constant of glucobrassicin and glucoraphanin, respectively, show that these semi-polar metabolites can partly predict the degradation rate constants of these two major GLs in *Brassica* vegetables. The use of different extraction solvents to cover a wider range of polarities and analyses by different analytical platforms, such as GC-MS and NMR, would enable the detection of more compounds and possibly increase the predictive power. Random forest regression was also applied by Eggink *et al.* [108] to predict sensory attributes of fresh sweet pepper. Eight different sensory attributes could be significantly predicted with a set of volatile and non-volatile metabolites with on average 46% of explained variance, whereas six sensory attributes could not be predicted. The explained variance was higher compared to our study; but the two studies have very different setups and concern very different attributes, and the absolute levels of explained variance are therefore difficult to compare. Unfortunately no closer related studies are available in literature yet.

For the degradation rate constant of glucobrassicin, 3 of the 17 metabolites which were significantly related could be annotated as derivatives of the flavonols kaempferol and quercetin. The quercetin-derivatives showed a positive correlation with the GL degradation rate constants, suggesting a higher GL degradation rate with increasing levels of these flavonols. Bellostas *et al.* [33] proposed a mechanism for non-enzymatic GL degradation in the presence of ferrous-ions (Fe^{2+}), thereby producing ferric ions (Fe^{3+}) and releasing nitriles (Figure 5.6). Hanschen *et al.* [34] showed an enhanced thermal degradation of the aliphatic GL sinigrin when added to a broccoli matrix compared to degradation in water at the same pH. The addition of ferrous ions to the water led to an increased thermal degradation, although the degradation rate was still lower than in the broccoli matrix at twice as high concentration of ferrous ions in the water, indicating that Fe^{2+} present in the broccoli was not the only determining factor. The simultaneous addition of ferrous ions and vitamin C led to a higher degradation of sinigrin compared to the addition of ferrous ions alone and the authors therefore proposed a synergistic effect of the vitamin C based on its Fe^{3+}-reducing ability resulting in redox-cycling of this metal ion. Macáková *et al.* [109] demonstrated that quercetin also exhibit Fe^{3+}-reducing activity, which is dependent on the pH, the ratio flavonoid/ferric ions and degree of glycosylation. The positive correlation of the quercetin-derivatives with the degradation rate constant of glucobrassicin may therefore relate to this redox-cycling activity. In the study of Hanschen *et al.* [34] the presence of vitamin C and ferrous ions in the buffered sinigrin solution could still not fully explain the degradation of sinigrin observed in the broccoli sprout matrix. Our findings that quercetin-derivatives are associated with increased GL degradation may at least partly explain the remaining gap in the study of Hanschen *et al.* [34]. Eight yet-unidentified metabolites were also positively correlated with the degradation rate constants of glucobrassicin, and it is tempting to speculate that their effect is also based on their Fe^{3+}-reducing power, analogous to the positively correlating quercetin-derivatives.
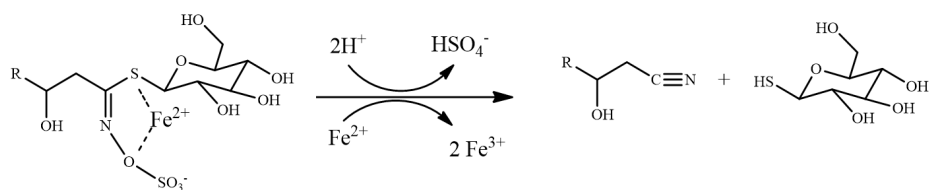

Figure 5.6: Mechanism of non-enzymatic glucosinolate degradation catalysed by ferrous ions (Fe^{2+}) leading to the formation of nitriles, as proposed by Bellostas *et al.* [33]

QTLs could be identified for nine of the ten metabolites that were positively correlated to the degradation rate constant of glucobrassicin and all these QTLs map either on linkage group C07, or on linkage group C09, or on both linkage groups, and hence they co-localise with the QTLs identified for the degradation rate constant of glucobrassicin. The co-localization of the QTLs further supports the identified association of these metabolites with the degradation rate. The two quercetin-derivatives both map on linkage group C07 and one additionally on linkage group C09. Metabolites of the same biochemical pathway often co-localise on the same linkage group, as it was shown for GLs by Lou *et al.* [87].

Seven metabolites related to the degradation rate constant of glucobrassicin show a negative correlation with the rate constant. One of these metabolites was identified as 4-methoxyglucobrassicin, which was degraded fast during 60 min of heating (Figure 5.4). This negative correlation may be explained by a competitive degradation mechanism in which 4-methoxyglucobrassicin is degraded faster and hence can reduce the degradation of glucobrassicin. Confirming this idea, Oerlemans *et al.* [32] showed that the degradation rate constant of 4-methoxyglucobrassicin in red cabbage was higher than that of glucobrassicin. The combination of both, positive and negative correlations of the degradation rate constants of glucobrassicin with metabolites illustrate the complexity of the GL degradation mechanism.

The QTLs of the degradation rate constants of glucoraphanin and glucobrassicin co-localise on linkage group C07 and the degradation rate constants of the two GLs are well correlated (Pearson correlation coefficient: 0.77). Therefore it was expected that some of the metabolites are associated to the degradation of both glucobrassicin and glucoraphanin. However, this was the case for two metabolites only (4-methoxyglucbrassicin, one not identified). The QTLs of four metabolites co-localised with both QTLs for the degradation rate constant of glucobrassicin and the QTLs of seven other yet unidentified metabolites associated with the degradation rate constant on linkage group C07. The co-localization of these metabolites with the two GLs in combination with the similar direction of their correlation coefficients suggests that similar groups of compounds are related to the GL degradation. On linkage group C01 the metabolite related to the degradation of glucoraphanin is positively correlated with the rate constant whereas the metabolites related to the degradation of glucobrassicin are negatively correlated; this observation requires further investigation.

QTLs of metabolites related to GL degradation were identified on linkage groups C01, C07 and C09. Although only a few of the associated metabolites have yet been identified, the markers underlying these loci can be used in marker-assisted breeding to produce new *Brassica* vegetable varieties with reduced GL degradation by heating during food processing.

5.5 Conclusion

In conclusion, the present study shows that untargeted metabolomics technologies, such as C18 reversed-phase LC-MS, combined with suitable statistical tools, such as random forest regression, are suitable to identify factors within the plant matrix that influence GL stability during heating. The use of a broader range of extraction solvents combined with different but complementary analytical platform might increase the explained variation of the GL degradation rate constants and identify different groups of metabolites associated with GL degradation rates. Our results of LC-MS profiling of crude methanol extracts from leaves of 85 DH lines from a *Brassica oleracea* DH population indicate that flavonols present in the same vegetable may increase the GL thermal degradation in the plant matrix, possibly by redox-cycling of endogenous iron molecules in the plant matrix. The fact that other metabolites present may retard the degradation of GLs illustrates the complexity of the vegetable system. Clearly, further research is needed to explain and control the thermal degradation of GLs in different vegetables. This approach is valuable for breeding new *Brassica* varieties with higher retention of GLs during food processing.

5.6 Acknowledgements

5

CHAPTER

General Discussion

6.1 Introduction

Phytochemicals in plant-based foods have been linked to a reduced incidence and progression of diseases [2, 3]. Glucosinolates (GLs) are phytochemicals that are typical for *Brassica* and other *Cruciferous* plants. The intake of GLs has been associated with lowered risks of several types of cancer and other diseases [15]. To reach a high level of GLs in the vegetables at the stage of consumption, research along the food chain aimed to maintain or increase the concentrations of certain GLs and to lower the losses during food processing; however, effects of the different steps in the food chain are mostly studied separately. In this thesis, an interdisciplinary approach, combining food science and plant science, was applied to explore the possibilities to retain GLs throughout the food chain. The objective of this thesis was to investigate genetic effects related to GL degradation during food processing in order to explore if food processing parameters can be used as phenotypic traits by plant breeders to breed new vegetable varieties with improved GL retention. Additionally, the identification of genetic regions associated with GL degradation provides a starting point to investigate molecular mechanisms causing GL losses. Several mechanisms have been described leading to GL reduction during food processing, one of them is thermal degradation, which is solely induced by heat and is the focus of this thesis. This discussion starts with a summary of the main findings, followed by methodological considerations and the discussion of the results. Finally, future prospects and main conclusions of the research are described.

6.2 Main findings

One challenge of integrating plant science and food science is the high number of samples to be analysed to apply quantitative genetics to technological traits, in particular when a kinetic approach is used. In **chapter 2** the analysis of GLs, as desulpho-GLs, was optimized to obtain reliable results using a high-throughput method. The effect of the sulphatase preparation on the accuracy of the enzymatic desulphation step in microtiter plates was tested. For the first time an inverse effect of the sulphatase concentration on the peak area of a GL, glucotropaeolin, which is often used as internal standard, was reported. Decreased peak areas of the internal standard led to a substantial overestimation of GL concentrations. We recommend the application of a purified sulphatase preparation to obtain good GL recoveries for a broad range of samples. This finding emphasizes the importance of having suitable analytical methods; in particular for the interdisciplinary

approach we advocate in this thesis, where actions in the various disciplines depend strongly on the analytical outcomes.

Kinetic modelling was applied to describe GL thermal degradation in a quantitative way in **chapter 3**. The first aim was to identify an appropriate model for the description of GL degradation in two genotypes and two different seasons. Model discrimination was performed for first- and n^{th}-order models and it was shown that a first order reaction was appropriate in both seasons and genotypes. The second aim was to investigate the effect of season and genotype on GL thermal degradation. No difference in degradation rate constants of structurally identical GLs was found between leaves of a broccoli and a Chinese kale genotype when grown in the same season. However, GL degradation rate constants were highly affected by the season (20% to 80% difference).

In **chapter 4**, the kinetic modelling approach was applied to a doubled haploid (DH) population, obtained by crossing the broccoli and Chinese kale genotypes investigated in chapter 3. GL degradation rate constants were determined for five aliphatic and two indolic GL and they varied by 3-fold throughout the population for most GLs. Quantitative trait loci (QTL) for the degradation rate constants were identified for the two GLs that occurred in all genotypes; two QTLs were identified for glucobrassicin (25% explained variation) and one for glucoraphanin (12% explained variation). A co-localization of the QTLs and correlations between the rate constants of different GLs indicate that similar factors in the plant matrix influence GL degradation. In addition to the genotypic effects, the growing conditions influence GL degradation as well.

An untargeted metabolomics approach was applied to test if metabolites are associated with GL thermal degradation rate constants in **chapter 5**. Out of 413 metabolites, 15 were significantly associated with the degradation rate constant of glucobrassicin, 6 with the degradation rate constant of glucoraphanin and 2 with both GL degradation rate constants. Among these 23 different metabolites, 3 were identified as flavonols (kaempferol-3-sophoroside-7-glucoside, quercetin-3-sophoroside-7-glucoside-2-caffeoyl, quercetin-3-glucoside-7-glucoside) and two as glucosinolates (4-methoxyglucobrassicin, gluconasturtiin). QTLs for two flavonols co-localise with the QTL for glucosinolate degradation in chapter 4.

6

6.3 Methodological considerations

6.3.1 Choice of plant population

Several types of plant populations have been described in literature, e.g. backcross populations, recombinant inbred line populations and doubled haploid (DH) populations [50]. For the work of this thesis a publically available DH population was used, which was developed by crossing a broccoli and a Chinese kale genotype [71]. The advantage of DH populations is that each DH line is homozygous so that genetically identical seeds can be produced to perform repeated phenotyping with identical genotypes. A disadvantage of DH populations is that they are derived from single gametes that they are induced to regenerate into diploid plants. Thus, a single meiosis per DH line occurred, which results in low recombination rates and consequently a genetic map with lower resolution compared to genetic maps for other types of populations. Recombinant inbred lines require, for example, self-pollination between six to eight generations resulting in more chromosomal recombination and higher resolution genetic maps compared to DH populations. As recombinant inbred lines are homozygous, they share that advantage with DH lines. However, for a bi-annual crop as *Brassica* up to 4 years are required to reach homozygous lines in these populations.

GL degradation of the parental lines of this DH population was investigated in chapter 3. Based on the results of Dekker *et al.* [35], who reported that GL thermal degradation differed in five *Brassica* vegetables, differences in GL thermal degradation between this Chinese kale and this broccoli genotype grown in the same season were expected; however, degradation of structurally identical GLs did not differ between the two parental genotypes. Despite similar degradation rate constants of the parental lines, the degradation rate constants segregated quantitatively throughout the DH population, which is a phenomenon called transgressive segregation (phenotypic values are beyond the range of the parents) and generally means that the trait is quantitatively inherited.

6.3.2 Experimental design

Genetic effects explained 12% and 25% of the variation of the degradation rate constants for the GLs glucoraphanin and glucobrassicin, respectively. The remaining variation of the rate constants is caused by environmental effects, experimental noise and minor genetic effects (QTL), which cannot be detected, due to the size (100 lines) of the used DH population. To reduce the experimental noise, five biological replicates were

grown randomised in the greenhouse and harvesting was performed in the mornings because variation in GL concentration throughout the day is highly significant [110]. Furthermore, the circadian clock influences the primary and secondary metabolism of plants [42, 111] which might influence GL thermal degradation. Due to the sample size, harvesting of the whole population had to be performed in five consecutive mornings where all biological replicates were harvested the same day and were subsequently pooled to prepare one homogenous sample for the GL degradation study. The assumption underlying the pooling was to prepare an average sample accounting for the variation in the five biological replicates and keep the number of samples manageable. Variation caused by the different harvesting days was shown to be not significant (chapter 4) for the degradation rate constants of most of the GLs. Harvesting one replicate per DH line per day, rather than harvesting all replicates the same day, performing the subsequent inactivation of myrosinase on each biological replicate separately and pool the material afterwards is an alternative to the chosen experimental setup, but it is more laborious.

Microwave treatment at high power for a short time has shown to fully inactivate myrosinase and retain GLs [78]. This treatment was chosen to study thermal degradation as the sole mechanism and exclude enzymatic GL degradation. The microwave treatment also alters other metabolites and proteins than myrosinase, which leads to different experimental conditions compared to common household/industrial preparation. Alternative ways to inactivate myrosinase could be adding hot water, adding enzyme inhibitors or using metabolically engineered plants that have negligible myrosinase activity. Hanschen *et al.* [31] added boiling water to dried broccoli sprout powder to inactivate myrosinase, but water was added by a factor five more compared to the natural amount of water occurring in broccoli. Competitive inhibitors of myrosinase have been described in literature with structural similarities to glucotropaeolin with 50% reduction of the enzyme activity at milli-molar and micro-molar level respectively [112, 113]. Transgenic *Brassica rapa* genotypes with ablated myrosin cells have been developed without affecting plant viability. Transgenic seedlings showed 84% less myrosinase activity compared to the wild type [114, 115], but GL profiles are altered and possibly other metabolites as well. All mentioned alternatives for the microwave treatment alter metabolite profiles. The least invasive method may be the enzyme inhibitors, though the inhibition is competitive and GL degradation may still take place. The microwave treatment is an additional heating step before the actual heat treatment and is, compared to the described alternatives, seen as the best option for this study,

6

although heat labile metabolites may be degraded. The effect of heating on the overall metabolite profile, as determined in chapter 5, was in some lines of this DH population three times smaller than the effect of genotype and environment.

To ensure that myrosinase was inactivated by the microwave treatment, a subset of the DH lines was tested for remaining myrosinase activity. The subset was selected based on the water loss during microwave treatment as a measure of the heating intensity. No myrosinase activity was detected during 40 min of incubation when 15% or more water evaporated during microwave treatment, which was the case for most of the DH lines (data not shown).

6.3.3 Kinetic modelling & quantitative trait loci mapping

Kinetic modelling was chosen as a tool to describe chemical changes in foods in a quantitative way [75]. Using degradation rate constants (k_d values) as phenotypic traits for the QTL mapping has the advantage that the overall reaction is characterised and includes more information than a single processing time point. Furthermore, the degradation rate constant is independent from the initial concentration, which resulted in different locations for the degradation QTLs compared to the QTLs for the initial GL concentration, illustrating that the identified genetic regions are true degradation QTLs (chapter 4).

For the degradation rate of glucobrassicin two QTLs were identified and for the degradation rate of glucoraphanin only one QTL was identified (chapter 4). Furthermore, more metabolites were associated with the degradation rates of glucobrassicin than for glucoraphanin (chapter 5). The degradation rate constants (k_d values) of glucobrassicin had lower standard deviations (8% on average over the whole population) than the rate constants of glucoraphanin (18% on average over the whole population, (from data in table 4-2, standard deviations not shown)). Less accurate k_d values of the glucoraphanin degradation, as indicated by the standard deviations, is one possible reason why less QTLs and less metabolites are associated with the degradation of glucoraphanin compared to glucobrassicin. Less accurate k_d values could be a result of a poorer fit of the first order model for glucoraphanin. The abundance of two GLs (4-methoxyglucobrassicin, gluconasturtiin) was associated with the degradation of glucoraphanin, whereas only one GL (4-methoxyglucobrassicin) was associated with the degradation of glucobrassicin (chapter 5, Table 5.2). Gluconasturtiin and 4-methoxyglucobrassicin degrade with increasing heating time, which suggests that the first order model needs to be corrected

for the degradation of the influencing compounds. The rank of importance to random forest model of the two GLs is higher for the glucoraphanin model (rank 4 & 7 of 8) than for glucobrassicin (rank 16 of 17). The higher rank of importance of GLs on the degradation rate constants of glucoraphanin than of glucobrassicin suggests that the first order model is appropriate for the degradation of glucobrassicin and less so for glucoraphanin. Regarding the residuals of the glucoraphanin degradation models, however, no clear trend was visible for DH lines with high 4-methoxyglucorbrassicin concentrations, which would indicate an inappropriate fit of the first order model. The first order model is an empirical model, which was applied since the molecular mechanism of thermal degradation is not known and selected based on the best fit. A change in the reaction conditions causes a change in the mechanism which will affect the order of a reaction.

Another factor that could affect the number of QTLs and the associated metabolites with the degradation of glucoraphanin is that small differences of the k_d values were found between the harvesting days, whereas no differences were found for glucobrassicin. ANOVA and subsequent post-hoc tests revealed significant differences between harvesting day one and harvesting day three for glucoraphanin, whereas the other days did not differ significantly from each other.

6.4 Discussion and interpretation of the results

6.4.1 Glucosinolate analysis

An official method for GL analysis was published in 1992 by the International Standard Organisation [73] with the aim to standardize the analyses in rapeseeds. This method has been adjusted to analyse other *Brassica* species and analysis of GLs as desulpho-GLs is still the most widely used [116]. The desulphation decreases the polarity of the GLs and hence improves chromatographic separation on reversed-phase columns [116]. The data presented in chapter 2 demonstrate that the analytical result strongly depends on the applied sulphatase concentration, on the chemical structure and concentration of the GLs in the sample. Furthermore, it has been shown that the sulphatase batch can influence the determined GL concentration as well. The lack of commercially available standards impedes the adjustments of methods and quality control and subsequently, can make the comparison of results obtained in different laboratories difficult.

The analysis of intact GLs is an alternative to avoid the desulphation step. Recent advances

6

in the development of column material allow the separation of the very polar intact GLs in short time [117]. The desulphation procedure on ion-exchange columns includes a concentration step of the extract, which is not required when highly sensitive detectors, such as QTOF-MS, are used. The use of these highly sensitive detectors, however, is more expensive than a common UV-detector [117]. Intact GLs are often detected with MS [118-121] which makes the identification of the GLs more reliable, but requires external standards as pure compounds for accurate quantification due to the strong influence of the matrix on the ionization [70]. An advantage of the desulpho-method combined with UV-detection is the quantification via established relative response factors (RRF) which recently have been summarized by Clarke [70]. The analysis of desulpho-GLs with UV- and MS-detection has been suggested for quantification and structural identification [116], but this method still requires a reliable desulphation procedure. The analysis of intact and desulpho-GLs have both strengths and weaknesses, but the lack of commercially available standards hinders the quantification and reliability for both methods.

6.4.2 Breeding for food quality traits

Up to now, only few studies in literature have combined breeding and food processing. In the following paragraphs existing genetic studies on several food quality traits are reviewed, with and without prior knowledge on mechanisms underlying the trait.

Genetic studies with prior knowledge of the trait

An example of a breeding strategy for affecting a quality attribute of a processed food product is the breeding for improving the colour of potato chips. Li *et al.* [122] identified QTLs for the colour of potato chips after frying. A too dark colour of potato chips, a negative quality attribute, is a result of the Maillard reaction, a reaction between free amino groups and reducing sugars occurring at high temperatures. A high amount of reducing sugars results in dark colour during frying and formation of the toxicologically suspect compound acrylamide, whereas a high starch content and a low amount of reducing sugars results in lighter colours. For the genotyping, DNA markers were partly designed based on known genes involved in sugar and starch metabolism. The authors identified several QTLs for chip quality and most significant QTLs were detected at genetic loci that encode enzymes involved in carbohydrate metabolism. Another example is given by Simons *et al.* [123] who studied the genetic effects of bread making quality traits in spring wheat using a segregating population. Gluten, consisting of the

6

main two proteins glutenin and gliadin, affects the rheological (viscoelastic) properties of dough and hence plays a major role in dough strength and baking performance. The authors complemented the genetic map with molecular markers based on the DNA sequences of high and low molecular weight glutenins. A major QTL, explaining 47 to 63% of the phenotypic variation for three dough mixing time parameters and bread making performance, was identified at the genomic region where also the high molecular weight glutenin gene mapped.

Both studies clearly illustrate the advantage of the use of mechanistic knowledge from food scientists to study the genetics related to food quality attributes and identify allelic variation of the genes causally related to the variation in these attributes.

Genetic studies without prior knowledge of the trait
In contrast to the studies presented in the previous section, the following examples of quantitative food processing traits have in common that the molecular mechanisms affecting the trait are not known. Dong *et al.* [116] mapped QTLs for the popping characteristics of popcorn in a segregating population that was evaluated for three different food processing traits, popping volume, popping fold and popping rate and identified QTLs for all three popping characteristics. The explained phenotypic variation by each QTL varied from 6.2% to 29.5%, depending on the trait and the environment in which the population was grown. Furthermore, the authors identified a significant genotype × environment interaction, showing that the QTLs are dependent on the environment.

Wright & Kelly [94] studied the canning quality of black beans (*Phaseolus vulgaris L.*), which is determined by a strong colour, intact beans and a desirable texture after canning. The final colour might be influenced by the initial anthocyanin levels in the black bean seed coat and the amount of leaching into the cooking water; however, the molecular mechanisms determining the leaching rate are unknown. A segregating black bean population, derived from a cross of a high yielding parent with low colour retention and a parent with acceptable colour retention, was evaluated in four years. Seven QTLs related to canned bean colour were identified, which shows the complexity of the trait. For most traits QTLs could be identified, despite a strong interaction between environment and traits.

6

QTL mapping for the rheological properties during dough making from barley flour was performed by Wang *et al*. [95] using a segregating population. A shorter time to peak viscosity and lower pasting temperature are favourable for malting and food processing. Lower pasting temperatures are related to starches with low amylose content, but also other components, like lipids, protein and beta-glucan, affect flour pasting properties. The study showed that flour pasting quality traits are quantitatively inherited, with many QTLs explaining between 4% and 15% of the phenotypic variation. Flour pasting quality is also affected by the growth environment as the identified sets of QTLs were depending on the growing years and environments.

The examples on traits without knowledge of the exact molecular mechanisms show that breeding for improvement of food processing traits does not necessarily require the detailed understanding of the biochemical pathways. The position of the QTL in the genome can assist in finding causal genes in the QTL region for species with genome sequence information and may even help to identify underlying mechanisms occurring during food processing.

6.4.3 Phytochemical concentration as a food quality attribute

The presented studies on breeding for food quality traits cover several aspects of food quality, mainly related to colour, texture and taste. Food quality can be generally described as "satisfaction of consumer expectations" and food quality attributes as product properties that are relevant to determine the quality [80]. Food quality can be decomposed into extrinsic and intrinsic quality attributes: extrinsic attributes are linked to the product but are not a property itself, e.g. price, acceptability for religious reasons, the way of production, brands, while intrinsic quality attributes are the properties of the product itself, e.g. texture, taste, nutritional value, chemical and microbiological safety [80]. A selection of intrinsic food quality attributes is illustrated in Figure 6.1, which are the focus of food scientists since they are measurable attributes [80]. Phytochemical content, in particular GL content, is the focus of this thesis, which belongs to the quality attribute "nutritional value".
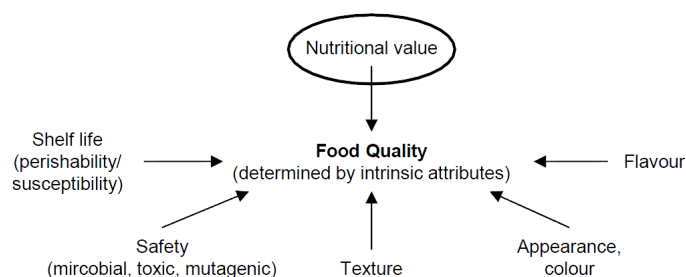
Figure 6.1: Intrinsic food quality attributes that are most important to food scientists [80].

6.4.4 Breeding for increased glucosinolate retention during food processing

Glucosinolate thermal degradation

The chemical structure of a GL, the time-temperature profile applied, the water content and the pH of the reaction environment influence GL thermal degradation [32, 36, 84]. Additionally, ferrous ions (Fe^{2+}) mediate thermal degradation [33]. The presence of both vitamin C and ferrous ions in a buffer solution even further enhanced the thermal degradation rate due to redox-cycling of ferric ions (Fe^{3+}) by vitamin C; however, the degradation in the buffer solution was lower compared to a broccoli sprout matrix [34]. Differences in thermal degradation speed of chemically identical GLs have been described in five different *Brassica* vegetables [35] and three heading cabbage cultivars [124]. The explanation for the observed differences is unknown since these vegetables were not only different in their genetic background (cultivar/type) but also in stage of maturity, organs processed (e.g. broccoli heads, cabbage leaves), growing locations and storage conditions.

To our knowledge, we are the first group investigating genetic effects of GL thermal degradation during food processing. It was shown that indeed multiple QTLs are associated with GL thermal degradation, explaining a part of the observed variation between the different genotypes. Furthermore, GL thermal degradation was influenced by the growing season, which is a common phenomenon of quantitatively inherited traits (chapter 3 & 4). To apply the identified QTLs for marker-assisted selection (MAS) the QTL must be stable across environments and time [125], which cannot be predicted from the current experimental data. The effect size of a QTL is another criterion to determine if MAS will be efficient for breeders. Hospital [125] concludes that there is no clear limit for the size of QTLs to be used for MAS and that some companies claim to use QTLs

explaining less than 10% of phenotypic variation for MAS. Each of the QTLs identified for GL degradation explained more than 10% of the variation and are according to Hospital suitable for MAS. For traits with low heritability and small effects, as it is the case for GL degradation, MAS would be beneficial because phenotypic effects might be hard to detect each season, but MAS is effective to introgress previously determined QTLs in elite germplasm. In addition, QTL × QTL and QTL × genetic background interaction have to be investigated to predict the success of MAS to avoid that the effect of the QTL is lost in another genetic background [125].

Leaching of glucosinolates

Leaching of GLs into the cooking water causes the majority of GL losses during boiling and is dependent on the type of vegetable, the size of (cut) vegetable pieces, the ratio water/vegetable, the boiling time and method [23, 24]. Many studies investigated the effect of boiling on retained GL levels; however, results across studies are difficult to compare due to the wide range of boiling methods applied, especially vegetable/water ratios. Also many studies do not study specifically leaching losses and only report the total loss observed due to all possible mechanisms. The studies that report quantitative information on leaching as a separate mechanism during processing are described below and possible phenotypic variation that could be associated with genetic variation is presented.

A higher extent of leaching was reported in a leafy *Brassica napus* type compared to four *Brassica oleracea* types, which is probably related to a thinner wax layer of the *B. napus* type [124]. The structure of the cabbage (kale or heading type, *B. oleracea*) revealed only small differences in leaching, but an increased surface by shredding promoted leaching [124]. In two broccoli cultivars, short time boiling significantly reduced the concentrations of glucoraphanin (aliphatic GL) and glucobrassicin (indolic GL), though the retention of both GLs was slightly higher in the cultivar Marathon compared to Booster [126]. For five different vegetables differences of 30% in leaching are reported, however boiling times varied slightly between the vegetables [127]. In contrast, Francisco *et al.* [128] did not find significant variety x cooking method interactions for any GL in five varieties of turnips greens (leaves) and turnip tops (young sprouting shoots) (*B. rapa*). Furthermore, for five different cauliflower cultivars boiling and blanching showed differences in GL losses only for some cultivars and some GLs [129].

The presented studies on leaching of GLs show variation among different vegetables and

to a lower extent between cultivars and indicate a potential to breed for less leaching of GLs; however, systematic and mechanistic studies are lacking. Hence, systematic studies with well-defined plant material, grown under the same conditions and several processing time points are needed to evaluate potential variation. In order to identify QTLs for leaching in future research, it is essential that the leaching losses of lines from a segregating population are quantified in a uniform way. A good way to do this is by a kinetic modelling approach as described by Sarvan *et al.* [44]. In short, this method is using the GL content of both the water and vegetable phase at different processing times (and temperatures if appropriate) at a known water/vegetable ratio. By fitting the kinetic model to the data, a rate constant for leaching for each vegetable is obtained. These rate constants can subsequently be used for QTL analysis in a similar way as described in our study of thermal degradation. If QTLs can be identified, breeding for vegetables with lower leaching losses can be done in an efficient way.

Myrosinase

The activity of the enzyme myrosinase is of minor influence during food processing, since it is inactivated below common cooking temperatures (chapter 1). However, if *Brassica* vegetables or sprouts are consumed raw, myrosinase is active and hydrolysis products can be formed in the mouth leading to 3 - 10 times higher concentrations of beneficial isothiocyanates compared to the ingestion of cooked vegetables [19-21]. Myrosinase associated proteins are cofactors that influence the type of breakdown products formed, especially epithiospecifier protein (ESP) directs the reaction towards the formation of nitriles, which are not associated with health benefits, on the expense of isothiocyanate formation. Not all *Brassica* species contain ESP; it was detected in turnip, Brussels sprouts and white cabbage, and absent in horseradish seeds and mustard seeds [130]. Differences in isothiocyanate and nitrile production between different cultivars were also reported in broccoli [131]. Breeding for e.g. broccoli sprouts with low amounts or absence of ESP is a potential breeding goal to increase the isothiocyanate production.

Another way to increase isothiocyanate production is mild heat treatment of e.g. broccoli, since ESP is more heat sensitive than myrosinase, leading to increased sulforaphane (isothiocyanate of glucoraphanin) production after mild heat treatments where myrosinase was still active, but ESP inactivated. Optimal steaming times (1 min to 3 min) to reach highest sulforaphane concentrations differed per cultivar, either caused by different amounts of myrosinase or by differences in structure of the broccoli

6

head resulting in different heat transfer [131]. These results suggest that some cultivars can retain more myrosinase activity after certain steaming times than others, which could be a potential breeding trait and mechanism to investigate. Furthermore, the selection of specific cultivars for the production of vegetable products could enhance product properties. For example, Oliviero *et al.* [43] suggested steam blanching of broccoli for 3 min, followed by a novel adsorption drying technique allowing lower drying temperatures (40 - 50°C) than conventional drying techniques to obtain a dried broccoli product with active myrosinase. The selection of broccoli cultivars with higher isothiocyanate production under these specific processing conditions could increase health promoting effects and the quality of the broccoli product.

Both genetic and environmental influences on the level of health promoting compounds formed during vegetable processing by the GL-myrosinase-system are illustrated in Figure 6.2. This summary can be used to identify factors influencing the different degradation mechanisms and study their genetic effects with the aim to reach a high content of health promoting compounds before the product is consumed.

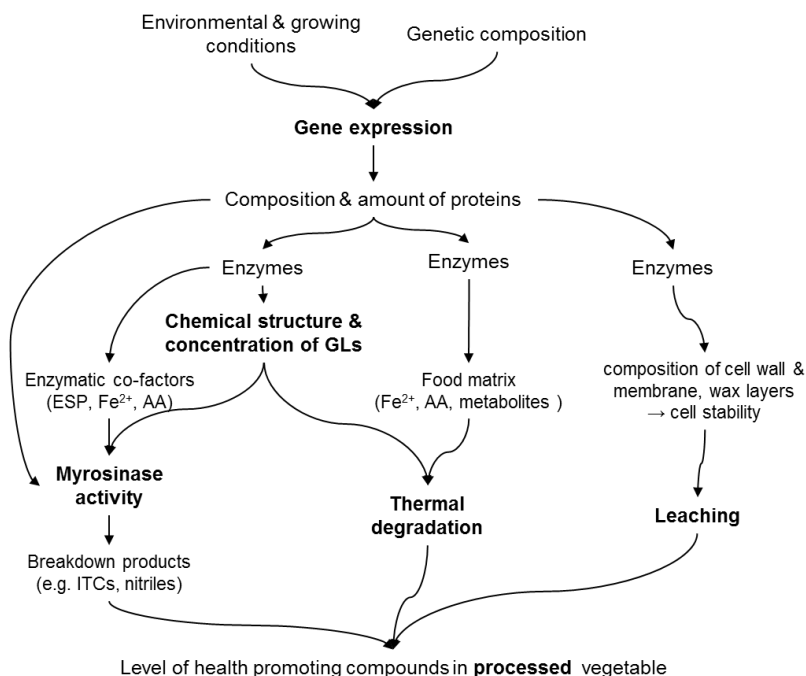

Figure 6.2: Possible genetic and environmental effects on factors influencing glucosinolate degradation during food processing, GLs = glucosinolates, ITC = isothiocyanates, MYR = myrosinase, ESP = epithiospecifier protein, AA = ascorbic acid

6.4.5 Challenges of breeding for vegetables with improved food processing properties

To study genetic effects of food processing traits, as suggested in the previous paragraphs, it is essential to quantitatively assess the processing properties for a large number of individuals in segregating populations. To obtain a population with substantial variation in the trait of interest, vegetable accessions and cultivars exhibiting a large difference in the trait can be crossed to develop segregating populations for genetic mapping studies. Various approaches are available to develop segregating populations.

Following the development of the populations, the major challenge for studying food processing traits is the number of samples involved. Farnham *et al.* [132] stated that the challenge in enhancing phytonutrient content is that a plant breeder needs to analyse a large number of plant samples and that the collaborators responsible for the characterization of phytonutrient content, such as nutritionists or biochemists, do not always understand the nature of plant breeding and the necessity of large populations in a breeding effort. Combining the requirement of large populations (100 to 200 individuals) with mechanistic studies of food processing, where several time points and replicates for processing are needed, will result in at least 1000 samples (>10 data points × >100 individual plants).

1. Development of a high-throughput method to determine concentration of phytochemicals

2. Screening of cultivars for differences in phytochemical retention during food processing

3. Development of a segregating population by crossing cultivars with maximum differences in phytochemical retention

4. Genetic map construction

5. Growing of the segregating population under certain conditions

6. Analysis of phtytochemical concentrations after several processing times & kinetic modelling of the results per individual line
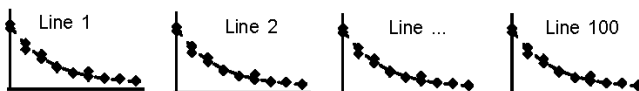
7. QTL mapping (with genotypic data and rate constant from kinetic modelling)

8. Test for environmental effects by repeating steps 5 -7 in a different season/year

Figure 6.3: Workflow for studying genetic regulation of phytochemical retention during food processing.

Studying the gene x environment interactions of food processing traits is another challenge leading to a further increase in the number of experiments and analyses as shown for the cooking time in beans [133] and for GL thermal degradation (chapter 3 & 4). Recent advances in development of high-throughput analytical methodologies will facilitate the analysis of such high sample numbers. A possible workflow for the identification of genetic loci regulating phytochemical retention is presented in Figure 6.3.

6.4.6 Metabolites influencing glucosinolate degradation

Up to now the full sequence of the *Brassica oleracea* genome has not been published, which hinders the identification of genes underlying the detected QTLs for GL degradation. Furthermore, fine mapping of the DH population would be required to obtain smaller genetic regions with fewer possible genes underlying the QTLs, since little prior knowledge about GL degradation exists. As an alternative approach, an untargeted metabolomics approach combined with random forest regression was applied to identify metabolites associated with GL thermal degradation to form new hypotheses about the underlying mechanism. Two GLs, three flavonols and 18 not annotated metabolites were significantly associated with thermal degradation rate constants. The two GLs possibly degrade faster and hence can reduce the degradation of glucobrassicin by a competitive mechanism. The flavonols possibly enhance degradation of glucobrassicin by the reduction of ferric ions (Fe^{3+}) to ferrous ions (Fe^{2+}) which mediates the degradation of GL under the formation of nitriles and ferric ions ([33], chapter 5). Vitamin C was identified to enhance the degradation of sinigrin in a buffered solution in the presence of ferrous ions due to the redox-cycling ability of vitamin C [34]. Vitamin C was not associated with GL thermal degradation in the experiments in this thesis because the LC-MS method applied did not detect vitamin C. Furthermore, vitamin C might have been degraded during the microwave treatment due to its susceptibility to heat.

Since ferrous ions are the actual compounds mediating the GL thermal degradation, their initial concentration may play a role in GL thermal degradation [33, 34]. The iron uptake in *Brassica* is dependent on genotype when grown in one location [134] and dependent on the vegetable, soil and climate conditions when grown in different locations [135]. Furthermore, a significant genotype × environment interaction for iron levels in broccoli was shown [136]. The alteration of iron levels in *Brassica* vegetables might affect GL thermal degradation; however, the reduction of iron levels would lead to lower mineral intake. Furthermore, an increased level of vitamin C or other metabolites

with reducing properties may increase GL thermal degradation even at low ferrous ion concentrations.

Until now, the enhanced GL thermal degradation in the presence of vitamin C and ferrous ions has only been shown for the aliphatic GL sinigrin [34] and in chapter 5 the three flavonols were only associated with glucobrassicin degradation and not with glucoraphanin. The correlation of the degradation rate constants of the different GLs (chapter 4) suggests that similar metabolites influence GL thermal degradation. The proposed mechanism of thermal GL degradation mediated by ferrous ions of Bellostas *et al.* [33] does not suggest an influence of the side chain on the reaction, because the ligands are expected to be the sulphur in the thioglucose and the oxygen in the sulphate group (Figure 5.6). Targeted analysis and model systems are required to confirm the findings that some flavonols influence glucobrassicin degradation. More investigations are needed, to test if the ferrous-ion-mediated mechanism shows the same reactivity towards the different GLs. Other factors, such as water content [84] and pH [36], influence GL thermal degradation as well. Until now it is unknown whether these factors affect the ferric-ion-mediated degradation or whether another unknown mechanism exists.

The QTL analysis of metabolites associated with the degradation rate constants of glucoraphanin and glucobrassicin revealed additional QTLs on linkage group C01 where no QTL for the degradation rate constants had been identified (chapter 5). Furthermore, the explained variation of each metabolite QTL is up to three times higher than the explained variation of the degradation rate QTLs (Figure 5.5), which is likely due to less experimental noise. For the metabolite QTLs intensity values were used directly, whereas the degradation rate constants were determined by kinetic modelling after several heating time points. A larger population size increases the power of the detected QTLs, possibly a QTL on linkage group C01 could have been detected for the degradation rate constants in addition to the QTL on linkage group C07 and C09. Sequence information of *Brassica oleracea* may help in the future to identify the genes underlying metabolite and degradation QTLs. The combination of sequence information and the identity of the metabolites allow a more specific selection of candidate genes for GL thermal degradation.

6

6.4.7 Which doses of glucosinolates are healthy?

The dose-response curve of phytochemicals on the risk of chronic diseases is not linear, it follows generally a U-shaped curve with low levels causing an increased risk, an optimal protective amount and excessive levels causing risk again [137]. The development of vegetables with increased GL concentrations, such as the high-glucoraphanin broccoli *Beneforte* [46], and breeding for high GL retention during food processing aim at a higher intake of GLs and hence requires the estimation of potential risks.

Epidemiological studies associate a decreased risk of cancer with high consumption of *Brassica* vegetables [138], but a good estimation of GL intake of the populations is difficult, as the huge level of variation throughout the food chain and the cooking method can lower GL concentrations in the consumed product substantially [23]. The physiological effect of each GL depends on its chemical structure, for example the hydrolysis product of progoitrin is goitrogenic at high concentrations which was not found for most other GLs [139]. Intake of broccoli sprouts extracts over seven days with varying glucoraphanin (24 and 96 mg/day) and sulforaphane (10 mg/day) concentrations did not show toxic side effects of liver and thyroid [[140] cited in [18]]. Contrarily, a limited number of *in vitro* and animal studies indicate a mutagenic potential of broccoli extracts, which might be caused by neoglucobrassicin in the presence of myrosinase. In humans, however, genotoxic effects have not been reported after consumption of broccoli [10]. Furthermore, *in vitro* and experimental animal studies indicate that certain nitriles might cause damage in specific organs when extremely high concentrations were applied, which exceed the intake of humans by far [10].

Current studies don't indicate that an increased intake of GLs will lead to toxic effects in humans, especially with focus on increased glucoraphanin concentrations. The majority of *Brassica* vegetables is consumed cooked where myrosinase has been inactivated; hence the microflora will convert GLs into bioactive isothiocyanates with a considerably lower conversion rate. Latté *et al.* [10] conclude from their risk-benefit assessments that benefits of broccoli consumption in modest quantities and processed form outweigh potential risks. Other preparations, such as broccoli-based dietary supplements, diets with extraordinary high intake and consumption as a raw vegetable, require further investigations to evaluate benefits and risks in the future.

6.5 Recommendations and implications

In this thesis genetic loci for GL thermal degradation have been identified and show that breeding for GL retention during processing is possible. Nevertheless, it was also revealed that GL thermal degradation is strongly influenced by the environment. Follow-up studies are needed with material grown in different seasons and biological replicates to test if the identified QTLs are stable across different environments. Data about biological variation may give further insight in the relative importance of the genetic effects. The degradation of chemically identical GLs in parental lines of the investigated DH population did not differ when grown in the same season. Systematic studies of different cultivars and genotypes, grown at the same location and environment will reveal the variation of GL degradation in commonly consumed *Brassica* vegetables. The genotypes with large variation in GL degradation could be used to develop new segregating populations, which may exhibit larger variation than the currently used population and may lead to larger genetic effects of GL thermal degradation than observed with this DH population. A bigger population size (>100 lines) could potentially lead to the identification of more QTLs. Fine mapping of the DH population used in this thesis could result in the identification of smaller genetic regions, which can be used to search for functional genes when the *Brassica oleracea* genome will be published.

Studies on thermal degradation products in water cress seeds and broccoli sprouts report nitriles, which don't show health promoting effects at dietary levels, as degradation products [7, 14, 36]. However, studies in vegetables at commercial maturity are still lacking to confirm nitriles as degradation product, but from this data GL thermal degradation can be regarded as an undesired mechanism during food processing. Related to the mechanism of thermal degradation by ferric ions, reconstitution experiments with flavonols are required to prove the findings in chapter 5 that three flavonols are positively correlated with the degradation rate constants of glucobrassicin. Furthermore, the use of different metabolomics platforms, such as GC-MS and NMR, and different extraction solvents to cover a wider range of polarities will provide more information on which classes of metabolites influence GL thermal degradation. Once metabolites that affect GL thermal degradation are identified, screening and breeding for those metabolites will be facilitated.

Next to thermal degradation, leaching of GLs into the cooking water is a major cause for GL losses during food processing. More mechanistic and systematic studies are needed

6

to explore possibilities to breed for leaching of GLs into the cooking water.

The combination of two disciplines is a promising, but also challenging approach for the future. Possible benefits for plant breeders and food scientists are summarized in Table 6.1. A higher extent of collaboration may reduce variability of plant based products along the food chain in the future. One big challenge is the sample size required for genetic studies combined with phenotyping for food quality traits, but the development of high-throughput analysis techniques will certainly help in achieving the goal. Studying food processing traits in a high number of genotypes allows the use of advanced statistical methods to identify new mechanisms and hypotheses using untargeted methods, as shown in chapter 5. Targeted methods are still required to prove the new hypotheses.

Table 6.1: Possible benefits from the collaboration of food scientists and plant breeders

Benefits of the collaboration for:

a) Food Scientists:

✓ Insight in genetic & environmental variation of quality attributes
✓ Tools to dissect genetic and environmental variation
✓ Protocols for optimal generation of plant material
✓ Quantitative trait loci for identification of causal genes that may accelerate research about mechanisms during food processing

b) Plant Breeders:

✓ Insight in the food chain about quality changes till product reaches consumer
✓ Insight in mechanisms taking place during food processing → may lead to more defined genetic regions
✓ Protocols for mechanistic phentoyping procedures for food processing traits

Breeding for food processing traits is a novel approach to improve food quality. The work in this thesis and some literature examples show that it can be a valuable approach to improve food quality in the future. The examples reported in literature aim mostly at food quality attributes that are directly perceived by the consumer, such as colour and texture. Phytochemical concentrations are not directly perceived by the consumer, unless they are related to colour, e.g. carotenoids, which may complicate marketing compared to an improved colour or texture. Vitamins and other micronutrients are known for a long time to be essential to our body [137], hence the consumer acceptance and willingness to buy a higher priced product might be higher than for a product

high in phytochemicals, from which health beneficial effects have been established only recently. The general approach to use specific cultivars to increase food quality attributes and to breed for improved quality attributes has the potential to improve food quality, but it is still at the beginning.

6.6 Main conclusions

The research conducted in this thesis demonstrates that GL thermal degradation is partly genetically regulated and the identified QTLs may be used in future marker-assisted selection (MAS) to breed for vegetables with improved GL retention during food processing. Furthermore, environmental factors, such as season and growing year, influence GL thermal degradation. However, stability of the QTLs across different seasons and growing locations is a requirement for efficient MAS.

The identified genetic regions associated with GL thermal degradation were not used to scan for orthologous genes with an assigned function in *Arabidopsis thaliana* because the identified regions are too large and no prior knowledge could help to assign candidate genes. Instead, an untargeted metabolomics approach was applied to investigate causes of the variation in GL thermal degradation. Metabolites associated with GL thermal degradation were identified, which partly explain the variation in the degradation rate constants throughout the population. The results indicate that three flavonols may increase GL thermal degradation and two GLs, other than glucoraphanin and glucobrassicin, may decrease GL thermal degradation in the matrix. Eighteen not-annotated metabolites were associated with GL thermal degradation, some with positive and some with negative correlations, which illustrates the complexity of factors influencing this reaction. Future research to confirm these findings and annotate unknown metabolites is required.

The findings provide necessary information to breed for vegetables with increased GL retention during food processing, but clearly more studies are required to reach that goal in the future. Furthermore, a methodology to identify metabolic factors causing the variation of GL thermal degradation in different vegetables was applied successfully. The combination of metabolite identities with genetic loci associated with GL thermal degradation and future sequence information of *Brassica* will help to identify candidate genes and molecular mechanisms for GL thermal degradation.

In order to improve specific quality attributes of plant foods, breeding for quantitative food processing traits is a promising and challenging approach. As shown in this chapter, it has potential in the improvement of the nutritional quality of food products by combining the disciplines food science and plant science to select and breed for varieties with not only higher initial amounts of phytochemicals but also with a high retention during processing.

6

REFERENCES

1. WHO. Diet, nutrition and the prevention of chronic diseases. Report of a joint Food and Agriculture Organization/WHO Expert Consultation. T*echnical Report Series*, WHO, **2003**, No. 916, Geneva.

2. Liu, R.H. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *American Journal of Clinical Nutrition*, **2003**, 78 (3 SUPPL.): p. 517S-520S.

3. Martin, C. The interface between plant metabolic engineering and human health. *Current Opinion in Biotechnology,* **2012**, 24 (2): p. 344-353.

4. Steinmetz, K.A.; Potter, J.D. Vegetables, fruit, and cancer prevention: a review. *Journal of the American Dietetic Association*, **1996**, 96 (10): p. 1027-1039.

5. Mithen, R.F.; Dekker, M.; Verkerk, R.; Rabot, S.; Johnson, I.T. The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. *Journal of the Science of Food and Agriculture*, **2000**, 80 (7): p. 967-984.

6. Verkerk, R.; Schreiner, M.; Krumbein, A.; Ciska, E.; Holst, B.; Rowland, I.; De Schrijver, R.; Hansen, M.; Gerhäuser, C.; Mithen, R. Glucosinolates in *Brassica* vegetables: The influence of the food supply chain on intake, bioavailability and human health. *Molecular Nutrition & Food Research*, **2009**, 53: p. 219-265.

7. Bones, A.M.; Rossiter, J.T. The enzymic and chemically induced decomposition of glucosinolates. *Phytochemistry*, **2006**, 67 (11): p. 1053-1067.

8. Fahey, J.W.; Zalcmann, A.T.; Talalay, P. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry, 2001, 56 (1): p. 5-51.

9. Kissen, R.; Rossiter, J.T.; Bones, A.M. The 'mustard oil bomb': Not so easy to assemble?! Localization, expression and distribution of the components of the myrosinase enzyme system. *Phytochemistry Reviews*, **2009**, 8 (1): p. 69-86.

10. Latte, K.P.; Appel, K.E.; Lampen, A. Health benefits and possible risks of broccoli - An overview. *Food and Chemical Toxicology*, **2011**, 49 (12): p. 3287-3309.

11. Textor, S.; Gershenzon, J. Herbivore induction of the glucosinolate-myrosinase defense system: major trends, biochemical bases and ecological significance. *Phytochemistry Reviews*, **2009**, 8 (1): p. 149-170.

12. Williams, D.J.; Pun, S. Glucosinolates in *Brassica* vegetables: role in bitterness and

hence significance. *Food Australia*, **2011**, 63 (9): p. 407-412.

13. Schonhof, I.; Krumbein, A.; Brückner, B. Genotypic effects on glucosinolates and sensory properties of broccoli and cauliflower. *Nahrung - Food*, **2004**, 48 (1): p. 25-33.

14. Williams, D.J.; Critchley, C.; Pun, S.; Chaliha, M.; O'Hare, T.J. Differing mechanisms of simple nitrile formation on glucosinolate degradation in *Lepidium sativum* and *Nasturtium officinale* seeds. *Phytochemistry*, **2009**, 70 (11-12): p. 1401-1409.

15. Traka, M.; Mithen, R. Glucosinolates, isothiocyanates and human health. *Phytochemistry Reviews*, **2009**, 8 (1): p. 269-282.

16. Jeffery, E.H.; Araya, M. Physiological effects of broccoli consumption. *Phytochemistry Reviews,* **2009**, 8 (1): p. 283-298.

17. Verhoeven, D.T.H.; Verhagen, H.; Goldbohm, R.A.; van den Brandt, P.A.; van Poppel, G. A review of mechanisms underlying anticarcinogenicity by *Brassica* vegetables. *Chemico-Biological Interactions*, **1997**, 103 (2): p. 79-129.

18. Herr, I.; Büchler, M.W. Dietary constituents of broccoli and other Cruciferous vegetables: Implications for prevention and therapy of cancer. *Cancer Treatment Reviews*, **2010**, 36 (5): p. 377-383.

19. Shapiro, T.A.; Fahey, J.W.; Wade, K.L.; Stephenson, K.K.; Talalay, P. Human metabolism and excretion of cancer chemoprotective glucosinolates and isothiocyanates of Cruciferous vegetables. *Cancer Epidemiology Biomarkers & Prevention*, **1998**, 7 (12): p. 1091-1100.

20. Conaway, C.C.; Getahun, S.M.; Liebes, L.L.; Pusateri, D.J.; Topham, D.K.W.; Botero-Omary, M.; Chung, F.L. Disposition of glucosinolates and sulforaphane in humans after ingestion of steamed and fresh broccoli. *Nutrition and Cancer*, **2000**, 38 (2): p. 168-178.

21. Vermeulen, M.; Klöpping-Ketelaars, I.W.A.A.; Van Den Berg, R.; Vaes, W.H.J. Bioavailability and kinetics of sulforaphane in humans after consumption of cooked versus raw broccoli. *Journal of Agricultural and Food Chemistry*, **2008**, 56 (22): p. 10505-10509.

22. Fahey, J.W.; Wehage, S.L.; Holtzclaw, W.D.; Kensler, T.W.; Egner, P.A.; Shapiro, T.A.; Talalay, P. Protection of humans by plant glucosinolates: efficiency of conversion

of glucosinolates to isothiocyanates by the gastrointestinal microflora. *Cancer Prevention Research*, **2012**, 5 (4): p. 603-611.

23. Dekker, M.; Verkerk, R.; Jongen, W.M.F. Predictive modeling of health aspects in the food production chain: A case study on glucosinolates in cabbage. *Trends in Food Science & Technology*, **2000**, 11 (4-5): p. 174-181.

24. Nugrahedi, P.Y.; Verkerk, R.; Widianarko, B.; Dekker, M. A mechanistic perspective on process induced changes in glucosinolate content in *Brassica* vegetables: a review. *Critical Reviews in Food Science and Nutrition*, **in press**.

25. Van Eylen, D.; Indrawati; Hendrickx, M.; Van Loey, A. Temperature and pressure stability of mustard seed (*Sinapis alba L.*) myrosinase. *Food Chemistry*, **2006**, 97 (2): p. 263-271.

26. Van Eylen, D.; Oey, I.; Hendrickx, M.; Van Loey, A. Kinetics of the stability of broccoli (*Brassica oleracea* cv. *italica*) myrosinase and isothiocyanates in broccoli juice during pressure/temperature treatments. *Journal of Agricultural and Food Chemistry,* **2007**, 55 (6): p. 2163-2170.

27. Ludikhuyze, L.; Ooms, V.; Weemaes, C.; Hendrickx, M. Kinetic study of the irreversible thermal and pressure inactivation of myrosinase from broccoli (*Brassica oleracea L-Cv. italica*). *Journal of Agricultural and Food Chemistry*, **1999**, 47 (5): p. 1794-1800.

28. Ghawi, S.K.; Methven, L.; Rastall, R.A.; Niranjan, K. Thermal and high hydrostatic pressure inactivation of myrosinase from green cabbage: a kinetic study. *Food Chemistry,* **2012**, 131 (4): p. 1240-1247.

29. Rungapamestry, V.; Duncan, A.J.; Fuller, Z.; Ratcliffe, B. Changes in glucosinolate concentrations, myrosinase activity, and production of metabolites of glucosinolates in cabbage (*Brassica oleracea* var. *capitata*) cooked for different durations. *Journal of Agricultural and Food Chemistry*, **2006**, 54 (20): p. 7628-7634.

30. Verkerk, R.; Knol, J.J.; Dekker, M. The effect of steaming on the glucosinolate content in broccoli. *Acta Horticulture (ISHS)*, **2010**, 867: p. 37-46.

31. Hanschen, F.S.; Rohn, S.; Mewis, I.; Schreiner, M.; Kroh, L.W. Influence of the chemical structure on the thermal degradation of the glucosinolates in broccoli sprouts. *Food Chemistry*, **2012**, 130 (1): p. 1-8.

32. Oerlemans, K.; Barrett, D.M.; Suades, C.B.; Verkerk, R.; Dekker, M. Thermal degradation of glucosinolates in red cabbage. *Food Chemistry*, **2006**, 95 (1): p. 19-29.

33. Bellostas, N.; Sorensen, A.D.; Sorensen, J.C.; Sorensen, H. Fe^{2+}-catalyzed formation of nitriles and thionamides from intact glucosinolates. *Journal of Natural Products,* **2008**, 71 (1): p. 76-80.

34. Hanschen, F.S.; Bauer, A.; Mewis, I.; Keil, C.; Schreiner, M.; Rohn, S.; Kroh, L.W. Thermally induced degradation of aliphatic glucosinolates: identification of intermediary breakdown products and proposed degradation pathways. *Journal of Agricultural and Food Chemistry*, **2012**, 60 (39): p. 9890-9899.

35. Dekker, M.; Hennig, K.; Verkerk, R. Differences in thermal stability of glucosinolates in five *Brassica* vegetables. *Czech Journal of Food Sciences*, **2009**, 27 (special issue): p. 85-88.

36. Hanschen, F.S.; Platz, S.; Mewis, I.; Schreiner, M.; Rohn, S.; Kroh, L.W. Thermally induced degradation of sulfur-containing aliphatic glucosinolates in broccoli sprouts (*Brassica oleracea* var. *italica*) and model systems. *Journal of Agricultural and Food Chemistry*, **2012**, 60 (9): p. 2231-2241.

37. Charron, C.S.; Saxton, A.M.; Sams, C.E. Relationship of climate and genotype to seasonal variation in the glucosinolate-myrosinase system. I. Glucosinolate content in ten cultivars of *Brassica oleracea* grown in fall and spring seasons. *Journal of the Science of Food and Agriculture*, **2005**, 85 (4): p. 671-681.

38. Grubb, C.D.; Abel, S. Glucosinolate metabolism and its control. *Trends in Plant Science*, **2006**, 11 (2): p. 89-100.

39. Rangkadilok, N.; Nicolas, M.E.; Bennett, R.N.; Premier, R.R.; Eagling, D.R.; Taylor, P.W.J. Developmental changes of sinigrin and glucoraphanin in three *Brassica* species (*Brassica nigra, Brassica juncea* and *Brassica oleracea* var. *italica*). *Scientia Horticulturae*, **2002**, 96 (1-4): p. 11-26.

40. Bellostas, N.; Kachlicki, P.; Sørensen, J.C.; Sørensen, H. Glucosinolate profiling of seeds and sprouts of *B. oleracea* varieties used for food. *Scientia Horticulturae*, **2007**, 114 (4): p. 234-242.

41. Sonderby, I.E.; Geu-Flores, F.; Halkier, B.A. Biosynthesis of glucosinolates - gene

discovery and beyond. *Trends in Plant Science*, **2010**, 15 (5): p. 283-290.

42. Farre, E.M.; Weise, S.E. The interactions between the circadian clock and primary metabolism. *Current Opinion in Plant Biology*, **2012**, 15 (3): p. 293-300.

43. Oliviero, T.; Verkerk, R.; Dekker, M. A research approach for quality based design of healthy foods: Dried broccoli as a case study. *Trends in Food Science & Technology*, **2013**, 30 (2): p. 178-184.

44. Sarvan, I.; Verkerk, R.; Dekker, M. Modelling the fate of glucosinolates during thermal processing of *Brassica* vegetables. *Lwt-Food Science and Technology*, **2012**, 49 (2): p. 178-183.

45. Mithen, R.; Faulkner, K.; Magrath, R.; Rose, P.; Williamson, G.; Marquez, J. Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. *Theoretical and Applied Genetics,* **2003**, 106 (4): p. 727-734.

46. Mithen, R.F. Enhancing the health benefits of broccoli - From concept to commercialisation. *Acta Horticulture (ISHS)*, **2012**, 939: p. 39-43.

47. Sarikamis, G.; Marquez, J.; Mac Cormack, R.; Bennett, R.N.; Roberts, J.; Mithen, R. High glucosinolate broccoli: a delivery system for sulforaphane. *Molecular Breeding*, **2006**, 18 (3): p. 219-228.

48. Cartea, M.E.; Velasco, P. Glucosinolates in *Brassica* foods: bioavailability in food and significance for human health. *Phytochemistry Reviews*, **2008**, 7 (2): p. 213-229.

49. Francisco, M.; Cartea, M.E.; Soengas, P.; Velasco, P. Effect of genotype and environmental conditions on health-promoting compounds in *Brassica rapa. Journal of Agricultural and Food Chemistry*, **2011**, 59 (6): p. 2421-2431.

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

51. Koornneef, M.; Alonso-Blanco, C.; Vreugdenhil, D. Naturally occurring genetic variation in *Arabidopsis thaliana. Annual Review of Plant Biology,* **2004**, 55: p. 141-172.

52. Capanoglu, E.; Beekwilder, J.; Boyacioglu, D.; Hall, R.; De Vos, R. Changes in antioxidant and metabolite profiles during production of tomato paste. *Journal of Agricultural*

and Food Chemistry, **2008**, 56 (3): p. 964-973.

53. Cevallos-Cevallos, J.M.; Reyes-De-Corcuera, J.I.; Etxeberria, E.; Danyluk, M.D.; Rodrick, G.E. Metabolomic analysis in food science: a review. *Trends in Food Science & Technology*, **2009**, 20 (11-12): p. 557-566.

54. Wathelet, J.P.; Iori, R.; Leoni, O.; Quinsac, A.; Palmieri, S. Guidelines for glucosinolate analysis in green tissues used for biofumigation. *Agroindustria*, **2004**, 3 (3): p. 257-266.

55. Fiebig, H. Desulphation of glucosinolates-potential pitfalls of the HPLC method for double zero rapeseed. *FETT Wissenschaft Technologie/FAT Science Technology*, **1991**, 93 (7).

56. Wathelet, J.-P.; Mabon, N.; Marlier, M. Determination of glucosinolates in rapeseed improvement of the official HPLC ISO method (precision and speed). *Rapeseed Congress,* **1999**, Canberra, Australia.

57. Kliebenstein, D.J.; Kroymann, J.; Brown, P.; Figuth, A.; Pedersen, D.; Gershenzon, J.; Mitchell-Olds, T. Genetic control of natural variation in arabidopsis glucosinolate accumulation. *Plant Physiology*, **2001**, 126 (2): p. 811-825.

58. Van Eylen, D.; Bellostas, N.; Strobel, B.W.; Oey, I.; Hendrickx, M.; Van Loey, A.; Sørensen, H.; Sørensen, J.C. Influence of pressure/temperature treatments on glucosinolate conversion in broccoli (*Brassica oleraceae L.* cv *italica*) heads. *Food Chemistry*, **2009**, 112 (3): p. 646-653.

59. Bjerg, B.; Sørensen, H. Quantitative analysis of glucosinolates in oilseed rape based on HPLC of desulfoglucosinolates and HPLC of intact glucosinolates, in *World Crops: Production, Ulilization, Description - Glucosinolates in Rapeseeds: Analytical aspects*. Editor J.P. Wathelet. **1987**, Martinus Nijhoff Publishers Dordrecht-Boston-Lancaster, p. 125-150.

60. Cartea, M.E.; Rodriguez, V.M.; de Haro, A.; Velasco, P.; Ordas, A. Variation of glucosinolates and nutritional value in nabicol (*Brassica napus pabularia* group). *Euphytica*, **2008**, 159 (1-2): p. 111-122.

61. Graser, G.; Oldham, N.J.; Brown, P.D.; Temp, U.; Gershenzon, J. The biosynthesis of benzoic acid glucosinolate esters in *Arabidopsis thaliana*. *Phytochemistry*, **2001**, 57 (1): p. 23-32.

62. 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*, **2010**, 119 (1): p. 423-428.

63. Kusznierewicz, B.; Bartoszek, A.; Wolska, L.; Drzewiecki, J.; Gorinstein, S.; Namiesnik, J. 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*, **2008**, 41 (1): p. 1-9.

64. Agerbirk, N.; Olsen, C.E.; Nielsen, J.K. Seasonal variation in leaf glucosinolates and insect resistance in two types of *Barbarea vulgaris* ssp. *arcuata*. *Phytochemistry*, **2001**, 58 (1): p. 91-100.

65. Jia, C.G.; Xu, C.J.; Wei, J.; Yuan, J.; Yuan, G.F.; Wang, B.L.; Wang, Q.M. Effect of modified atmosphere packaging on visual quality and glucosinolates of broccoli florets. *Food Chemistry,* **2009**, 114 (1): p. 28-37.

66. Krumbein, A.; Schreiner, M. Asia-Blattgemüse – gesund und lecker: Analyse der Glucosinolate und deren Abbauprodukte mit HPLC-DAD-MSn und GC-MS. *Deutsche Lebensmittel-Rundschau*, **2011**, 107: p. 30-34.

67. Verkerk, R.; Dekker, M.; Jongen, W.M.F. Post-harvest increase of indolyl glucosinolates in response to chopping and storage of *Brassica* vegetables. *Journal of the Science of Food and Agriculture*, **2001**, 81 (9): p. 953-958.

68. Agerbirk, N.; De Vos, M.; Kim, J.H.; Jander, G. Indole glucosinolate breakdown and its biological effects. *Phytochemistry Reviews*, **2009**, 8 (1): p. 101-120.

69. Hrncirik, K.; Velisek, J.; Davidek, J. Comparison of HPLC and GLC methodologies for determination of glucosinolates using reference material. *Zeitschrift für Lebensmitteluntersuchung und -Forschung - International Journal of Food Research and Technology*, **1998**, 206 (2): p. 103-107.

70. Clarke, D.B. Glucosinolates, structures and analysis in food. *Analytical Methods,* **2010**, 2 (4): p. 310-325.

71. Bohuon, E.J.R.; Keith, D.J.; Parkin, I.A.P.; Sharpe, A.G.; Lydiate, D.J. Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*. *Theoretical and Applied Genetics*, **1996**, 93 (5-6): p. 833-839.

72. Kiddle, G.; Bennett, R.N.; Botting, N.P.; Davidson, N.E.; Robertson, A.A.B.; Wallsgrove,

R.M. High-performance liquid chromatographic separation of natural and synthetic desulphoglucosinolates and their chemical validation by UV, NMR and chemical ionisation-MS methods. *Phytochemical Analysis*, **2001**, 12 (4): p. 226-242.

73. ISO. Rapeseed - determination of glucosinolates content - Part 1: Method using high-performance liquid chromatography *ISO 9167-1*, **1992**: p. 1-9.

74. Kwon, S.H.; Kang, M.J.; Huh, J.S.; Ha, K.W.; Lee, J.R.; Lee, S.K.; Lee, B.S.; Han, I.H.; Lee, M.S.; Lee, M.W.; Lee, J.; Choi, Y.W. Comparison of oral bioavailability of genistein and genistin in rats. *International Journal of Pharmaceutics*, **2007**, 337 (1-2): p. 148-154.

75. Van Boekel, M.A.J.S. Statistical aspects of kinetic modeling for food science problems. *Journal of Food Science*, **1996**, 61 (3): p. 477-485.

76. Shapiro, T.A.; Fahey, J.W.; Wade, K.L.; Stephenson, K.K.; Talalay, P. Chemoprotective glucosinolates and isothiocyanates of broccoli sprouts: Metabolism and excretion in humans. *Cancer Epidemiology Biomarkers & Prevention*, **2001**, 10 (5): p. 501-508.

77. Rosa, E.A.S.; Heaney, R.K.; Portas, C.A.M.; Fenwick, G.R. Changes in glucosinolate concentrations in *Brassica* crops (*B. oleracea* and *B. napus*) throughout growing seasons. *Journal of the Science of Food and Agriculture*, **1996**, 71 (2): p. 237-244.

78. Verkerk, R.; Dekker, M. Glucosinolates and myrosinase activity in red cabbage (*Brassica oleracea L.* var. *Capitata f. rubra DC.*) after various microwave treatments. *Journal of Agricultural and Food Chemistry*, **2004**, 52 (24): p. 7318-7323.

79. Hennig, K.; Verkerk, R.; Bonnema, G.; Dekker, M. Pitfalls in the desulphation of glucosinolates in a high-throughput assay. *Food Chemistry*, **2012**, 134 (4): p. 2355-2361.

80. Van Boekel, M.A.J.S. Kinetic Modeling of Reactions in Foods, **2009** CRC Press, p. **1**-1 - **1**-6; **4** -16 - **4**-17; **7**-29 - **7**-40.

81. De Levie, R. Advanced Excel for scientific data analysis. 2nd edition, **2008** Oxford University Press.

82. De Levie, R. (2011). http://www.bowdoin.edu/~rdelevie/ excellaneous/#downloads. Access Date: November 24, 2011.

83. Velasco, P.; Cartea, M.E.; Gonzalez, C.; Vilar, M.; Ordas, A. Factors affecting the glucosinolate content of kale (*Brassica oleracea acephala* group). *Journal of*

Agricultural and Food Chemistry, **2007**, 55 (3): p. 955-962.

84. Oliviero, T.; Verkerk, R.; Dekker, M. Effect of water content and temperature on glucosinolate degradation kinetics in broccoli (*Brassica oleracea* var. *italica*). *Food Chemistry,* **2012**, 132 (4): p. 2037-2045.

85. Kliebenstein, D.J.; Gershenzon, J.; Mitchell-Olds, T. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics*, **2001**, 159 (1): p. 359-370.

86. Feng, J.; Long, Y.; Shi, L.; Shi, J.Q.; Barker, G.; Meng, J.L. Characterization of metabolite quantitative trait loci and metabolic networks that control glucosinolate concentration in the seeds and leaves of *Brassica napus*. *New Phytologist*, **2012**, 193 (1): p. 96-108.

87. Lou, P.; Zhao, J.; He, H.; Hanhart, C.; Pino Del Carpio, D.; Verkerk, R.; Custers, J.; Koornneef, M.; Bonnema, G. Quantitative trait loci for glucosinolate accumulation in *Brassica rapa* leaves. *New Phytologist*, **2008**, 179 (4): p. 1017-1032.

88. Wang, H.; Wu, J.; Sun, S.L.; Liu, B.; Cheng, F.; Sun, R.F.; Wang, X.W. Glucosinolate biosynthetic genes in *Brassica rapa*. *Gene*, **2011**, 487 (2): p. 135-142.

89. John Innes Centre News. (2012) British research leads to UK launch of Beneforté broccoli. http://news.jic.ac.uk/2011/10/british-research-leads-to-uk-launch-of-beneforte-broccoli/?utm_source=John+Innes+Centre&utm_campaign=c779743867-Advances_18&utm_medium=email. Access Date: July 4, 2012.

90. Hennig, K.; Verkerk, R.; Bonnema, G.; Dekker, M. Rapid estimation of glucosinolate thermal degradation rate constants in leaves of Chinese kale and broccoli (*B. oleracea*) in two seasons. *Journal of Agricultural and Food Chemistry*, **2012**, 60 (32): p. 7859-7865.

91. Sebastian, R.L.; Howell, E.C.; King, G.J.; Marshall, D.F.; Kearsey, M.J. An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. *Theoretical and Applied Genetics*, **2000**, 100 (1): p. 75-81.

92. Van Ooijen, J.W. Software for the mapping of quantitative trait loci in experimental populations of diploid species **2009**, Kyazma B.V. : Wageningen, The Netherlands.

93. Mao, D.H.; Liu, T.M.; Xu, C.G.; Li, X.H.; Xing, Y.Z. Epistasis and complementary gene action adequately account for the genetic bases of transgressive segregation of kilo-grain weight in rice. *Euphytica*, **2011**, 180 (2): p. 261-271.

94. Wright, E.M.; Kelly, J.D. Mapping QTL for seed yield and canning quality following processing of black bean (*Phaseolus vulgaris L.*). *Euphytica*, **2011**, 179 (3): p. 471-484.

95. Wang, J.M.; Yang, J.M.; McNeil, D.; Zhou, M.X. Mapping of quantitative trait loci controlling barley flour pasting properties. *Genetica*, **2010**, 138 (11-12): p. 1191-1200.

96. 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. The genetics of plant metabolism. *Nature Genetics*, **2006**, 38 (7): p. 842-849.

97. Hennig, K.; Verkerk, R.; Dekker, M.; Bonnema, G. Quantitative trait loci analysis of non-enzymatic glucosinolate degradation rates in *Brassica oleracea* during food processing Theoretical and Applied Genetics, **in press**, DOI: 10.1007/s00122-013-2138-1.

98. De Vos, C.H.R.; Schipper, B.; Hall, R.D. High-performance liquid chromatography-mass spectrometry analysis of plant metabolites in *Brassicaceae*, in *Plant Metabolomics : Methods and Protocols*. Editors: N.W. Hardy and R.D. Hall. **2012**, Humana Press: Totowa, NJ, p. 111-128.

99. Lommen, A. MetAlign: Interface-Driven, Versatile Metabolomics Tool for Hyphenated Full-Scan Mass Spectrometry Data Preprocessing. *Analytical Chemistry*, **2009**, 81 (8): p. 3079-3086.

100. Tikunov, Y.M.; Laptenok, S.; Hall, R.D.; Bovy, A.; de Vos, R.C.H. MSClust: A tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. *Metabolomics*, **2012**, 8 (4): p. 714-718.

101. Vallejo, F.; Tomás-Barberán, F.A.; Ferreres, F. Characterisation of flavonols in broccoli (*Brassica oleracea L.* var. *italica*) by liquid chromatography–UV diode-array detection–electrospray ionisation mass spectrometry. *Journal of Chromatography A*, **2004**, 1054 (1–2): p. 181-193.

102. Sumner, L.W.; Amberg, A.; Barrett, D.; Beale, M.H.; Beger, R.; Daykin, C.A.; Fan, T.W.M.;

Fiehn, O.; Goodacre, R.; Griffin, J.L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A.N.; Lindon, J.C.; Marriott, P.; Nicholls, A.W.; Reily, M.D.; Thaden, J.J.; Viant, M.R. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*, **2007**, 3 (3): p. 211-221.

103. Saeed, A.I.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M.; Sturn, A.; Snuffin, M.; Rezantsev, A.; Popov, D.; Ryltsov, A.; Kostukovich, E.; Borisovsky, I.; Liu, Z.; Vinsavich, A.; Trush, V.; Quackenbush, J. TM4: A free, open-source system for microarray data management and analysis. *BioTechniques*, **2003**, 34 (2): p. 374-378.

104. Breiman, L. Random forests. *Machine Learning*, **2001**, 45 (1): p. 5-32.

105. Nakamura, Y.; Asahi, H.; Altaf-Ul-Amin, M.; Kurokawa, K.; Kanaya, S. (2004) KNApSAcK: A Comprehensive Species-Metabolite Relationship Database. http://kanaya.naist.jp/KNApSAcK/KNApSAcK.php. Last Update Date: 2012/12/27. Access Date: April 2013.

106. Kim, J.; Jung, Y.; Song, B.; Bong, Y.-S.; Ryu, D.H.; Lee, K.-S.; Hwang, G.-S. Discrimination of cabbage (*Brassica rapa* ssp. *pekinensis*) cultivars grown in different geographical areas using 1H NMR-based metabolomics. *Food Chemistry*, **2013**, 137 (1–4): p. 68-75.

107. Pino Del Carpio, D.; Basnet, R.; Vos, R.H.; Maliepaard, C.; Visser, R.; Bonnema, G. The patterns of population differentiation in a *Brassica rapa* core collection. **2011**, 122 (6): p. 1105-1118.

108. Eggink, P.M.; Maliepaard, C.; Tikunov, Y.; Haanstra, J.P.W.; Bovy, A.G.; Visser, R.G.F. A taste of sweet pepper: Volatile and non-volatile chemical composition of fresh sweet pepper (*Capsicum annuum*) in relation to sensory evaluation of taste. *Food Chemistry*, **2012**, 132 (1): p. 301-310.

109. Macáková, K.; Mladěnka, P.; Filipský, T.; Říha, M.; Jahodář, L.; Trejtnar, F.; Bovicelli, P.; Proietti Silvestri, I.; Hrdina, R.; Saso, L. Iron reduction potentiates hydroxyl radical formation only in flavonols. *Food Chemistry*, **2012**, 135 (4): p. 2584-2592.

110. Rosa, E.A.S.; Heaney, R.K.; Rego, F.C.; Fenwick, G.R. The variation of glucosinolate concentration during a single day in young plants of *Brassica oleracea* var. *acephala*

and *capitata. Journal of the Science of Food and Agriculture*, **1994**, 66 (4): p. 457-463.

111. Nascimento, N.C.D.; Fett-Neto, A.G. Plant secondary metabolism and challenges in modifying its operation: An overview, in *Plant Secondary Metabolism Engineering, Methods in Molecular Biology*. Editor A.G. Fett-Neto. **2010**, Springer Science & Business Media, p. 1-13.

112. Bourderioux, A.; Lefoix, M.; Gueyrard, D.; Tatibouet, A.; Cottaz, S.; Arzt, S.; Burmeister, W.P.; Rollin, P. The glucosinolate-myrosinase system. New insights into enzyme-substrate interactions by use of simplied inhibitors. *Organic & Biomolecular Chemistry,* **2005**, 3 (10): p. 1872-1879.

113. Cerniauskaite, D.; Gallienne, E.; Karciauskaite, H.; Farinha, A.S.F.; Rousseau, J.; Armand, S.; Tatibouët, A.; Sackus, A.; Rollin, P. A simple O-sulfated thiohydroximate molecule to be the first micromolar range myrosinase inhibitor. *Tetrahedron Letters*, **2009**, 50 (26): p. 3302-3305.

114. Borgen, B.H.; Ahuja, I.; Thangstad, O.P.; Honne, B.I.; Rohloff, J.; Rossiter, J.T.; Bones, A.M. 'Myrosin cells' are not a prerequisite for aphid feeding on oilseed rape (*Brassica napus*) but affect host plant preferences. *Plant Biology*, **2012**, 14 (6): p. 894-904.

115. Borgen, B.H.; Thangstad, O.P.; Ahuja, I.; Rossiter, J.T.; Bones, A.M. Removing the mustard oil bomb from seeds: transgenic ablation of myrosin cells in oilseed rape (*Brassica napus*) produces MINELESS seeds. *Journal of Experimental Botany*, **2010**, 61 (6): p. 1683-1697.

116. Dong, Y.B.; Zhang, Z.W.; Shi, Q.L.; Wang, Q.L.; Zhou, Q.; Li, Y.L. Quantitative trait loci mapping and meta-analysis across three generations for popping characteristics in popcorn. *Journal of Cereal Science*, **2012**, 56 (3): p. 581-586.

117. Glauser, G.; Schweizer, F.; Turlings, T.C.J.; Reymond, P. Rapid profiling of intact glucosinolates in arabidopsis leaves by UHPLC-QTOFMS using a charged surface hybrid column. *Phytochemical Analysis*, **2012**, 23 (5): p. 520-528.

118. Gratacos-Cubarsi, M.; Ribas-Agusti, A.; Garcia-Regueiro, J.A.; Castellari, M. Simultaneous evaluation of intact glucosinolates and phenolic compounds by UPLC-DAD-MS/MS in *Brassica oleracea L.* var. *botrytis. Food Chemistry*, **2010**, 121 (1): p. 257-263.

119. Maldini, M.; Baima, S.; Morelli, G.; Scaccini, C.; Natella, F. A liquid chromatography-mass spectrometry approach to study "glucosinoloma" in broccoli sprouts. *Journal of Mass Spectrometry,* **2012**, 47 (9): p. 1198-1206.

120. Mellon, F.A.; Bennett, R.N.; Holst, B.; Williamson, G. Intact glucosinolate analysis in plant extracts by programmed cone voltage electrospray LC/MS: Performance and comparison with LC/MS/MS methods. *Analytical Biochemistry*, **2002**, 306 (1): p. 83-91.

121. Mohn, T.; Cutting, B.; Ernst, B.; Hamburger, M. Extraction and analysis of intact glucosinolates - a validated pressurized liquid extraction/liquid chromatography-mass spectrometry protocol for Isatis tinctoria, and qualitative analysis of other *Cruciferous* plants. *Journal of Chromatography A*, **2007,** 1166 (1-2): p. 142-151.

122. Li, L.; Paulo, M.J.; Strahwald, J.; Lübeck, J.; Hofferbert, H.R.; Tacke, E.; Junghans, H.; Wunder, J.; Draffehn, A.; Van Eeuwijk, F.; Gebhardt, C. Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield. *Theoretical and Applied Genetics*, **2008**, 116 (8): p. 1167-1181.

123. Simons, K.; Anderson, J.A.; Mergoum, M.; Faris, J.D.; Klindworth, D.L.; Xu, S.S.; Sneller, C.; Ohm, J.B.; Hareland, G.A.; Edwards, M.C.; Chao, S. Genetic mapping analysis of bread-making quality traits in spring wheat. *Crop Science*, **2012**, 52 (5): p. 2182-2197.

124. Rosa, E.; Heaney, R. The effect of cooking and processing on the glucosinolate content: studies on four varieties of Portuguese cabbage and hybrid white cabbage. *Journal of the Science of Food and Agriculture*, **1993**, 62 (3).

125. Hospital, F. Challenges for effective marker-assisted selection in plants. *Genetica*, **2009**, 136 (2): p. 303-310.

126. Jones, R.B.; Frisina, C.L.; Winkler, S.; Imsic, M.; Tomkins, R.B. Cooking method significantly effects glucosinolate content and sulforaphane production in broccoli florets. *Food Chemistry*, **2010**, 123 (2): p. 237-242.

127. Cieslik, E.; Leszczynska, T.; Filipiak-Florkiewicz, A.; Sikora, E.; Pisulewski, P.M. Effects of some technological processes on glucosinolate contents in cruciferous vegetables. *Food Chemistry*, **2007**, 105 (3): p. 976-981.

128. Francisco, M.; Velasco, P.; Moreno, D.A.; García-Viguera, C.; Cartea, M.E. Cooking

methods of *Brassica rapa* affect the preservation of glucosinolates, phenolics and vitamin C. *Food Research International*, **2010**, 43 (5): p. 1455-1463.

129. Volden, J.; Borge, G.I.A.; Hansen, M.; Wicklund, T.; Bengtsson, G.B. 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*, **2009**, 42 (1): p. 63-73.

130. Petroski, R.J.; Tookey, H.L. Interactions of thioglucoside glucohydrolase and epithiospecifier protein of cruciferous plants to form 1-cyanoepithioalkanes. *Phytochemistry*, **1982**, 21 (8): p. 1903-1905.

131. Wang, G.C.; Farnham, M.; Jeffery, E.H. Impact of thermal processing on sulforaphane yield from broccoli (*Brassica oleracea L.* ssp. *italica*). *Journal of Agricultural and Food Chemistry*, **2012**, 60 (27): p. 6743-6748.

132. Farnham, M.W.; Simon, P.W.; Stommel, J.R. Improved phytonutrient content through plant genetic improvement. *Nutrition Reviews*, **1999**, 57 (9): p. S19-S26.

133. Garcia, R.A.V.; Rangel, P.N.; Bassinello, P.Z.; Brondani, C.; Melo, L.C.; Sibov, S.T.; Vianello-Brondani, R.P. QTL mapping for the cooking time of common beans. *Euphytica*, **2012**, 186 (3): p. 779-792.

134. Czech, A.; Pawlik, M.; Rusinek, E. Contents of heavy metals, nitrates, and nitrites in cabbage. *Polish Journal of Environmental Studies*, **2012**, 21 (2): p. 321-329.

135. Asdeo, A.; Loonker, S. A comparative analysis of trace metals in vegetables. *Research Journal of Environmental Toxicology*, **2011**, 5 (2): p. 125-132.

136. Farnham, M.W.; Keinath, A.P.; Grusak, M.A. Mineral concentration of broccoli florets in relation to year of cultivar release. *Crop Science*, **2011**, 51 (6): p. 2721-2727.

137. Holst, B.; Williamson, G. Nutrients and phytochemicals: from bioavailability to bioefficacy beyond antioxidants. *Current Opinion in Biotechnology*, **2008**, 19 (2): p. 73-82.

138. Verhoeven, D.T.H.; Goldbohm, R.A.; Van Poppel, G.; Verhagen, H.; Van Den Brandt, P.A. Epidemiological studies on *Brassica* vegetables and cancer risk. *Cancer Epidemiology Biomarkers and Prevention*, **1996**, 5 (9): p. 733-748.

139. Holst, B.; Williamson, G. A critical review of the bioavailability of glucosinolates and related compounds. *Natural Product Reports*, **2004**, 21 (3): p. 425-447.

140. Shapiro, T.A.; Fahey, J.W.; Dinkova-Kostova, A.T.; Holtzclaw, W.D.; Stephenson, K.K.; Wade, K.L.; Ye, L.X.; Talalay, P. Safety, tolerance, and metabolism of broccoli sprout glucosinolates and isothiocyanates: A clinical phase I study. *Nutrition and Cancer-an International Journal*, **2006**, 55 (1): p. 53-62.

SUMMARY

Background

Phytochemicals are non-essential nutrients occurring in plant-based foods, which have been linked to a reduced incidence and progression of diseases. Glucosinolates (GLs) are phytochemicals that are typical for *Brassica* and other *Cruciferous* plants, such as cabbage, broccoli, Brussels sprouts, Chinese cabbage, mustard and horseradish. Chemically, GLs are β-thioglucosides *N*-hydroxysulphates with a sulphur-linked β-D-glucopyranose moiety. Up to 120 different GLs have been identified with a common core structure, but differing in their side chain (aliphatic, aromatic, indolic). However, only a restricted number (3-10 GLs) occurs in commonly consumed vegetables. The intake of GLs has been associated with lowered risks of several types of cancer and other diseases. To reach a high level of GLs in the vegetables at the stage of consumption, research along the food chain aims to increase the initial concentrations of certain GLs and to lower GL losses; however, effects of the different steps in the food chain are mostly investigated separately. In this thesis an interdisciplinary approach, combining food science and plant science, was applied to explore the possibilities to improve the retention of GLs throughout the food chain.

Aim

The objective of this thesis was to investigate genetic effects related to GL degradation during food processing in order to test if food processing parameters can be used as phenotypic traits and breed for vegetables with improved GL retention. The identification of genetic regions associated with GL degradation provides a starting point to investigate molecular mechanisms causing GL losses. Variation in the trait of interest is required to identify genetic loci associated with it. Several mechanisms have been described leading to GL reduction during food processing; one of them is thermal degradation, which is solely induced by heat. Variation of GL thermal degradation rates of chemically identical GLs has been shown among different vegetables, which provides the basis to study genetic effects and was the focus of this thesis.

Results

One challenge of integrating plant science and food science is the high number of samples to be analysed to apply quantitative genetics to food technological traits. In the first study (described in **chapter 2**), the analysis of GLs, as desulpho-GLs, was optimized to reach accurate results using a high-throughput method. The effect of the sulphatase preparation on the accuracy of the enzymatic desulphation step in microtiter plates was tested. The desulphation procedure applied to determine GLs as desulpho-GLs was found to be crucial for the analytical result. For the first time an inverse effect of the sulphatase concentration on the peak area of a GL, glucotropaeolin, which is often used as internal standard, was shown, leading to a substantial overestimation of GL concentrations. We recommend the application of a purified sulphatase preparation to obtain accurate results for a broad range of samples. This finding emphasises the importance of having suitable analytical methods, in particular for the interdisciplinary approach, where actions in the various disciplines depend strongly on the analytical outcomes.

Kinetic modelling was applied as a tool to describe GL thermal degradation in a quantitative way in a broccoli (*B. oleracea* var. *italica*) and a Chinese kale (*B. oleracea* var. *alboglabra*) genotype in **chapter 3**. The first aim was to identify an appropriate model for the description of GL degradation in the two genotypes and two different seasons. Model discrimination was performed for first- and n^{th}-order models and it was shown that a first order reaction was appropriate in both seasons and genotypes. The first and n^{th}-order models are empirical models, which were applied since the molecular mechanism of thermal degradation is not known. The first order model was selected based on the best fit through the measured data. The second aim of this study was to investigate the effect of season and genotype on GL thermal degradation. Degradation rate constants obtained from the modelling were higher for indolic GLs than for aliphatic GL. Furthermore, we have previously demonstrated that thermal degradation rate constants of chemically identical GLs differed in five different *Brassica* vegetables; hence differences in GL thermal degradation between this Chinese kale and this broccoli genotype grown in the same season were expected. However, the resulting degradation rate constants of chemically identical GLs did not differ between the broccoli and the Chinese kale genotype when grown in the same season, but were strongly affected by the season (20% to 80% difference). In this study, defined genotypes were grown under

the same conditions, whereas in literature different vegetables, at different maturity, with unknown growing conditions and storage history were used.

A publically available segregating population (doubled-haploid (DH)), developed by crossing the broccoli and the Chinese kale genotype studied in chapter 3, was investigated for GL thermal degradation in **chapter 4**. Thermal degradation rate constants were determined in 100 DH lines and were combined with molecular marker information to identify genetic regions associated with GL thermal degradation (quantitative trait loci (QTL)). Despite the fact that the degradation rate constants of chemically identical GLs did not differ in the parental genotypes of the segregating population, rate constants varied by 3-fold throughout the population for most GLs. The phenomenon that the phenotypic values of the segregating population are beyond the range of the parental lines is called transgressive segregation and generally means that the trait is quantitatively inherited (controlled by several genes). QTLs were identified for the two GLs which occurred in all genotypes; two QTLs were identified for glucobrassicin on linkage groups C07 and C09 (25% explained variation) and one for glucoraphanin on linkage group C07 (12% of explained variation). A co-localisation of the QTLs on linkage group C07 and correlations between the rate constants of different GLs indicate that similar factors in the plant matrix influence GL degradation. In addition to the genotypic effects, the growing year influences GL degradation as well, which is a common phenomenon of quantitatively inherited traits. The identified QTL can be used for marker-assisted selection (MAS), which is a selection method based on the genotype at a specific position to increase breeding efficiency. The identified QTL must be stable across environments and time to be applied efficiently to MAS, which cannot be predicted from the current experimental data. Hence, further experiments are required to apply the identified GL thermal degradation QTL to MAS and breed for vegetables with increased GL retention during food processing in the future.

Up to now the full sequence of the *Brassica oleracea* genome has not been published, which hinders the identification of genes underlying the detected QTL for GL degradation. Furthermore, fine mapping of the DH population would be required to obtain smaller genetic regions with fewer possible genes within the QTL. The selection of candidate genes is difficult, since little prior knowledge about the molecular mechanism of GL degradation exists. As an alternative, an untargeted metabolomics approach combined with random forest regression was applied to identify metabolites associated with GL thermal degradation rate constants to form new hypotheses about the underlying

mechanism (**chapter 5**). Metabolites associated with the variation in the GL thermal degradation rate constants throughout the DH population were identified. Out of 413 analysed metabolites, 15 were significantly associated with the degradation rate constants of glucobrassicin, 6 with the degradation rate constants of glucoraphanin and 2 with both GL degradation rate constants. Due to the co-localisation of the degradation QTLs, a higher number of metabolites associated with both GL degradation rate constants was expected. Among these 23 different metabolites, 3 were identified as flavonols (1 kaempferol- and 2 quercetin-derivatives) and 2 as GLs (gluconasturtiin, 4-methoxyglucobrassicin). The flavonols were positively correlated with the degradation rate constants of glucobrassicin and possibly enhance the degradation by the reduction of ferric ions (Fe^{3+}) to ferrous ions (Fe^{2+}) that mediate the degradation of GL under the formation of ferric ions. One GL was negatively correlated with the degradation rate constant of glucobrassicin and 2 GLs with the rate constant of glucoraphanin, which may reduce the degradation of glucobrassicin and glucoraphanin by a competitive mechanism. Furthermore, 18 metabolites could not be annotated, some showing positive and some showing negative correlation. Possibly enhancing and retarding effects of metabolites on GL degradation illustrate the complexity of the system. Further research is required to prove the findings and to investigate the mechanisms of thermal GL degradation in more detail.

QTL analysis of metabolites associated with the GL degradation rate constants revealed metabolite QTL on linkage groups C07 and C09, which co-localise with the QTL identified for GL degradation rate constants. An additional QTL was detected on linkage group C01 where no QTL for the degradation rate constants had been identified (**chapter 5**). The future availability of sequence information of *Brassica oleracea* may help to identify the genes underlying metabolite and degradation QTLs. The combination of sequence information and the identity of the metabolites will allow a more specific selection of candidate genes.

The combination of plant science and food science to improve food quality in general is discussed in **chapter 6**. The main challenge to breed for vegetables with improved food processing properties is the number of samples to be analysed. Recent advances in development of high-throughput analytical methodologies will facilitate the analysis of such high sample numbers. Both disciplines will benefit from such collaboration and it may reduce variability and increase the quality of plant based products along the food chain in the future.

Conclusions

The research conducted in this thesis demonstrates that GL thermal degradation is partly genetically regulated. In addition, environmental factors, such as season and growing year, influence GL thermal degradation. The findings provide necessary information to breed for vegetables with increased GL retention during food processing, but more studies about environmental stability of the genetic regions are required to use this genetic regions in MAS in the future. Furthermore, a methodology to identify metabolic factors causing the variation of GL thermal degradation in different vegetables was applied successfully. Three flavonols, two GLs and 18 not annotated metabolites were associated with GL thermal degradation, which also requires future research to confirm these findings. The use of other metabolomics/ proteomics platforms may reveal more metabolites associated with thermal degradation. The combination of metabolite identities with genetic loci associated with GL thermal degradation and future sequence information of *Brassica oleracea* will help to identify candidate genes and molecular mechanisms for GL thermal degradation.

In order to improve specific quality attributes of plant foods, breeding for quantitative food processing traits is a promising and challenging approach. It has potential in the improvement of the nutritional quality of food products by combining the disciplines food and plant science to select and breed for varieties that not only have higher initial amounts of phytochemicals but that also have a high retention during storage and processing.

Samenvatting

Achtergrond

Fytochemicaliën zijn in planten aanwezige nutriënten die voor de mens niet essentieel zijn, maar welke waarschijnlijk wel gezondheidsbevorderende eigenschappen hebben die een rol spelen bij het voorkomen van, en/of de vermindering van het risico op ziekten zoals kanker en hart- en vaatziekten. Glucosinolaten (GLn) zijn een groep van fytochemicaliën die alleen in *Brassica* en andere groenten uit de familie van de kruisbloemigen (cruciferen), zoals bijvoorbeeld rode en boerenkool, broccoli, spruitjes, Chinese kool, mosterd en mierikswortel, aanwezig zijn. Chemisch gezien zijn GLn *β*-thioglucoside *N*-hydroxysulphaten met een *β*-D-glucopyranose groep die aan een zwavel-atoom gebonden is. Er zijn ongeveer 120 verschillende GLn geïdentificeerd, met een identieke kernstructuur, maar met ieder een verschillende alifatische, aromatische of indolische zijketen. In specifieke *Brassica*-groenten komt slechts een beperkt aantal (3-10) GLn voor. Het eten van *Brassica*-groenten en hiermee dus de inname van specifieke GLn is gekoppeld aan een vermindering van het risico op verschillende vormen van kanker en andere ziekten. Verschillende stappen in de levensmiddelenproductieketen beïnvloeden de concentraties van GLn in de groenten. De concentratie na de oogst is afhankelijk van het genotype, het klimaat, de bodem en andere omstandigheden. Door opslag en huishoudelijk of industrieel koken kan de concentratie van GLn sterk gereduceerd worden. Door onderzoek aan verschillende stappen in de levensmiddelenproductieketen kunnen de GL-concentraties in het geconsumeerde product verhoogd worden. Dit kan bijvoorbeeld door met behulp van veredeling hogere concentraties van bepaalde GLn in het geoogste gewas te verkrijgen, maar ook door te zorgen voor lagere verliezen van GLn tijdens de verwerking en bereiding van groenten. Meestal worden effecten van verschillende stappen in de levensmiddelketen afzonderlijk van elkaar onderzocht. In dit proefschrift zijn de expertises van plantenwetenschappers gecombineerd met die van levensmiddeltechnologen om zo met een interdisciplinaire onderzoekbenadering door de hele levensmiddelenproductieketen GLn beter te behouden.

Doel van dit proefschrift

Het doel van dit proefschrift is het onderzoeken of er genetische factoren zijn die een rol spelen bij de afbraak van GLn tijdens de bereiding. Voor het bepalen van de genetische regio's is onderzocht of procesparameters tijdens de bereiding van levensmiddelen als fenotypische eigenschappen kunnen worden gebruikt om groenten te kunnen veredelen met een hoger behoud van GLn tijdens de voedselverwerking en –bereiding. De identificatie van gebieden in het kool-genoom die geassocieerd zijn met GL-afbraak biedt een uitgangspunt om moleculaire mechanismen die de GL-afbraak veroorzaken te onderzoeken. Deze gebieden in het genoom worden aangeduid als Quantitative Trait Loci (QTL), welke bepaald kunnen worden in genetische studies aan populaties; deze gebieden verklaren een deel van de variatie in de te meten eigenschap, wat veroorzaakt wordt door genen die in dit gebied liggen met direct effect op die eigenschap. Om de QTL die aan de eigenschappen gekoppeld zijn te identificeren, is het van belang dat variatie in de eigenschappen aanwezig is in planten met verschillende genetische achtergrond. Verschillende mechanismen voor GL-afbraak tijdens de voedselbereiding zijn beschreven, één van deze mechanismen is de thermische afbraak die bij hoge temperatuur plaatsvindt. Variatie in de thermische afbraak van chemisch identieke GLn in verschillende koolgroenten was beschreven en biedt dus de basis om genetische factoren die aan deze thermische GL-afbraak gekoppeld zijn te bepalen en verder te onderzoeken.

Onderzoek en resultaten

Het koppelen van de onderzoeksgebieden op het gebied van plantenveredeling en levensmiddelentechnologie vereist de analyse van grote aantallen monsters om kwantitatieve genetica op levensmiddelentechnologische eigenschappen toe te kunnen passen. In **hoofdstuk 2** is de optimalisatie van een GL-analysemethode (als desulfo-GLn) beschreven. De geoptimaliseerde methode geeft nauwkeurige resultaten in een snelle procedure geschikt voor grote aantallen monsters. De invloed van de sulfataseconcentratie en de monstervoorbereiding op de nauwkeurigheid van de enzymatische desulfatatie in microtiter- platen is hiervoor in detail bestudeerd. De details van de toegepaste desulfatatiemethode is cruciaal voor de nauwkeurigheid van het analyseresultaat. Voor de eerste keer is hierbij een negatief effect gevonden van de sulfataseconcentratie op het piekoppervlak van een GL, glucotropaeolin, dat vaak

als interne standaard wordt gebruikt, Hierdoor bleken er systematisch te hoge GL-concentraties te worden gevonden. Het gebruik van een gezuiverd sulfatase-preparaat wordt hierdoor aanbevolen om nauwkeurige resultaten in monsters met uiteenlopende GL-concentraties te verkrijgen. Dit resultaat benadrukt het belang van geschikte analytische methoden, vooral bij interdisciplinair onderzoek waarbij de analyse-resultaten in sterke mate de vervolgstappen in de verschillende disciplines bepalen.

Kinetisch modelleren is als hulpmiddel gebruikt om de thermische afbraak van GLn in de gebruikte ouderlijnen van een uitsplitsende populatie, broccoli broccoli (*B. oleracea* var. *italica*) en Chinese broccoli of Gai Lan (*B. oleracea* var. *alboglabra*), kwantitatief te beschrijven (**hoofdstuk 3**). Het eerste doel van deze studie was een geschikt model voor de kinetiek van afbraak in de twee genotypen voor twee seizoenen te identificeren. Door middel van modeldiscriminatie werd aangetoond dat voor beide genotypen en seizoenen een 1^e-orde kinetiek de experimentele data beter beschreef dan een n^e-orde kinetiek. Beide onderzochte modellen zijn empirische modellen, welke zijn gekozen omdat het moleculaire mechanisme van thermische GL-afbraak niet bekend is. Het tweede doel van deze studie was om de invloed van genotype en seizoen op de thermale afbraak van de individuele GL te onderzoeken. De bepaalde 1^e-orde snelheidsconstanten waren hoger voor GLn met een indolische zijketen dan voor de GLn met een alifatische zijketen. In een eerdere studie was aangetoond dat de snelheidsconstanten voor thermische afbraak van chemisch identieke GLn in vijf verschillende *Brassica*-groenten verschillend zijn. Hierdoor werd een verschillende thermische afbraak snelheid tussen de broccoli en Chinese broccoli genotypen in hetzelfde seizoen verwacht. De gevonden snelheidsconstanten verschilden echter niet tussen de onderzochte broccoli en Chinese broccoli genotypen in hetzelfde seizoen. Daarentegen was de invloed van het seizoen op de thermische afbraak van chemisch identieke GL groot (20% tot 80% verschil). In deze studie zijn goed gedefinieerde organen (bladeren van 8 weken oude planten) van bekende genotypen welke onder dezelfde condities geteeld zijn, gebruikt. In de eerdere studie daarentegen, zijn verschillende groenten gebruikt met vele, vaak onbekende, verschillen, zoals: leeftijd, orgaantype, teeltlocatie en -omstandigheden, opslagcondities.

De snelheidsconstantes van de thermische GL-afbraak is verder onderzocht in honderd lijnen van een vrij beschikbare, splitsende, verdubbelde haploïde populatie (DH), die door kruising van de broccoli met de Chinese broccoli is verkregen (**hoofdstuk 4**). Deze ouderlijnen zijn in hoofdstuk 3 voor thermische GL-afbraak onderzocht. De

gevonden afbraakconstanten zijn gecombineerd met moleculaire marker data van al de honderd lijnen van de DH populatie om genetische regio's (quantitative trait loci (QTL)) te identificeren die met de thermische GL-afbraak geassocieerd zijn. Hoewel de thermische afbraak van chemisch identieke GLn in de oudergenotypen van deze populatie niet verschilden, varieerden de snelheidsconstanten met een factor drie voor de splitsende populatie van nakomelingen. Het fenomeen dat de fenotypische waarden buiten het bereik van de waarden van de ouder genotypen liggen wordt transgressieve segregatie genoemd. Dit betekent dat een eigenschap kwantitatief overerft en wordt bepaald door meer dan één gen. QTL voor de GLn die in alle genotypen aanwezig waren, zijn geïdentificeerd: twee QTL voor glucobrassicin op chromosomen C07 en C09 (25% verklaarde variatie) en één QTL voor glucoraphanine op chromosoom C07 (12% verklaarde variatie). Het feit dat de QTL op chromosoom C07 voor twee verschillende GLn op dezelfde positie liggen en dat de afbraakconstanten van verschillende GLn in dezelfde planten gecorreleerd zijn, wijst erop dat soortgelijke factoren in de plantmatrix de thermische afbraak van deze verschillende GLn beïnvloeden. Naast de genotypische effecten heeft het jaar van de teelt ook invloed op de GL-afbraak. Dit is een gebruikelijk fenomeen voor kwantitatief overervende eigenschappen. De geïdentificeerde QTL kunnen gebruikt worden voor merkergestuurde selectie (MAS), een indirecte selectiemethode die gebruik maakt van moleculaire merkers die de QTL flankeren. Dit verhoogt de effectiviteit van de veredeling voor stabielere GLn, daar merkeranalyse vele malen minder kostbaar en arbeidsintensief is, dan het bepalen van GL-afbraak aan vele genotypen van splitsende populaties. Bovendien is MAS niet afhankelijk van seizoeneffecten en teeltcondities, en kunnen QTL die alleen in bepaalde teelt condities tot expressie komen, toch altijd geselecteerd worden. Of de geïdentificeerde QTL voor de thermische GL-afbraak effectief zijn onder verschillende teeltomstandigheden, kan met de huidige data nog niet voorspeld worden. Meer experimenten zijn noodzakelijk om de thermische afbraak QTL in MAS in de toekomst toe te kunnen passen voor het veredelen van groenten met een verhoogde stabiliteit van GLn.

De sequentie van het *Brassice oleracea*-genoom is tot nu toe niet gepubliceerd, hetgeen de identificatie van de genen op het genoom onder de QTL voor thermische GL-afbraak verhindert. Bovendien is gedetailleerde genetische kartering noodzakelijk om kleinere genetische regio's te verkrijgen met minder opties voor kandidaatgenen in de QTL voor GL- afbraak. Het vinden van kandidaatgenen is moeilijk omdat weinig bekend is over de factoren die de GL-afbraak beïnvloeden.

Een alternatieve benadering om de onderliggende factoren voor de thermische GL-afbraak te identificeren is "untargeted metabolomics" gecombineerd met "random forest" regressie. Deze methode is gekozen om metabolieten die met de GL-afbraaksnelheidsconstanten geassocieerd zijn te identificeren en om hiermee nieuwe hypothesen over de mechanismen van de thermische GL-afbraak te kunnen formuleren (**hoofdstuk 5**). Een aantal metabolieten die met de GL-afbraak in de DH-populatie geassocieerd zijn, zijn met deze methode gevonden. Van de 413 metabolieten welke geïdentificeerd zijn in de DH-populatie, waren er 15 significant geassocieerd met de snelheidsconstante van glucobrassicine, 6 met de snelheidconstanten van glucoraphanin en 2 met die van allebei GLn. De verwachting was dat een groter aantal metabolieten met beide GL-afbraaksnelheidsconstanten geassocieerd zou zijn, omdat de QTL op koppelingsgroep C07 voor beide GLn is gevonden. Uit deze 23 metaboliten zijn 3 als flavonolen (een kaempferol- and twee quercetine-derivaten) en twee als GLn (4-methoxyglucobrassicine, gluconasturtiine) geïdentificeerd. De flavonolen zijn positief gecorreleerd met de snelheidsconstanten van glucobrassicin. Het zou kunnen dat flavonolen de afbraaksnelheid van glucobrassicin verhogen doordat ze Fe^{3+} ijzerionen kunnen reduceren naar Fe^{2+} ionen, die de afbraak van GLn onder vorming van Fe^{3+} kunnen versnellen. Één GL was geassocieerd met de snelheidsconstante van glucobrassin en twee GLn met de afbraaksnelheidsconstante van glucoraphanin, welke mogelijk de afbraak van glucobrassicine en glucoraphanine door een competitief mechanisme vertragen. De identiteit van 18 metabolieten kon niet bepaald worden, sommige hiervan zijn positief en sommige negatief met de GL-afbraaksnelheidsconstanten gecorreleerd. Het feit dat er zowel versnellende als vertragende effecten van metabolieten gevonden zijn toont de complexiteit van de GL-afbraakmechanismen aan. Vervolgstudies zijn noodzakelijk om deze resultaten te bevestigen en om de moleculaire mechanismen van de GL-afbraak op te helderen.

De QTL-analyse van de metabolieten die met de GL-afbraaksnelheidsconstanten gecorreleerd waren, heeft QTL op chromosomen C07 en C09 aangetoond, die op dezelfde posities liggen als de QTL voor de GL-afbraaksnelheden zelf. Verder is een QTL op chromosoom C01 geïdentificeerd waar geen QTL voor GL-afbraak gevonden zijn (**hoofdstuk 5**). Sequentie-informatie van *Brassica oleracea* kan in de toekomst de zoektocht naar genen die onder de QTL voor metabolieten en GL-afbraak liggen ondersteunen. De combinatie van sequentie-informatie en de identiteit van metabolieten kan het bepalen van kandidaatgenen betrokken bij de thermische afbraak van GLn faciliteren.

De combinatie van plantenwetenschappen en levensmiddelentechnologie om de kwaliteit van levensmiddelen te verbeteren is bediscussieerd in *hoofdstuk 6*. De grootste uitdaging bij de plantenveredeling voor levensmiddelentechnologische eigenschappen is het grote aantal monsters dat geanalyseerd moet worden voor kwantitatief genetische analyses. Nieuwe ontwikkelingen in "high-throughput" analyse technieken zullen analyse van grote monsterhoeveelheden makkelijker maken. Allebei de disciplines zullen van deze samenwerking profiteren en wellicht kan hierdoor de grote variatie in de kwaliteit van op planten gebaseerde producten hierdoor in de toekomst worden verminderd.

Conclusies

Het onderzoek in dit proefschrift toont aan, dat de thermische afbraak van GLn deels genetisch gereguleerd is. Hiernaast beïnvloeden omgevingsfactoren, bijvoorbeeld seizoen en jaar van de teelt, de afbraaksnelheid van GLn. De resultaten van dit proefschrift geven de noodzakelijke informatie om groenten te kunnen veredelen met een verhoogde GL-stabiliteit tijdens de voedselverwerking en bereiding. Wel zijn er meer studies nodig om de stabiliteit van de gevonden QTL in verschillende omgevingen te onderzoeken en dus de effectiviteit van MAS voor de GLn stabiliteit te bepalen. Verder is een methologie ontwikkeld om metabole factoren die de thermale GL-afbraak kunnen beïnvloeden te identificeren. Drie flavonolen, twee GLn en 18 ongeïdentificeerde metabolieten zijn met GL-afbraak geassocieerd. Door andere metabolomics/ proteomics-platforms te gebruiken kunnen mogelijk meer metabolieten geïdentificeerd worden die de thermische GL-afbraak beïnvloeden. De combinatie van metabolietinformatie, de genetische loci die de GL-afbraak beïnvloeden en sequentie- informatie van *Brassica oleracea* zal in de toekomst meer mogelijkheden bieden voor het identificeren van kandidaatgenen en moleculaire mechanismen voor thermale GL-afbraak.

De combinatie van plantenveredeling en kwantitatieve informatie vanuit de levensmiddelentechnologie is een veelbelovende en uitdagende nadering om de kwaliteit van levensmiddelen te verbeteren. Er is potentie om de nutritionele kwaliteit van levensmiddelen te verbeteren door deze interdisciplinaire aanpak, met als doel gewassen te veredelen die niet alleen een hoger uitgangsgehalte aan fytochemicaliën hebben, maar waarin deze componenten ook stabieler zijn tijdens verwerking, bereiding en opslag.

ZUSAMMENFASSUNG

Hintergrund

Glucosinolate sind sekundäre Pflanzenstoffe, die typischerweise in den Arten der *Brassicacea* und anderen Kreuzblütengewächsen vorkommen, wozu Gemüse wie Brokkoli, Rotkohl, Rosenkohl, Chinakohl, aber auch Raps, Senf und Meerrettich gehören. Glucosinolaten und deren Abbauprodukten werden gesundheitsfördernde Eigenschaften, insbesondere die Hemmung von Krebserkrankungen, nachgesagt. Außerdem tragen die Abbauprodukte von Glucosinolaten zum Geschmack und Aroma von *Brassica*-Gemüse bei. Chemisch gehören Glucosinolate zu den *β*-Thioglucosid-*N*-hydroxysulfaten. Diese bestehen aus einem Agluconrest der über ein Schwefelatom mit einer Glucoseeinheit verbunden ist. Über 120 verschiedene Glucosinolate, die eine gemeinsame Grundstruktur haben und sich im Aufbau des Agluconrestes unterscheiden, wurden bisher identifiziert. Basierend auf der chemischen Struktur des Agluconrestes werden Glucosinolate in drei Untergruppen eingeteilt, in aliphatische, aromatische und indolische Glucosinolate. Jede Gemüsesorte enthält allerdings nur eine bestimmte Anzahl an Glucosinolaten (3-10).

Aufgrund ihrer gesundheitsfördernden Eigenschaften wird eine hohe Glucosinolatkonzentration zum Zeitpunkt des Verzehrs angestrebt. Diese kann durch verschiedene Schritte in der Lebensmittelkette modifiziert werden. So haben zum einen der Genotyp und die Anbaubedingungen einen Einfluss auf die Ausgangskonzentration der Glucosinolate und zum anderen können die Lagerung und die Zubereitung der Gemüse den Glucosinolatgehalt in Gemüsesorten erheblich reduzieren. Das Ausmaß der Verluste wird durch die Art und Dauer der Zubereitung beeinflusst. Bisher wurde der Einfluss der verschiedenen Schritte in der Lebensmittelkette getrennt voneinander betrachtet. Der Gegenstand dieser Arbeit ist es deshalb Pflanzenzüchtung und Lebensmitteltechnologie zu verbinden, um Möglichkeiten zu erforschen, hohe Glucosinolatkonzentrationen innerhalb der gesamten Lebensmittelkette aufrechtzuerhalten.

Zielstellung

Das Ziel dieser Arbeit war es, genetische Effekte des Glucosinolatabbaus während der Gemüsezubereitung zu untersuchen. Dabei sollte geprüft werden, ob lebensmitteltechnologische Parameter als phänotypische Merkmale eingesetzt werden können, um Gemüse zu züchten, die Glucosinolatkonzentrationen länger aufrechterhalten. Die Bestimmung von genetischen Regionen kann als Ausgangspunkt verwendet werden, um molekulare Mechanismen des Glucosinolatabbaus zu untersuchen. Unterschiede in einer phänotypischen Eigenschaft sind Voraussetzung für die Identifizierung von zugehörigen genetischen Regionen. Verschiedene Mechanismen wurden bereits in der Literatur beschrieben, die Glucosinolatverluste während der Gemüsezubereitung beeinflussen. Dazu gehört u.a. der thermisch induzierte Abbau, für den gezeigt werden konnte, dass chemisch identische Glucosinolate in verschiedenen Gemüsesorten unterschiedlich stabil sind. Diese Unterschiede bilden die Grundlage, um genetische Effekte zu untersuchen und sind Schwerpunkt dieser Arbeit.

Ergebnisse

Die hohe Probenanzahl, die notwendig ist, um quantitative Genetik und lebensmitteltechnologische Eigenschaften zu verbinden, stellt eine besondere Herausforderung dar. Deshalb wurde die Bestimmung von Glucosinolaten optimiert, um genaue Ergebnisse mit einem hohen Probendurchsatz zu erzielen (**Kapitel 2**). Glucosinolate werden üblicherweise als Desulfo-Glucosinolate bestimmt, um ihre Polarität zu verringern und somit die chromatographische Trennung zu verbessern. Die enzymatische Desulfierung wurde in Mikrotiterplatten durchgeführt. Dabei wurde sowohl der Einfluss der Sulfatasekonzentration als auch der Einfluss der Reinigung der Sulfatase auf Richtigkeit des Analysenergebnisses untersucht. Die Ergebnisse zeigen, dass die Konzentration und die Reinigung der Sulfatase einen entscheidenden Einfluss auf das Analysenergebnis haben. Erstmalig wurde ein negativer Effekt der Sulfatasekonzentration auf die Peakfläche eines Glucosinolates, dem Glucotropaeolin, gezeigt. Glucotropaeolin wird oft als interner Standard verwendet, deshalb führen geringere Peakflächen zu falsch positiven Ergebnissen aller Glucosinolate. Ein kurzer Reinigungsschritt der Sulfatase wird empfohlen, um genaue Analysenergebnisse in verschiedenen Proben mit einem breiten Konzentrationsbereich zu erzielen.

Für die Quantifizierung des thermisch induzierten Glucosinolatabbaus in einem

Brokkoli-Genotyp (*B. oleracea* var. *italica*) und einem chinesischen Brokkoli-Genotyp (*B. oleracea* var. *alboglabra*) wurde die kinetische Modellierung (**Kapitel 3**) angewendet. Als Erstes wurde ein geeignetes Modell zur Beschreibung des Glucosinolatabbaus in den zwei Genotypen, die in zwei verschiedenen Jahreszeiten angebaut wurden, identifiziert. Es wurden empirische Modelle der Reaktionen erster und n-ter Ordnung angewendet, da der molekulare Abbaumechanismus der Glucosinolate nicht bekannt ist. Zur Unterscheidung, welche der beiden Modelle die Daten besser beschreibt, wurde die Methode der kleinsten Quadrate angewendet und der resultierende Fit zwischen den beiden Modellen verglichen. Das Modell erster Ordnung wurde für weitere Modellierungen ausgewählt, da die bestimmten Parameter genauer als mit dem Modell n-ter Ordnung waren. Als Zweites wurde der Einfluss des Genotyps und der Jahreszeit auf den Glucosinolatabbau untersucht. Die mit dem Modell erster Ordnung bestimmten Abbauraten waren im Fall der indolischen Glucosinolate größer als für aliphatische Glucosinolate. Die Abbauraten chemisch identischer Glucosinolate waren in diesem Brokkoli-Genotyp und in diesem chinesischen Brokkoli-Genotyp gleich groß, wenn sie in der gleichen Jahreszeit angebaut wurden. Die Jahreszeit hatte einen großen Einfluss auf die thermische Stabilität von Glucosinolaten in den zwei untersuchten Genotypen (20%ige bis 80%ige Unterschiede). Im Gegensatz dazu, wurde in vorhergehenden Untersuchungen gezeigt, dass chemisch identische Glucosinolate in verschiedenen Gemüsesorten unterschiedliche thermische Stabilitäten aufweisen. Für die in dieser Arbeit beschriebenen Experimente wurden zwei Genotypen verwendet, die unter gleichen Bedingungen angebaut wurden. Hingegen wurden in den in der Literatur beschriebenen Experimenten verschiedene Gemüse mit unbekanntem Genotyp und Herkunft, unterschiedlichen Anbau- und Lagerbedingungen und Reife verwendet. Diese Faktoren können ein Grund für die unterschiedlichen Ergebnisse der beiden Studien sein.

Durch Kreuzung der in Kapitel 3 untersuchten Brokkoli- und chinesischen Brokkoli-Genotypen wurde eine doppelhaploide Pflanzenpopulation entwickelt. In **Kapitel 4** wurde die thermische Stabilität von verschiedenen Glucosinolaten in dieser Population untersucht. Thermische Abbauraten von Glucosinolaten wurden in 100 Pflanzenlinien bestimmt und die Daten mit genetischen Marker-Informationen kombiniert, um Genabschnitte zu identifizieren, die mit thermischen Glucosinolatabbau assoziiert sind (quantitative trait loci [QTL]). Obwohl die Abbauraten chemisch identischer Glucosinolate in den elterlichen Genotypen der Pflanzenpopulation nicht verschieden waren, variierten

die Abbaukonstanten der meisten Glucosinolate innerhalb der Pflanzenpopulation dreifach. Wenn die phänotypischen Merkmale der Population jenseits der elterlichen Werte liegen, bedeutet das, dass eine Eigenschaft quantitativ vererbt wird, das heißt von mehreren Genen beeinflusst wird. Es konnten QTL für die Abbauraten von zwei Gluosinolaten bestimmt werden: Zwei QTL wurden für Glucobrassicin auf den Chromosomen C07 und C09 (25% Varianzaufklärung) und ein QTL für Glucoraphanin auf dem Chromosom C07 (12% Varianzaufklärung) identifiziert. Die ermittelten QTL für Glucobrassicin und Glucoraphanin auf Chromosom C07 liegen auf dem gleichen Genabschnitt. Zusätzlich korrelieren die Abbauraten von verschiedenen Glucosinolaten, was ein Indiz dafür ist, dass ähnliche Faktoren innerhalb der Pflanzenmatrix den Abbau chemisch verschiedener Glucosinolate beeinflussen. Neben genetischen Effekten wird der Glucosinolatabbau von Umweltbedingungen beeinflusst, was für quantitativ vererbte Eigenschaften üblich ist. Die identifizierten QTL können zur Präzisionszucht (*marker assisted selection*) verwendet werden, wobei die Genotypen, welche zur Züchtung der nächsten Generation verwendet werden, anhand von molekularen Markern ausgewählt werden, die die QTL umgeben. Diese Art der Auswahl der Elterngeneration erhöht die Effektivität der Züchtung, da eine Analyse der genetischen Marker weniger arbeits- und kostenintensiv als die Bestimmung des Glucosinolatabbaus in vielen Genotypen einer Pflanzenpopulation ist. Die Vorrausetzung für die Präzisionszucht ist allerdings, dass die QTL unter verschiedenen Anbaubedingungen stabil sind, was anhand der durchgeführten Experimente nicht vorhergesagt werden kann. Weitere Untersuchungen sind notwendig, um die identifizierten QTL für den Glucosinolatabbau zur Präzisionszucht anwenden zu können, und um in Zukunft Gemüse mit einer erhöhten Stabilität von Glucosinolaten während der Lebensmittelzubereitung züchten.

Bisher wurde die Genomsequenz von *Brassica oleracea* nicht veröffentlicht, wodurch die Identifizierung von Genen, die den Glucosinolatabbau-QTL zugrunde liegen, erschwert wird. Außerdem ist eine Feinkartierung der genetischen Karte dieser Pflanzenpopulation notwendig, um kleinere genetische Regionen zu identifizieren und damit die Anzahl möglicher Gene zu verringern. Weiterhin ist der Mechanismus des Glucosinolatabbau nicht bekannt und bietet somit keinen Anhaltspunkt für die Auswahl von Kandiatengenen.

Als alternative Herangehensweise wurde eine Metabolomik-Analyse mit der *Random Forest Regression* kombiniert, um Metabolite, die mit den Glucosinolatabbaukonstanten korrelieren, zu identifizieren und neue Hypothesen über mögliche Abbaumechansimen

zu formulieren (**Kapitel 5**). Mit Hilfe dieser beiden Methoden konnten Metabolite, die mit den Glucosinolatabbaukonstanten assoziiert sind, bestimmt werden. Von 413 bestimmten Metaboliten waren 15 signifikant mit den Abbaukonstanten von Glucobrassicin, sechs signifikant mit den Abbaukonstanten von Glucoraphanin und zwei signifikant mit beiden Glucosinoaten assoziiert. Es wurde eine höhere Anzahl an Metaboliten, die mit beiden Glucosinolatabbaukonstanten assoziiert sind, erwartet, da ein gemeinsamer QTL auf Chromosom C07 identifiziert wurde. Von diesen 23 insgesamt Metaboliten wurden drei als Flavonole (1 Kaempferol- und 2 Quercetinderivate) und zwei als glucosinolate (Gluconasturtiin, 4-methoxyglucobrassicin) identifiziert. Die Flavonole waren positiv mit den Glucosinolatabbaukonstanten korreliert und erhöhen somit möglicherweise den Glucosinolatabbau durch Reduktion von Fe^{3+}- zu Fe^{2+}-Ionen, welche den Glucosinolat-abbau katalysieren. Ein Glucosinolat war negativ mit den Abbaukonstanten von Glucobrassicin korreliert und zwei Glucosinolate mit den Abbaukonstanten von Glucoraphanin. Diese negative Korrelation könnte auf einen kompetitiven Mechanismus hindeuten und somit den Abbau bestimmter Glucosinolate verlangsamen. Weitere 18 Metabolite konnten nicht identifiziert werden, wovon einige negativ und andere positiv mit den Glucosinolatabbaukonstanten korrelierten. Effekte von Metaboliten, die den Glucosinolatabbau möglicherweise steigern und andere Metabolite, die ihn möglicherwiese verlangsamen, verdeutlichen die Komplexität des Reaktionssystems. Weitere Studien sind notwendig, um diese Ergebnisse zu bestätigen und den Mechanismus des Glucosinolatabbaus genauer zu beschreiben.

Die QTL-Analyse von Metaboliten, die mit den Glucosinolatabbaukonstanten assoziiert waren, zeigte Metaboliten-QTL auf den Chromosomen C07 und C09 auf, welche auf der gleichen Genregion liegen wie die Glucosinolatabbau-QTL. Weiterhin wurde ein weiterer Metaboliten-QTL auf dem Chromosom C01 identifiziert, wo kein Glucosinolatabbau-QTL nachgewiesen wurde (**Kapitel 5**). Die Sequenzierung des *Brassica oleracea* Genoms könnte die Identifikation der Gene, die den Metaboliten-QTL und den Glucosinolatabbau-QTL zugrunde liegen, in der Zukunft erleichtern. Weiterhin kann die Kombination der Sequenzinformationen mit der Identität von Metaboliten die Auswahl von spezifischen Kandidatengenen erleichtern.

Der allgemeine Ansatzpunkt, um die Qualität von Lebensmitteln durch die Verknüpfung von Pflanzenzüchtung und Lebensmittelchemie zu verbessern, wurde in **Kapitel 6** diskutiert. Jüngste Fortschritte in der Entwicklung von Analysenmethoden mit hohem Probendurchsatz werden die Analyse notwendiger Probenanzahlen, welche die größte

Herausforderung darstellt, weiter erleichtern. Durch die gemeinsame Betrachtung von biochemischen Veränderungen innerhalb der Lebensmittelkette können beide Forschungsdisziplinen profitieren und damit möglicherweise die Qualität von pflanzlichen Lebensmitteln innerhalb der Lebensmittelkette in Zukunft verbessern

Schlussfolgerungen und Ausblick

Die im Rahmen dieser Arbeit durchgeführten Experimente zeigen, dass thermischer Glucosinolatabbau teilweise genetisch bedingt ist. Außerdem beeinflussen Umweltfaktoren, wie Anbausaison und -jahr, die Glucosinolatabbaugeschwindigkeit. Die Ergebnisse dieser Arbeit zeigen erstmals, dass *Brassica*-Gemüsesorten mit einer erhöhten Glucosinolatstabilität während der Lebensmittelzubereitung gezüchtet werden können. Weitere Studien sind notwendig um die Stabilität der genetischen Regionen zu prüfen, um diese effektiv in der Präzisionszucht anwenden zu können. Weiterhin wurde eine Herangehensweise erarbeitet, um Metabolite zu identifizieren, die möglicherweise den thermischen Glucosinolatabbau beeinflussen. Drei Flavonole, zwei Glucosinolate und 18 nicht identifizierte Metabolite waren mit dem Glucosinolatabbau assoziiert. Durch die Anwendung anderer Metabolomik- und Proteomik-Plattformen könnten weitere Stoffe identifiziert werden, die den Glucosinolatabbau beeinflussen könnten. Die Kombination von genetischen Regionen, Metaboliten und zukünftig der Genomsequenz von *Brassica oleracea* wird die Identifikation von Genen und Mechanismen des Glucosinolatabbaus erleichtern.

Die Verknüpfung von Pflanzenzüchtung und quantitativen lebensmitteltechnologischen Parametern ist eine herausfordernde und gleichzeitig vielversprechende Herangehensweise, um die Qualität von pflanzlichen Lebensmitteln zu verbessern. Sie bietet die Möglichkeit, den Nährwert von pflanzlichen Lebensmitteln zu erhöhen, indem Sorten gezüchtet werden, die nicht nur einen höheren Ausgangswert an sekundären Pflanzenstoffen besitzen, sondern diese auch während der Zubereitung aufrechterhalten.

ACKNOWLEDGEMENTS

IT IS DONE, with this thought I am writing the very last part of my thesis, looking back to four years I spent in a small town which is hard to find on a map, but where I met many nice people from all over the world and who shaped my personality. However, a PhD degree is not the work of one person, because without guidance, help and support from many people I could not have written this book.

I am very grateful to my supervisors **Guusje** Bonnema, **Ruud** Verkerk, **Matthijs** Dekker and **Tiny** van Boekel, who gave me the opportunity to work on this PhD project, who guided me through these years and who taught me a lot. I discovered that two disciplines, food technology and plant breeding, which do not seem be so far from each other speak different languages and have different ways to conduct experiments. This "special" project broadened my mind considerably and I experienced to work with a dataset from hundreds of plants and keeping the overview about the data. Writing was not my favourite type of work, but with your help every paper went better and I learned how to step back from the details and see the full picture. Thank you for always being there for advice when needed, but also giving me the freedom to take own decisions and become an independent researcher. I enjoyed a lot working with you on this project.

I want to express my appreciations to **Ric** de Vos who taught me how to conduct metabolomics experiments, how to treat the data and interpret the results, which was a very interesting experience for me. I thank **Chris** Maliepaard, who helped me in the statistical analysis of the metabolomics data and in other statistical questions related to the data in this thesis. **Bert** Schipper and **Jeroen** van Arkel are acknowledged for their technical assistance.

Of course, the members of the PDQ group (I will never get used to FQD) also contributed to the work accomplished in thesis. I am grateful for the help of **Jenneke** who prepared about 1000 samples for my study in chapter 4. Without your help my thesis had been delayed for some months. Furthermore, I acknowledge **Charlotte**, **Xandra**, **Frans** and **Geert** for their technical assistance in the lab. I also want to thank my students **Marloes** Roordink, **Martin** Hog and **Ruth** Sluimer for their contribution and the for the experience of being a supervisor. **Teresa**, **Irmela**, **Radhika** and **Probo** started at the same time in PDQ or shortly after me. We became a very nice "glucosinolate group" discussing problems and helping each other when the experiments and statistics did not what they were supposed to do, but also shared happy moments when things worked out or papers got finally accepted. Thank you for all your support, I enjoyed

a lot working with you, but also spending free-time together. **Radhika**, nice that you had similar working hours as me, it made the Biotechnion less scary in the evening. Thank you **Teresa** and **Radhika** for being my paranymphs and for helping me with the organisation of my defence and party. Before the "glucosinolate group" moved into one office, **Dieuwerke**, **Van Anh** and **Teresa** were my office mates. Thank you for the nice time we had, for the discussions and motivation during the difficult moments. **Van Ahn,** thank you for choosing me as your paranymph, it made me seeing the stage "from the other side" and get a feeling for my own defence.

PhD trips are a very nice tradition of the food science groups, thanks to the organisation of **Jenneke**, **Dieuwerke**, **Van Anh** and **Carol** we went in 2010 to Australia. It was a great experience to travel to the other end of the world and present my work there. The following PhD trip went to England in 2012, where I enjoyed being part of the organisation team together with **Teresa**, **Radhika**, **Daylan** and **Klementina**. I also thank **Sarn**, **Marielle**, **Lu Jing**, **Etske**, **Elsa**, **Lina**, **Fahui**, **Grace**, **Marine** and all other **PDQ employees** for the nice time we had at "labuitjes", Christmas dinners and for "gezellig" daily coffee breaks. Sharing the same floor with PDQ, the **Food Physics colleagues** contributed to an enjoyable time on a daily basis. Thank you **Harry** and **Elke** for organizing "borrels" and parties and thanks to **Elisabete**, **Silvia**, **Dilek** and **Yul** for joining them.

I spent also nice time in my "other" group, the ***Brassica* team** of the Laboratory of Plant Breeding. Although, I was never there full-time I felt as part of the team. Thank you, **Johan** for your help in the greenhouse and showing me the small tricks so that I don't lose the tiny *Brassica* seeds or mix them up. I appreciate the support of **Ningwen** and **Ram** who taught me the QTL mapping and with whom I had fruitful discussions about statistics. From the weekly presentations I learned a lot about plant breeding, but also how to give presentations to non- food scientists. Thanks to all the team members for their input to my work and the nice time at parties we had.

Many many thanks to **Guillaume**, who always had an open ear, motivating words and supported me with his programing skills. Thank you for "magic macro", without I would be still modelling my data, for graphs and statistics in R and for proof-reading this thesis, finding many typos.

Peter, **Lorena**, **Jule**, **Alex**, **Jeab**, **Gunnar**, **Bipna**, **Alejandra**, **Iliana**, **Jordi**, **Ellen**, **Lea**, **Anna**, **Ronald**, **Jacqui**, **Tila**, **Nicole**, **Mary** and **Kim**, thank you for your friendship! Spending time with you outside work was great and I have many nice memories of it. Thank you **Anna** for the discussions around my metabolomics experiments and help in statistics. Thank you **Iliana** for giving the final finish to my thesis by designing the cover and layout.

Nicht zuletzt möchte ich **meinen Eltern** danken, die mir meine Ausbildung ermöglicht haben um nun dort zu sein wo ich bin. Vielen Dank dafür! Außerdem danke ich **meine Eltern**, meiner **Schwester** und **meinen Freunden** für ihre Unterstützung und gut zu wissen, dass sie für einen da sind, wenn man sie braucht.

Thank you all, let's stay in touch and maybe work together in the future again!

-Kristin-

About the Author

Curriculum vitae

Kristin Hennig was born on 22nnd March 1985 in Dresden, Germany. She grew up in Dresden and finished high school in 2003. Subsequently, Kristin started her studies in Food Chemistry at the Dresden University of Technology. Besides her studies, she worked part-time as a student assistant at the Institute of Food Chemistry of the University in the years 2006 and 2007. In October 2007, Kristin went as a participant of the Erasmus program to Wageningen University (The Netherlands), where she conducted research for a six-months thesis project entitled "Fate of glucosinolates during mastication – study on broccoli" in the Product Design and Quality Management Group. Afterwards, she returned to the Institute of Food Chemistry at Dresden University of Technology, where she worked on another thesis topic entitled "Enzymatic deglycation of Amadori products by bacteria". After her graduation in January 2009 Kristin went back to Wageningen University to start her PhD studies in a collaboration of the Product Design and Quality Management Group (recently renamed to Food Quality and Design) and the Laboratory of Plant Breeding. The project investigated the degradation of the health promoting glucosinolates during food processing in different genotypes and tested the hypothesis that genetic regions can be identified for food processing parameters. She participated as well in teaching activities and supervised thesis students. Currently, Kristin is working as a post-doctoral fellow on trace analysis of steroid hormones at the Laboratoire d'Etude des Résidus et Contaminants dans les Aliments (LABERCA) in Nantes, France.

List of publications

Publications is peer-reviewed journals

Dekker, M.; **Hennig, K.**; Verkerk, R.; Differences in thermal stability of glucosinolates in five *Brassica* vegetables. *Czech Journal of Food Sciences*. **2009**, *27*, 85-88

Hennig, K.; Verkerk, R.; Bonnema, G.; Dekker, M.; Pitfalls in the desulphation of glucosinolates in a high-throughput assay. *Food Chemisty* **2012**, *134*, 2355-2361

Hennig, K.; Verkerk, R.; Bonnema, G.; Dekker, M.; Rapid estimation of glucosinolate thermal degradation rate constants in leaves of Chinese kale and broccoli (*Brassica oleracea*) in two seasosns. *Journal o. Agricultural and Food Chemistry* **2012**, *60*, 7859–7865

Hennig, K.; Verkerk, R., Dekker, M. & Bonnema, G.; Quantitative trait loci analysis of non-enzymatic glucosinolate degradation rates in *Brassica oleracea* during food processing. *Theoretical Applied Genetics* **in press**, DOI 10.1007/s00122-013-2138-1

Submitted publications

Hennig, K.; de Vos, R.; Maliepaard, C., Dekker, M.; Verkerk, R.; Bonnema, G.; A metabolomics approach to identify factors influencing glucosinolate thermal degradation rates in *Brassica* vegetables.

Hennig, K.; Verkerk, R.; van Boekel, M.A.J.S; Dekker, M.; Bonnema, G.; Plant science meets food science: A case study on improved nutritional quality by breeding for glucosinolate retention during food processing.

Conference abstracts and proceedings

Hennig, K.; Verkerk, R.; Dekker, M.; Bonnema, G.; Linking plant sciences and food sciences: Variation of thermal stability of glucosinolates in a segregating *Brassica* population. *28th International Horticultural Congress, Lisbon, Portugal; 22-26 August, 2010* (Poster presentation)

Hennig, K.; Verkerk, R.; Dekker, M.; Bonnema, G.; Linking Plant Sciences and Food Sciences: Quantitative genetic analyses of glucosinolate degradation during food processing in *Brassica oleracea. Euro Food Chem XVI, Gdansk, Poland;* 6 – 8 July, 2011 (Oral presentation)

Education Statement of the Graduate School
Experimental Plant Sciences

The Graduate School **EXPERIMENTAL PLANT SCIENCES**
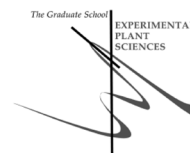
Issued to: **Kristin Hennig**
Date: **30 August 2013**
Group: **Laboratory of Plant Breeding & Product Design and Quality Management Group**
Wageningen University & Research Centre

1) Start-up phase	date
► **First presentation of your project (higly recommended)**	
Genetic Determinants of glucosinolate degradation during Food Processing	June 21, 2010
► **Writing or rewriting a project proposal**	
► **Writing a review or book chapter**	
► **MSc courses**	
Genetic analsis, tools and concepts	Nov-Dec 2009
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*7.5 credits**

2) Scientific Exposure	date
► **EPS PhD student days (highly recommended)**	
2nd joint International Retreat of PhD students in Plant Sciences, Cologne (Germany)	Apr 15-17, 2010
EPS PhD Student Day, Wageningen University	May 20, 2011
PDQ PhD student trip to Australia	Feb 04-20, 2010
PhD student trip with PDQ group to England	May 14-24, 2012
► **EPS theme symposia (highly recommended)**	
EPS theme 3 symposium 'Metabolism and Adaptation', Wageningen University	Feb 10, 2011
► NWO Lunteren days and other National Platforms	
► Seminars (series), workshops and symposia	
Qualitatitve work with the Trace DSQ GC-MS system, Interscience	Feb 20 & 22, 2012
► **Seminar plus**	
► International symposia and congresses (highly recommended)	
28th International Horticultural Congress, Lisbon (Portugal)	Aug 22-26, 2010
EuroFoodChem XVI, Gdansk (Poland)	Jul 06-08, 2011
Conference on Food metabolomics, Karlsruhe (Germany)	Oct 09-11,2011
► Presentations (highly recommended)	
Poster: 28th International Horticultural Congress, Lisbon (Portugal)	Aug 26, 2010
Oral: EuroFoodChem XVI in Gdansk (Poland)	Jul 07, 2011
Oral: 2nd PhD retreat of PhD students in Plant Sciences, Cologne (Germany)	Apr 16, 2010
► **IAB interview**	Feb 17, 2011
► Excursions	
Subtotal Scientific Exposure	*14.5 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses (higly recommended)**	
Reaction kinetics in Food Scinece	Oct 26-30, 2009
Advanced Food Analysis	Nov 15-19, 2010
Microscopy and Spectroscopy in Food and Plant Sciences	May 07-11, 2012
Nordic graduate school course on metabolomics	Aug 20-24, 2012
► **Journal club**	
► **Individual research training**	

Subtotal In-Depth Studies	*5.8 credits**

4) Personal development	*date*
► **Skill training courses**	
Supervising thesis students	Jun 04-05, 2009
PhD competence assesment	Oct 13, 2009
Project and time management	Mar 09, 23 & Apr 20, 2010
Techniques for writing and presenting a scientific paper	Jun 29-Jul 02, 2010
Career perspectives	Sep-Oct 2012
Scientific writing	May 11- Jun 29, 2011
► **Organisation of PhD students day, course or conference**	
Organisation of the PhD student trip of PDQ to the UK	Aug 2011-May, 2012
► **Membership of Board, Committee or PhD council**	

Subtotal Personal Development	*8.7 credits**

TOTAL NUMBER OF CREDIT POINTS*	**36.5**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

A credit represents a normative study load of 28 hours of study.