Project title:	Understanding the genetic and physiology controls of 'crumbly' fruit in red raspberry (*Rubus idaeus*)
Project number:	SF 167
Project leader:	Julie Graham – The James Hutton Institute
Report:	Final report
Previous report:	Annual report, 24th month
Key staff:	Luca Scolari
Location of project:	Invergowrie, Dundee, DD2 5DA, Scotland, UK
Industry Representative:	Louise Sutherland Freiston Associates Ltd.
Date project commenced:	October 2016

DISCLAIMER

While the Agriculture and Horticulture Development Board seeks to ensure that the information contained within this document is accurate at the time of printing, no warranty is given in respect thereof and, to the maximum extent permitted by law the Agriculture and Horticulture Development Board accepts no liability for loss, damage or injury howsoever caused (including that caused by negligence) or suffered directly or indirectly in relation to information and opinions contained in or omitted from this document.

AUTHENTICATION

We declare that this work was done under our supervision according to the procedures described herein and that the report represents a true and accurate record of the results obtained.

[Name]

[Position]

[Organisation]

Signature ..... Date .....

[Name]

[Position]

[Organisation]

Signature ..... Date .....

Report authorised by:

[Name]

[Position]

[Organisation]

Signature ..... Date .....

[Name]

[Position]

[Organisation]

Signature ..... Date .....

CONTENTS

GROWER SUMMARY 1

Headline 1

Background and expected deliverables 1

Summary of project and main conclusions 2

Financial benefits 3

Action points for growers 3

SCIENCE SECTION 4

Introduction 4

Materials and methods 11

Results 21

Discussion 40

Conclusions 49

Knowledge and Technology Transfer 52

References 53

Appendices 59

GROWER SUMMARY

Headline

- A genetic marker associated with the crumbly fruit condition has been identified for use in the UK raspberry breeding consortium at the James Hutton Institute.

Background and expected deliverables

Understanding the triggers of crumbly fruit in raspberry will allow breeding of more robust genotypes using molecular markers, modified agronomic practices based on project findings and new improved methodologies for plant health certification, leading to reduced wastage.

Raspberry fruit is an aggregation of multiple fertilized ovaries (drupelets) which in the crumbly fruit condition are abnormal. Crumbly fruit is a generic term used to describe the visual appearance of the abnormal fruit. This project identified two different forms of crumbly fruit, depending on the severity levels:

1. **Crumbly Fruit Condition (CFC)** – defined as plants where all fruits are symptomatic every season.
2. **Malformed Fruit Disorder (MFD)** - defined as plants that display crumbly fruit, but symptoms are intermittent within a season or across seasons.

The work in this project examined MFD which is both genetically and environmentally controlled. It is caused by a partial failure in one or more physiological processes concerned with fruit development, which results in the drupelets not forming properly. This is an increasing problem for the global raspberry industry.

The research in the project was based on a hypothesis that the growth of the fertilised ovaries is regulated and synchronised by a hormonal coordination process which takes place in the receptacle of the flower. The receptacle acts as a hub which controls molecular signals (plant hormones) and establishes an intense hormonal crosstalk with the fertilised ovules that helps to coordinate the fruit development.

This work aimed to study MFD through different approaches.

- A study of the differences in gene expression between crumbly and normal fruit in the progeny of a mapping population (Glen Moy x Latham) known to segregate for the crumbly fruit condition across different seasons (Graham et al. 2015).

- A study of the hormone regulation behind fruit development in artificially induced ‘crumbly’ fruits.
- A study of flower damage by bees and excess nectar from flowers preventing pollination, to determine if this may be playing a role in crumbly fruit development.

The work aimed to deliver improved understanding of what causes Malformed Fruit Disorder (MFD) and how growers might be able to overcome it in their production process. It was also hoped to identify molecular markers associated with crumbly fruit, which could be utilised in the UK raspberry breeding consortium at the James Hutton Institute to breed new cultivars with resistance to Crumbly Fruit Condition (CFC).

Summary of project and main conclusions

In studying differences in gene expression between crumbly and normal fruit, the project identified differences in gene regulation in fruit development. As a result, a genetic marker associated with the crumbly fruit condition has been identified for use in the raspberry breeding programme.

Studies of the hormone regulation allowed differences in hormone control of fruit development between crumbly and normal fruit to be identified. This could allow the development of an on-farm strategy involving hormone sprays to reduce crumbly fruit. Interfering with the regular growth of the fruit due to damage to the receptacle was shown to induce the ‘crumbly fruit’ condition producing artificial misshapen crumbly like berries.

Studies on raspberry plants grown under glasshouse conditions, with the assistance of commercial beehives to increase pollination efficiency, highlighted an interesting phenomenon. Excess floral nectar not collected by bees was observed to drip over the flower carpels (made up of the stigma, style and ovary). These became impregnated with the sticky substance and could not be pollinated. Bees are the major insect pollinator used for commercial purposes with bumble bees (*Bombus ssp*) and honeybees (*Apis mellifera*) dominating the market. Bumble bees are recommended for greenhouse-grown plants because they are better adapted to closed spaces (Andrikopoulos and Cane 2018). To further encourage bumble bees to focus exclusively in collecting pollen, commercial hives are equipped with small tanks containing artificial nectar to supply the bees at any time with this substance so they do not need to collect nectar and can concentrate on foraging only pollen. This particular hive design was specifically studied to increase pollination efficiency. In contrast to bumble bees, honeybees look for nectar when foraging.

Further investigation is required on this topic. However, the scientists suggest that it may be prudent to combine bumble bees and honeybees when pollinating commercial raspberry flowers to exploit their differing foraging behaviour. This could not only increase pollination efficiency but potentially contribute to a reduction in misshapen fruit caused by excessive nectar impregnating flower carpels.

Financial benefits

This project is unlikely to deliver any immediate financial benefit to growers. However, using our knowledge of the molecular marker for Crumbly Fruit Condition (CFC) in the raspberry breeding programme at the James Hutton Institute, will lead to the release of raspberry cultivars which do not suffer from the condition, leading to higher grade-out of Class 1 fruit in commercial raspberry plantations.

The potential to improve our knowledge and management of different bee pollinators may also help reduce the levels of Malformed Fruit Disorder (MFD) in future production systems.

Action points for growers

- Look out for the release of new raspberry cultivars in future which are known to be free from Crumbly Fruit Condition (CFC).
- Consider combining use of both bumble bees and honeybees to improve efficiency of raspberry flower pollination. This may also help to reduce the level of Malformed Fruit Disorder (MFD), although further study is required to confirm this hypothesis.

SCIENCE SECTION

Introduction

Raspberry fruits are formed from an aggregation of multiple fertilized ovaries each of which is referred to as a drupelet as it becomes fleshy. A drupe fruit is defined as one which develops entirely from a single ovary, and Rubus fruit are aggregates of drupelets formed by the ripening together of a number of ovaries all from the same flower and adhering to a common receptacle: in a sense each drupelet is a complete fruit in itself and a miniature homolog of such drupe fruits as the cherry, plum or peach. It is thought that 'crumbly' fruit is caused by a partial failure in one or more physiological processes concerned with fruit development, which results in the drupelets not forming properly. This is an increasing problem for the global raspberry industry. There have been a number of different causes suggested for the condition. It is known that infection with certain viruses (i.e. Raspberry Dwarf Virus, Raspberry Leaf Mottle Virus and Raspberry Latent Virus), alone or in all possible combinations, are responsible for very severe forms of crumbly fruit (Jennings 1988; Quito-Avela et al. 2014). Raspberry Bushy Dwarf Virus (RBDV) affects pollen reducing its capacity to induce fruit-set and can lead to failure of almost half of all drupelets to set (Murant et al. 1974; Daubeny et al. 1978). A genetic cause has been demonstrated where the crumbly phenotype arises from virus-tested mother plants (Jennings 1988). Studies have also shown that extensive tissue culturing of plants may increase the emergence of the condition (Jennings 1988). Additionally, the environmental conditions appear to play an important role with variations in the extent of crumbliness apparent from year to year (A. Dolan pers. comm.).

Jennings (1967b) stated that 'crumbly' fruit is an indication of a partial failure in one or more physiological processes concerned with fruit development. It is clear that a better understanding of the physiology and molecular processes behind fruit development is required to help develop a strategy for controlling, or ideally eradicating, 'crumbly' fruit and this forms the subject of this work.

One hypothesis in this work concerns a hormonal coordination process regulating and synchronizing the growth of all the fertilized ovaries. Without this system the first fertilized ovaries, starting to grow earlier would create a size gap with those fertilized later that will be never filled up, resulting in fruits misshapen and/or crumbly. The centre of this regulatory process is postulated to be the receptacle, that acting as a hub by means of molecular signals (e.g. plant hormones), establishes an intense hormonal crosstalk with the fertilized ovules that helps to coordinate the fruit development. This allows the last ovules to be fertilised to start growing simultaneously with those that received the pollen earlier. In simple terms, the receptacle acts like a switchboard operator taking note of all the calls it receives immediately

after an ovary gets fertilised. In this manner, soon after the receptacle has received a number of “calls/messages” equalling the minimum number required to develop a proper fruit, it contacts all the fertilised ovules at exactly the same time and with its signal triggering their simultaneous growth. No published scientific papers are available to support this hypothesis and a range of experiments have been specifically designed to address this.

Transcriptomic analysis

Despite botanical differences, all fruits undergo similar developmental steps of fruit set, growth, maturation and ripening. Fruit set is the first stage of development after fertilisation. It is followed by an active cell division and expansion phase, called growth during which the fruit attains its maximum size. This in turn is followed by ripening, the last of its development stages (Kumar et al. 2014). The current model of regular fruit set implies that ovary growth is blocked before pollination and that auxin is a key regulator of ovary growth de-repression at fruit set. Other phytohormones have been shown to have a role in fruit initiation and development (i.e. gibberellin, cytokinin, brassinosteroids, ethylene and abscisic acid) (Goetz et al., 2007, Pandolfini et al. 2007).

Plant hormones with their unquestionable role (Ozga and Reinecke, 2003, Pandolfini et al. 2007, Nagpal et al. 2005, Kumar et al. 2014, Vanstraelen and Benkova, 2012) in regulating all the processes related to fruit development, as well as the molecular processes behind the development of the flower through fertilization of the ovules to fruit development are all important aspects of study in raspberry because they could potentially bring to light the mechanisms responsible for the formation of ‘crumbly’ like misshapen fruit in red raspberry.

Transcriptomics technologies are techniques that are used to study an organism transcriptome (sum of all RNA transcripts) allowing gene expression analysis to identify genes whose expression levels vary due to specific treatments, different developmental stages, between different tissues etc. RNA-microarray technology is a tool containing thousands of DNA fragments (probes) of known sequence arrayed on a chip. This easily allows the collection of thousands of data points related to levels of gene expression between samples in formats that can be used for bioinformatic and statistical analysis (Waters et al. 2005). Microarray analysis has been applied with success to gene expression analyses of different plant species (i.e. tomato, strawberry, peach, pear and grape) during fruit development, in particular fruit ripening, allowing the identification of genes specifically involved in different stages of fruit growth (Waters et al., 2005). It constitutes a relatively easy and convenient way to select, for instance, potential genes of interest for breeding programmes (Slater et al. 2008).

'Crumbly' Quantitative Trait Loci (QTL)

Understanding genetics is key to improving our knowledge of many aspects of plant biology and has advanced greatly through an understanding of the principles of heredity, with Mendelian genetics forming the basis of plant breeding. Together with an understanding of biometrical genetics which applies to traits showing continuous variation and controlled by more than one gene, this has allowed significant understanding of quantitative traits (Mackay 2001; Rajon and Plotkin 2013). Using this knowledge, understanding plant trait control has been made possible by the development of genetic linkage maps where markers linked to the gene(s) or quantitative trait loci (QTLs) underlying a trait can be identified.

For map construction, individual marker loci are identified and genetically characterized in a segregating population from a single cross and the recombination rate of alleles at each pair of loci can be determined using classical linkage analysis. These marker loci represent DNA polymorphisms between different individuals. These polymorphisms can be of many different types from single nucleotide changes, large or small insertions and deletions or length variation in repeat sequences.

SNPs are the most abundant markers in a genome and can be used for many different genome analyses (He at al. 2014). These marker loci can be applied to a biparental cross and ordered into a linkage map and distance between loci can be expressed as recombination units given in centiMorgans (cM) where one cM is equal to 1% recombination. Once a sufficient number of markers have been mapped, the number of linkage groups (LGs) should equal the haploid number of chromosomes. Computer programs are available to quickly generate a map once markers have been applied to a segregating population (e.g. JoinMap software by Kazyama®).

Compared to other crops, there have been few genetic and genomic resources available for *Rubus* until recently when the development of high-density genetic maps and markers and other genomic resources have become available.

The James Hutton Institute played a role in contributing to the development of new genomic resources, for instance creating a database that provides access to the predicted genes from an assembly of whole-genome-shotgun sequence from raspberry (cv. Glen Moy). The genes were predicted from the mapping of RNA-sequencing data from twenty-two varieties of raspberry to the genomic assembly (http://camel.hutton.ac.uk/raspberry/). For reviews of markers, linkage maps and QTL developments in *Rubus* see (Graham et al. 2009; McCallum et al. 2018; Foster et al. 2019).

The developments of Genotype by Sequencing (GbS) mapping has allowed high density maps linked to genome scaffolds to be available for trait dissection (Foster et al. 2019). The

new generation of sequencing technologies allows high-throughput at lower costs; moreover, due to their accuracy and simplicity they are becoming commonly used in both bi-parental mapping and genome wide association studies (GWAS). The GbS technology is quite straightforward when it comes to small genomes while for large ones, enrichment or restriction strategies must be applied to guarantee enough overlap of sequence coverage. While enrichment strategies are time consuming and extremely expensive for large genomes, the reduction of the genome complexity with restriction enzymes is easy, site specific, reproducible and facilitate to access genome regions inaccessible to sequence capture techniques. Above all, when appropriate restriction enzymes are selected, repetitive regions of the genome as well as lower copy regions are easily avoided simplifying the computational alignment in species with high level of genetic variation (Elshire et al. 2011).

The use of GbS to identify new SNP markers and construct new linkage maps is becoming increasingly common in genetic studies especially for analysis of quantitative trait loci (QTLs); these are segments of genome that show statistical significant association with quantitative traits (Moose and Mumm 2008). GbS analysis proved to be successful in the identification of segregation distortion in raspberry, allowing the detection of deleterious alleles responsible for inbreeding depression in *Rubus idaeus* (Ward et al. 2013).

In terms of ‘crumbly’ fruit, previous QTL mapping of the ‘crumbly’ fruit trait in the Latham x Glen Moy population revealed two QTLs important for the genetic control of the condition located on linkage group 1 and 3, (LG1) and (LG3) see Graham et al. (2015). Contrary to the suggestion by (Jennings 1976b) that ‘crumbly’ fruit was related to the gene H region, no genetic association with this region on LG2 could be identified with the ‘crumbly’ fruit syndrome (Jennings 1967). However, there was an association with ripening, with the longer the fruit takes to achieve fruit set and reach the green fruit stage, the more likely it is to be crumbly. This may explain the association hypothesized by Jennings as the Hh genotype of gene H is associated with a slowing down of ripening across all stages from open flowers to the green/red stage compared to the hh genotype (Graham et al. 2009).

Hormones profiling

Plants have evolved many different strategies to cope with the challenging environmental conditions they encounter (Wasternack and Kombrink 2010). Plants cannot move to escape stress factors whether abiotic or biotic, and for this reason they have evolved a series of traits that allow them, for instance, to regenerate damaged organs and tissues and/or to re-direct growth in response to external stimuli (Vanneste and Friml 2009).

Small endogenous signalling molecules, called phytohormones (plant hormones), are responsible for the control and coordination of the physiological processes that plants activate to react to external environmental factors, as well as to regulate their growth (Santner et al. 2009; Wasternack and Kombrink 2010). Such compounds act locally at very low concentrations, at or near the site of synthesis or even in distant tissues (Santner et al. 2009). In general plant hormones act through extensive crosstalk between themselves and/or other signalling pathways and the results of these interactions are additive, synergistic or antagonistic actions that determine specific and complex physiological outcomes (Vanneste and Friml 2009; Pan et al. 2008).

Phytohormones, according to their structure and function, are divided into nine classes: abscisic acid (ABA), auxins (AUXs), cytokinins (CKs), gibberellins (GAs), jasmonates (JAs), salicylates (SAs), brassinosteroids (BRs), strigolactones (SLs) and ethylene (Santner et al. 2009; Zwanenburg et al. 2016). Although each class of phytohormones is linked with specific and typical biological effects (e.g. SA for plant defence, GAs for organ elongation and flowering time, CKs for germination, etc.), the biological processes are regulated by complex networks involving different hormonal signals (Cao et al. 2016). In plants, hormone interactions can occur at least at two different levels, hormones distribution (i.e. the opposing action of AUXs and CKs during lateral root initiation) and gene expression (i.e. AUXs and BRs repress the same genes suggesting coordination between the two signalling pathways) (Santner et al. 2009). The activity of plant hormones depends on their availability which is in turn affected by their metabolism, localisation, transport and signal transduction; the modulation at any of these levels, and there are myriads of possible combinations, can determine different physiological processes (Simura et al. 2018).

LC-MS, especially in the last decade, has become a valid technique to trace plant hormones because of higher sensitivity/lower detection limits that can be reached with these systems (Cao et al. 2016). There are many different LC-MS techniques available for accurate analyte quantification but LC-triple quadrupole mass spectrometry, although featuring lower resolution compared to the time-of-flight-MS or the ability to screen for unknown compounds like the orbitrap-MS, represents the best choice when it comes to detection/quantification of target compounds by virtue of its greater sensitivity, repeatability and a wider dynamic range (Cao et al. 2016).

One of the main challenges for future work is the network plus crosstalk of the hormonal circuits underlying the whole fruit development process. An analytical method designed to detect six groups of plant hormones (i.e. auxins, cytokinins, gibberellins, jasmonates, salicylates and abscisic acid), although only partially covering the whole spectrum of plant hormones (nine classes in total), represents a valid starting point to help understand the

metabolic pathways that regulate fruit growth in raspberry undergoing normal developmental processes compared with fruits induced to express the ‘crumbly’ fruit phenotype by mechanical damage to the flowers before pollination. The majority of data available in the scientific literature indicate the involvement of different hormones in the regulation of fruit development. Although the concerted action of auxins and/or gibberellins and/or cytokinins, through their biosynthesis and/or signalling pathway, seems to play a major role, on the basis of the high complexity of the whole process, the involvement of other phytohormones cannot be excluded a priori. In the raspberry ‘crumbly’ fruit, where berry abnormalities occur, different hormonal pathways might be involved in the regulation of these misshapen fruits and then a larger spectrum of different plant hormones must be taken into account to better understand the whole process. An analytical method designed to detect eighteen different plant hormones, covering six groups represents a good starting point to help understand the hormonal molecular circuits behind the fruit development.

In this work the use of microarray analysis is described to identify variations in gene expression in samples of red raspberry from three different developmental stages (i.e. closed bud, open flower and green berry) and for two different phenotypes (i.e. mostly and never ‘crumbly’).The study focused on the expression levels of predicted *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology annotations related to: flower development, hormone, pollen and transport (https://www.geneontology.org/). The choice to target primarily these four specific factors was driven by suggestion in the literature that in some varieties (e.g. Sumner) the ‘crumbly’ fruit phenotype might be caused by mutations producing the homozygous state for two recessive pairs of genes that retard the development of the embryo sac and reduce the production of fertile pollen (Jennings 1988). This suggested that those genes having functions that affect pollen production and functions such as those involved in flower development must be considered since they might be directly involved in processes and/or functions potentially causing ‘crumbly’ fruit. Hormones are the main plant growth regulators and thus their involvement cannot be excluded. Gene ontology terms related to transport were chosen also because, according to the hypothesis behind this work, a crosstalk between receptacle and fertilized ovaries would be needed to guarantee the synchronised growth of all the ovaries that form the final fruit. Without such a regulating process, mediated by the receptacle, the late fertilised ovary would never be able to reach the same size of those fertilised earlier with potential consequence of misshapen fruits.

The probes of the ‘crumbly’ microarray experiment (see chapter 3 for details) that mapped inside the ‘crumbly’ QTLs identified using the GbS map (Hackett et al. 2018) were selected and analysed. The idea was to identify potentially interesting genes related to ‘crumbly’ fruit whose relevance would have been strengthen by their location inside a genome region

already tightly associated to the ‘crumbly’ fruit phenotype such as the ‘crumbly’ QTLs. The aim was to try to find putative gene markers and in fact the new GbS linkage map allowed the identification of new and more significant ‘crumbly’ markers associated with the ‘crumbly’ QTLs. Moreover, the new GbS linkage map was used by Hackett et al. (2018) to re-analyse the fruit ripening scores (Graham et al. 2009) and identify QTLs associated to the ripening process. In this work, putative association between ‘crumbly’ fruit and fruit ripening were studied by analysing the functions of those predicted *A. thaliana* genes orthologs of the *R. idaeus*’ ones matched by the microarray probes mapped inside the overlapping regions of ‘crumbly’ and fruit ripening QTLs. The extension of this project focused on identifying potentially interesting genes related to ‘crumbly’ fruit whose relevance would have been strengthen by their location inside a genome region already tightly associated to the ‘crumbly’ fruit phenotype such as the ‘crumbly’ QTLs. In this work, data is presented on the genetic basis of the ‘crumbly’ phenotype using the GbS map of Hackett et al. (2018) to re-analyse the previous data from Graham et al. (2015). The aim was to try to find putative gene markers and in fact the new GbS linkage map allowed the identification of new and more significant ‘crumbly’ markers associated with the ‘crumbly’ QTLs. Eight of these markers were selected and subjected to a Genome Wide Association Study (GWAS) to try identifying one or more marker/s strongly associated with ‘crumbly’ fruit across a wider gene pool that could be used for molecular markers assisted breeding and for diagnostic purposes.

Plant hormones are the main growth regulators, they are indispensable for all the aspects related to plants life cycle. A better understanding of the leading phytohormones for what concerns the regulation of fruit growth and development could help to explain potential causes of ‘crumbly’ fruit in raspberry. In this work, a simple analytical method was developed to detect 18 phytohormones representing six of the most important classes of plant hormones: abscisic acid (ABA), salicylates (benzoic acid BA, cinnamic acid CA, salicylic acid SA, methyl benzoate MeBA, methyl cinnamate MeCA and methyl salicylate MeSA), jasmonates (jasmonic acid JA, methyl jasmonate MeJA and 13-epi-12-oxo-phytodienoic acid OPDA), auxins (indole-3-acetic acid IAA, indole-3-butyric acid IBA, Indole-3-carboxylic acid ICA and methyl indole-3-acetate MeIAA), gibberellins (gibberellic acid A1 GA1, gibberellic acid A3 GA3, gibberellic acid A4 GA4) and cytokinin (zeatin). The method was developed to allow the analysis of the raspberry fruit extract at two different stages, green berry and red berry both in the receptacle and in the drupelets to try to verify the hypothesis that fruit development is coordinated and regulated by the receptacle acting as a leading hub, synchronizing the growth of the many fertilized ovaries that will constitute the fruit drupelets. Samples of ‘crumbly’ induced fruit (flowers mechanically damaged) and control (normal developed fruits) were collected. Drupelets and receptacle for each berry were separated with the aim of

verifying which phytohormones were primarily involved in the fruit development and what were, if any, the differences between artificial ‘crumbly’ fruit and a normal fruit (control) and how, if the hypothesis was correct, the receptacle fulfilled its function of raspberry fruit regulator.

Materials and methods

RNA isolation and quality assessment

RNA extraction is complicated from samples like raspberry containing high levels of polysaccharides, polyphenols and protein contaminants (Jones et al., 1997). RNA was isolated from three different parts (i.e. closed bud, open flower and green berry) of full sib family generated from a cross between the European red raspberry cv. Glen Moy and the North American red raspberry cv. Latham. Minimum 70 to maximum 90 mg of plant material were weighed and then fine ground with mortar and pestle using liquid nitrogen. The extraction/purification was performed using the RNeasy Plant Mini Kit (Qiagen, Germany) and with the RNAse free and DNAse I Set (Qiagen, Germany), following the manufacturer’s instructions. RNA quality and its concentration were estimated spectrophotometrically from all samples using a NanoDrop™ spectrophotometer (NanoDrop, USA). Before proceeding with the microarray analysis, the quality of the RNA samples was further examined using the RNA Integrity Number (RIN) algorithm with a 2100 Bioanalyzer system (Agilent, USA) that analysed, after electrophoretic separation, the ribosomal RNA of the samples. Essentially this instrument gave a more reliable and accurate measure of RNA integrity; RIN values of ten stated intact RNA while RIN values of 1 indicated completely degraded RNA (Schroder et al., 2006). The RIN range values of the raspberry samples were between 8.90 to 9.90 indicating a very good quality RNA material.

Microarray experimental design and data analysis

The *Rubus idaeus* microarray was developed using the Agilent platform (Agilent Technologies, USA). The microarray contains in total 55,708 single 60-mer oligonucleotide probes representing unique transcripts. The samples analysed consisted of 14 individual progeny from a Glen Moy x Latham mapping population of which seven were labelled as mostly ‘crumbly’ phenotype since they consistently produced, across many years of scoring, primarily fruits with uneven shape (though not in every season thus the mostly crumbly designation), while the other seven plants were labelled as never ‘crumbly’ phenotype since they always produced regular shape fruits.

For each plant and phenotype, three different development stages were examined, closed bud (CB), open flower (OF) and green berry (GB). Four biological replicates were collected for each stage giving a total of 168 samples from which the RNA was extracted and processed. These 168 samples were pooled according to their phenotype, stage and biological replicate in order to form 24 pools in total. In practice the pools were named as A, B, C and D for the four biological replicates, so for instance all the 7 biological replicates A for the mostly 'crumbly' were pooled together as were the 7 biological replicates A for the phenotype never 'crumbly'; the same was done for each of the three stages and for the other three biological replicates. This specific design was chosen in order to reduce the effect of the environment and genetic differences not associated with the crumbly trait on the expression level. Although this created an artificial level of expression, the ranking between the genes remained unchanged as the pooling effect was the same for all the genes.

In total 24 microarrays were processed, data were extracted using Feature Extraction software (Agilent Technologies, USA) and then imported into GeneSpring software (version 7.3) (Agilent Technologies, USA) for data pre-processing and normalization. To make sure labelling differences were considered, the 1-colour normalization with the default setting was performed. All the probes with signal indistinguishable or too small to be significantly different from the background signal were removed from the data.

Heatmap tree clustering

The analysis was performed using GeneSpring software (v. 7.3) (Agilent Technologies, USA) using similarity measure of Pearson's correlation with clustering algorithm of average linkage. The analysis was then focused only on a limited number of probes (107 in total), those matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had specific gene ontology annotations related to: flower development, hormones, pollen and transport.

Gene Ontology (GO) term annotation

The probes differently expressed between the two different phenotypes (i.e. mostly and never 'crumbly') were first analysed through the Glen Moy genomic assembly browser (http://camel.hutton.ac.uk/raspberry/) to identify the correct matching *Rubus idaeus* gene. For each identified *R. idaeus* gene, the browser reported the corresponding *Arabidopsis thaliana* ortholog gene and to make sure the correspondence of the genes between the two different species was correct, the full length genomic sequence (https://www.arabidopsis.org) was copied and then blasted in the Glen Moy genomic assembly browser to ensure that the

selected *A. thaliana* was the correct ortholog of the *R. idaeus* corresponding gene. The selected genes were then analysed through (https://www.arabidopsis.org/tools/bulk/go), the *Arabidopsis thaliana* online browser that display in detail all the Gene Ontology (GO) annotations done to each selected gene.

'Crumbly' QTLs identification and their mapping on Glen Moy x Latham linkage groups; Re-analysis of 'crumbly' data on new GbS map

The crumbly data from the Latham x Glen Moy population (Graham et al. 2015) was re-analysed on the GbS map as described for ripening traits in Hackett et al. (2018) and here data from the 'crumbly' fruit microarray experiment were integrated in the analysis. The 'crumbly' data QTL positions were identified using a hidden Markov model (HMM) approach adapted from similar work for QTL mapping in autotetraploid species (Hackett et al. 2013), as initial QTL mapping using interval mapping in MapQTL 5 (Van Ooijen, 2004) gave logarithm of the odds (LOD) profiles that were unexpectedly irregular, given the high-density map, resulting in uncertainty in locating the peak LOD score. This was due to the differences in the number of parental heterozygous markers (Hackett pers. Comm.) thus HMM was employed. The hidden Markov model utilising all the marker information on a chromosome allowed the derivation of genotypic probabilities at each position, producing more accurate and precise peaks with smoother profiles (Hackett et al. 2018).

Identification of differentially expressed genes within 'crumbly' QTLs

All probes from the microarray were located onto the new GbS map (Hackett et al. 2018) to look for co-location with the 'crumbly' QTLs. A list with all the probes, mapped inside the three 'crumbly' QTLs, the two identified by Graham et al. (2015) on linkage groups (LGs) 1 and 3 and the new QTL again LG3 identified here, was produced by using the proprietary *Rubus idaeus* genome browser at The James Hutton Institute (http://camel.hutton.ac.uk/raspberry/index.html) to find the probe position within the genome scaffolds. An orientation file that identified the position of genome scaffolds for each linkage group was then utilised. A total of 1,375 probes, 307 from the QTL, identified in this work, on LG3, 851 and 216 probes respectively from the original 'crumbly' QTLs on LG1and LG3 identified by Graham et al. (2015) were identified. The selected probes were subjected to analysis of variance (ANOVA) with a phenotype (i.e. mostly and never 'crumbly) per stage (i.e. closed bud, open flower and green berry) interaction design. A very stringent threshold (p-value <0.001) to select the probes differently expressed was applied. The analysis was

performed in GenStat (VSN international, UK) to identify those showing differential patterns of expression.

The 165 probes selected as being differentially expressed between the two phenotypes (i.e. mostly and never 'crumbly') were used to find the predicted genes from the JHI Rubus genome data base. The Glen Moy genome browser provides for each probe, its length and its position, in base pair (bp), along the scaffold and if it was located within any of the *Rubus idaeus* genes. For each scaffold, the database provides, within other information, the sequence that was imported into the DNA sequence analysis software Sequencher® version 5.4.6 (Gene Codes Corporation, Ann Arbor, MI USA). The software displays the exact position of each base along the scaffold and so the region corresponding to the *R. idaeus* gene of interest could be easily identified. The gene sequence was then examined using blast into The Arabidopsis Information Resource (TAIR) database (https://www.arabidopsis.org/) to identify the best *Arabidopsis thaliana* ortholog gene for the *Rubus idaeus* selected one. Once the *A. thaliana* gene was identified, its sequence, available from TAIR database was then blasted back into the Glen Moy browser to double check the validity of the selected *Arabidopsis thaliana* gene. The predicted ortholog *A. thaliana* genes were then analysed with the gene ontology (GO) annotations function of The Arabidopsis Information Resource (TAIR) database.

'Crumbly' fruit induction experiments

Glen Ample long canes were purchased from EU plants Ltd. (Abingdon, UK). Plants were cultured in controlled environment, plant growth room (Nijssen, The Netherlands) and for the first two weeks the following parameters were set: minimum temperature 10°C, maximum temperature 14°C, relative humidity 70% and daylight length 16 hours (from 7:00 AM to 11:00 PM); such conditions helped the plants to acclimatize. After the two weeks of acclimatization, all the environmental parameters remained unchanged except the maximum temperature that was increased by two degrees, from 16 to 18°C. The first flowers, immediately after anthesis, were collected and dried at room temperature for two days to assist in pollen maturity. The dried flowers were closed in Petri dishes and stored in the fridge; they were used as pollen reservoir for the 'crumbly' induction experiments. Once the plants were established under these conditions, all unopened flowers one day from anthesis were emasculated. The emasculations were performed according to the method available at The James Hutton Institute (N. Jennings, James Hutton Ltd. Raspberry breeder pers. comm.). All flower buds were emasculated by means of a scalpel blade by simply following the contour of the five sepals, in this manner: petals, sepal and stamens were cut away. After the two days required

for the stigmas to become receptive to pollen, the emasculated flowers were first damaged and then hand pollinated. The damage, carried out on the receptacle, was accomplished by pinning its tip and side with a needle and for each damaged flower (treatment) a control was carried out by simply hand pollinating the emasculated flower. Each treated bud (damaged receptacle and control) was tagged reporting the date of the emasculation, the kind of treatment (damaged receptacle or control) carried out and the number of the plant.

Fruits derived from the crumbly fruit induction experiments (i.e. damaged receptacle and control), were collected at two different stages, green and red berry and immediately frozen in liquid nitrogen. They were then stored at -80°C to be used for the hormone profiling analysis.

Phytohormones analysis

Chemical analysis of plant material (receptacle and drupelets) was performed on an Agilent 1260 high performance liquid chromatography (HPLC) system consisting of quaternary pump, a Diode Array Detector (DAD), a Temperature Control Device and a solvent Thermostat module (Agilent Infinity 1290) coupled to an Agilent 6460A Triple Quadrupole Mass Spectrometer (Agilent Technologies, Santa Clara, CA, USA). The sample extract (5 µL) was injected onto a 150 x 2 mm (5µm) Gemini RP C18 (110 Å) column fitted with a Gemini C18 4 x 2 mm Security guard cartridge (Phenomenex, Cheshire, UK). Samples were eluted at a flow rate of 0.3 ml min^{-1}, using a gradient separation with two mobile phases A = 0.1% formic acid in deionized water and B = 0.1% formic acid in methanol. The elution gradient lasted 29 min in total and was as follows: A/B 95/5 (v/v) hold for 2 min, ramped up to 35% B in 1 min, followed by further ramping up from 35% to 55% in 15.5 min and further ramped up from 55% up to 100% in 3 min and hold for 2 min. Within 0.5 min the gradient was returned to the initial composition of 5% and held for 5 min.

Mass detection was carried out in negative or positive ion mode depending on the phytohormone (Tables A1 and A2 in appendix) using a jet stream electrospray ionization (ESI) interface coupled to the triple quadrupole system (Agilent, USA). For ESI, the gas temperature, gas flow, nebulizer pressure, sheath gas temperature, sheath gas flow, capillary cap and nozzle voltage were optimized with the help of Agilent Source Optimizer Software and set to 350°C, 10 L min-1, 30 psi, 4000°C, 11 L min-1 and 3500 V (both ion modes). Collision energies and fragmentor voltages for transition states of the 18 phytohormones were optimized using Agilent Optimizer Software (Agilent Technologies, USA). Hereby the most sensitive transitions (i.e. transitions with the highest intensity of the product ions were chosen to build the final multiple reactions monitoring (MRM) method. Forty-five transitions were part

of the MRM mode, each with a dwell time of 10 ms and a delay time of 3.5 ms, leading to a total cycle time of 445.5 ms and thus 2.5 Hz (2.5 cycles per second). Peak areas of the 18 phytohormones and of the nine isotopically labelled internal standards were integrated with Agilent MassHunter Quantitative Software (Agilent, USA; version B.07.00).

Statistical analysis

The 'crumbly' microarray data, after filtering, were normalised (i.e. transformed in logarithm on base 2), the probes were then subjected to statistical analysis (i.e. two-way ANOVA) with stage (i.e. closed bud, open flower and green berry) and phenotype (i.e. mostly and never 'crumbly') as factors, the analysis was performed using GeneSpring software (version 7.3) (Agilent Technologies, USA) to identify genes with significant differential expression dependent on the two factors and their interactions. A stringency threshold, p-value ≤0.05 and the Benjamini & Hochberg multiple correction test was applied for the analysis and from these, 827 probes exhibiting significant changes in the expression levels were identified.

The analysis of the variance for those probes mapped inside the three 'crumbly' QTLs was performed in GenStat 18th edition (VSN International, UK). The ANOVA with stage (i.e. closed bud, open flower and green berry) per phenotype (i.e. mostly and never 'crumbly') interaction was effectuated for verifying the significative differences, at 99.9% confidence levels, in the expression levels of the 'crumbly' microarray probes mapped inside the three 'crumbly' QTLs.

For the metabolomic analysis (hormones profiling), to assess whether or not the two different kind of treatments (i.e. damage of the receptacle and control) were significantly different, and because the artificially induced 'crumbly' fruits were the number of fruits that crumbled out of those treated, a generalised linear model with a binomial distribution was performed (GenStat, VSN International, UK). All the statistical analysis of the metabolomic measurements (i.e. ANOVA, principal component analysis, linear regression, principal component analysis) were run in GenStat (VSN International, UK). The analysis of variance, of the 120 samples, was performed, assuming a completely randomised experiment design, with a blocking structure (i.e. plant + batch) and with a treatment (i.e. treatment*stage*part) design.

Extension work:

Molecular protocols

Nucleic acid extraction (DNA isolation)

Genomic DNA was isolated from buds of 63 different genotypes (see Table A3 in appendix for the full list) of raspberry (*Rubus idaeus*). Minimum 70 to maximum 90 mg of plant material were weighed and then fine ground with mortar and pestle using liquid nitrogen. The extraction/purification was performed using the DNeasy Plant Mini Kit (Qiagen, Germany), following the manufacturer instructions.

Analysis of nucleic acid

Gel electrophoresis

Agarose gels were prepared by mixing 0.75 g of agarose with 50 mL of 1X Tris-borate/EDTA (TBE) buffer. The mixture was heated in a microwave on medium power for 1 minute to allow the agarose gel to dissolve. The mixture was then cooled to ca. 50-60 °C before adding 1 drop of ethidium bromide. The gel was cast in a tank with the required size comb and allowed to set under a fume hood for 1 hour. Once the gel set, the comb was removed and sufficient TBE buffer was added to ensure the gel was fully submerged. Samples were then loaded onto the gel, in the wells formed by the comb, and separated for 50 minutes by electrophoresis at 40 Volts. Imaging of the gel was under UV light, using the UVITech transilluminator (UVITech, Cambridge, UK).

Enzymatic manipulation of nucleic acids

Design of primers

Gene specific primers for PCR were designed using the online software Primer3web version 4.1.0. (primer3.ut.ee). Primer sequences for Sanger sequencing were purchased from Eurofins Genomics (Ebersberg, Germany) while those primers for Fragment analysis (genotyping) were purchased from Sigma-Aldrich (USA) and of these, the forward primers were HEX (hexachloro-fluorescein) labelled.

Polymerase Chain Reaction (PCR)

PCR reactions were carried out in 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) using a final volume of 25 µL containing 0.2 µL (5 U/µL) of Taq DNA Polymerase and 2.5 µL of PCR buffer (F. Hoffmann-La Roche, Switzerland), gene specific forward and reverse primers (0.1 µL each primer) at a final concentration of 0.2 µM, 2.5 µL of dNTPs (Invitrogen™ Corporation, USA) at a final concentration of 0.2 mM and about 50 ng of DNA template (10 µL of DNA stock solution 1:10 dilution) (table 1). The final volume of 25 µL was reached by adding 9.6 µL of sterile distilled water (SDW). Thermal cycling conditions were as follows: 5

minutes denaturation at 95 °C followed by 35 cycles of 94 °C for 1 minute, 57 °C (melting temperature for the primer pair) for 1 minute at 72 °C for the extension of the expected fragments.

Table 1. Polymerase Chain Reaction (PCR) solution mix for 100 reactions. Volume of the different reagents for the final of 1.5 mL suitable for 100 PCR reactions.

PCR reagents	volume (µL)
dNTPs mix[a]	250
reaction buffer	250
Taq polymerase	10
forward primer	20
reverse primer	20
sterile distilled water (sdw)	960

[a]mix, in equal amount, of the four deoxynucleotides (i.e. dATP, dGTP, dCTP and dTTP)

Enzymatic clean-up protocol for Sanger sequencing

5 µL of PCR reaction for each DNA sample were transferred to a new 0.2 mL non-skirted 96-well PCR plates. To each well, containing the amplified DNA, were added 1 µL of Shrimp Alkaline Phosphatase (rSAP) and 1 µL of Exonuclease I both from New England BioLabs Inc. (USA). The PCR plate was then moved to the thermocycler, for the second and last step of the clean-up protocol, with the following settings, 37 °C for 15 minutes and 80 °C for another 15 minutes; the latter temperature for the complete and irreversible inactivation of the enzymes.

Protocol for preparation of samples for fragment analysis (genotyping)

For 100 reactions, 10 µL of Gene Scan™ 500 ROX™ and 990 µL of highly deionized (Hi-Di) formamide both Applied Biosystems S.A. (USA) were mixed together. Two µL of SDW diluted PCR reaction were mixed with 8 µL of ROX mix reaction buffer and transferred to a new 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) for the fragment analysis (genotyping).

Sanger sequencing of DNA

The fluorescent Sanger sequencing was carried out by the Genome Technology lab at The James Hutton Institute (Dundee, Scotland, UK). DNA was extracted (see section 2.2.2) from all the sixty-three different raspberry cultivars and selections (see Table A.1 appendix). A 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) was prepared containing 0.2 mL

of DNA stock solution (1:10 dilution) for each of the sixty-three samples. A Polymerase Chain Reaction (PCR) to amplify the region to be sequenced was performed as followed, 10 µL of DNA stock solution (1:10) were mixed with 15 µL of PCR reaction mix. The mix prepared for 100 reactions contained 250 µL of dNTPs, Deoxynucleoside triphosphates (Invitrogen™ Corporation, USA), 250 µL of buffer mix and 20 µL of Taq DNA polymerase both (10 µM) and both from (F. Hoffmann-La Roche, Switzerland), 10 µL forward primer and 10 µL reverse primer Eurofins Genomics (Ebersberg, Germany) and 960 µL sterile distilled water. The plate with the PCR reactions was then loaded in a thermocycler, Alfa Thermo Cycler (PCR max, UK) with the following reaction settings (see Table 2).

After amplification, the DNA samples were cleaned-up to remove any residual dNTPs that could affect the sequencing reaction; 5 µL of PCR reaction for each DNA sample were transferred to a new 0.2 mL non-skirted 96-well PCR plates and the enzymatic clean-up protocol was performed by adding, 1 µL of Shrimp Alkaline Phosphatase (rSAP) and 1 µL of Exonuclease I both from New England BioLabs Inc (USA) to each sample.

Table 2. PCR settings for the amplification of the genomic regions to be sequenced for markers detection.

Temperature °C	Time (min)	no. of cycles
95	5	1
94		
57	1	35
72		
72	8	1

The PCR plate was then moved in the thermocycler, for the second and last step of the clean-up protocol, with the following settings, 37 °C for 15 minutes and 80 °C for another 15 minutes; the latter temperature to completely and irreversibly inactivation of the enzymes. Once the clean-up reaction was complete, the samples were sent to the Genome Technology lab at The James Hutton Institute for the Sanger sequencing; the ideal amount of DNA for the sequencing for amplified fragments of size between 200 and 500 bp would be 3-5 ng; an estimation of the amount of DNA was done by running few random samples on a 1.5% agarose gel and then comparing the DNA band intensity with those of known amount of DNA used as reference.

For the sequencing reaction, the 5 µL of samples were further processed by adding 1 µL of (10 µM) of forward primer, 1 µL of Big Dye Terminator (version 3.1) reaction mix Applied Biosystems S.A. (USA), 1.5 µL of 5X dilution buffer Applied Biosystems S.A. (USA) and distilled water to bring final reaction volume to 10 µL. The plate containing the samples was

then transferred to TETRAD thermal cycler, Applied Biosystems S.A. (USA), using the following programme, hold at 96 °C for 1 minute and then a cycle to be repeated 25 times at 96 °C for 10 seconds, 50 °C for 5 seconds and 60 °C for 4 minutes. After amplification, the samples were cleaned up by adding for each 10 µL of reaction 2.5 µL of EDTA (125 mM; ph 8.0) and after a brief spinning 30µl of 95% ethanol. The sample were vortexed and spun briefly, incubated at room temperature for 15 minutes and then spun for 30 min at 3000 revolutions per minute (rpm) by keeping the sample at 4 °C. After that, the samples were spun upside down at 100 g for 10 sec, a further cleaning up step was performed by adding 150 µL of 70% ethanol. The sample were briefly vortexed and spun for 10 minutes at 3000 rpm at 4 °C, again this step was followed by spinning the samples upside down at 100 g for 10 seconds; this whole cleaning step was then repeated and after that the samples were left drying at room temperature. Once dried the samples were re-suspended in 10 µL of highly deionized (Hi-Di) formamide, Applied biosystems S.A. (USA) and then analysed on the capillary sequencer ABI3730 DNA analyser Applied biosystem S.A. (USA) with a 36 cm capillary array and Applied Biosystem's POP7 polymer. The samples were run using Applied Biosystems 3730 Data collection software version 4.0 and the sequencing Data coming off the machine were analysed using Applied Biosystem's Sequence Analysis version 6.0; all software Applied Biosystem S.A. (USA).

Fragment analysis for genotyping

The DNA SSR genotyping was carried out by the Genome Technology Lab at The James Hutton Institute (Dundee, Scotland, UK). The DNA samples, used for this analysis, were those utilised for the sequencing so the same DNA extraction procedure was applied. The PCR amplification, of the genome regions containing the markers, was performed in the same way described above; the only difference was the use of fluorescently labelled forward primers. After amplification, 5 µL of the PCR reactions were run on an electrophoretic gel (1.5 % agarose) to estimate the amount of DNA and then, accordingly, to calculate the dilution rate; usually 280 µL of sterile distilled water (SDW) per PCR reaction were added. To prepare the samples for the analysis, a reaction mix containing 10 µL of Gene Scan™ 500 ROX™ and 990 µL of highly deionized (Hi-Di) formamide both Applied Biosystems S.A. (USA); the 1 mL volume (ROX + Hi-Di) reaction mix refers to 100 reactions. Two µL of SDW diluted PCR reaction were mixed with 8 µL of ROX mix reaction buffer and transferred to a new 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) for the fragment analysis (genotyping). The samples were processed with the Genetic Analyzer 3730 Applied Biosystems S.A. (USA), the data were collected by means of the Genetic Analyzer 3730 data collection software version 4.0 Applied Biosystems S.A. (USA). The final data analysis was performed with GeneMapper software version 5.0 Applied Biosystems™ (USA).

Selection of a population of raspberry genotypes for a Genome Wise Association study (GWAS)

For GWAS, a list of 63 different genotypes (see Table A.4.1 in appendix) was produced to allow us to examine variability in the allele status and the identification of any associations with the phenotype and the potential 'crumbly' markers that could be of economic importance for molecular breeding programs and for diagnostic purposes (i.e. plant health certification). The population of genotypes was carefully selected following consultation with Nikki Jennings (raspberry breeder at The James Hutton Limited), however this selection was not straight forward as some conflicting information was available on the status of the genotypes depending on location and season and this was a noted limitation in the GWAS performed here. Forty-five genotypes, a mix of cultivars and selections, both primocane and floricane fruiting forms, were selected as plants that had been reported as showing 'crumbly' symptoms. The remaining eighteen genotypes, circa 25% of the total, were again a mix, of cultivars and selections both floricane and primocane but these were selected because they were never reported to have displayed 'crumbly' symptoms anywhere in any season and this was the reason for a low number of non 'crumbly' genotypes being included. Where possible, the plant pedigree was recorded because one of the criteria for the analysis to identify potential robust 'crumbly' markers, was to test them on a population of unrelated genotypes. Such approach would permit the selection of markers with a broad applicability.

Statistical analysis

Statistical analyses were performed in GenStat 18th edition (VSN International, UK). The chi-square analysis and the permutation test were performed to identify statistically significative association between the selected 'crumbly' markers and the population of 63 genotypes loosely related to each other with about ¾ of them labelled as showing 'crumbly' symptoms while the remaining ¼ being never 'crumbly'.

Results

'Crumbly' microarray experiment

The microarray experiment on mostly and never 'crumbly' material across three developmental stages (i.e. closed bud, open flower and green berry) allowed the identification of eleven key genes with clear roles in leading ovule fertilization and fruit development. Starting from the list of genes, of the microarray experiment, that were significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly'), screening was done

selecting only genes with ontology annotation related to flower development, hormones, pollen and transport. Flower and pollen genes were chosen as during these stages it might be expected that any alterations in the growth processes here could lead to the formation of misshapen ‘crumbly’ like fruit. The other two groups of genes (i.e. hormones and transport) were considered for their role as plants growth regulators, and transport because they focus on the hormonal interplay between receptacle and fertilised ovaries, proposed in this work to regulate the fruit growth and development.

The first screening identified 107 differentially expressed probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* equivalent, according to ontology annotations, were distributed between the classes: (12) flower development (56) hormones, (21) pollen and 20 transport (see Figure 1); two genes were shared in two different groups.

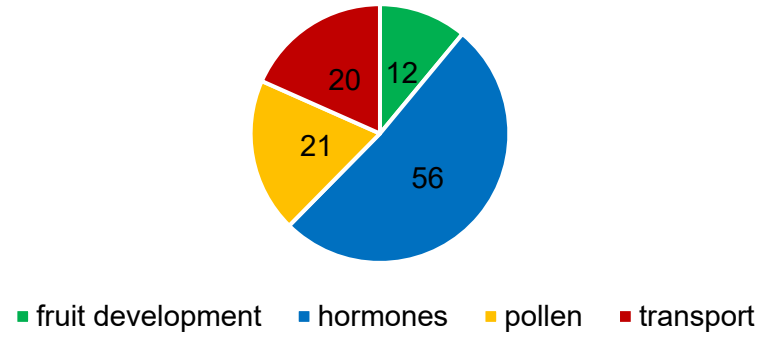

Figure 1. Pie chart of the 107 genes differentially expressed between the two different phenotypes (i.e. mostly and never ‘crumbly’). From the original list of 827 microarray probes matching Arabidopsis thaliana equivalent genes, 107 genes were selected from the four groups (i.e. flower development, hormones, pollen and transport) accordingly to their gene ontology annotations. The majority of the genes 56 (≈ 51%) had gene ontology terms related to hormones, 21 (≈ 19%) related to pollen, 20 (≈ 18%) to transport and 12 (11%) to flower development. Two genes were shared in two different groups.

Three groups (i.e. flower development, pollen and transport) of genes were individually subjected to tree cluster heatmap analysis while the group containing ontology terms related to hormones was further divided in three subgroups (i.e. response to hormones, hormones biosynthesis and ‘other’). The tree clustering heatmap analysis allowed further screening of the list of potentially interesting genes linked to ‘crumbly’ fruit to 69. In Figure 2, as example, was reported the tree cluster heatmap for the 107 ‘crumbly’ microarray probes differently expressed between the two phenotypes (i.e. mostly and never ‘crumbly’) and matching predicted *A. thaliana* genes with gene ontology annotations related to hormones, flower development, pollen and transport.

From these 69 *Arabidopsis thaliana* equivalent genes, identified via tree clustering heatmap analysis, a further screening was performed by searching, for each gene, for useful information (i.e. description, function and relevant scientific publication) in The Arabidopsis Information Resource (TAIR) https://www.arabidopsis.org/ database. The information was considered if it had relevance to any potential mechanism/process linked to the formation of flower and/or fruit, that could reasonably be associated to 'crumbly' like misshapen fruits (e.g. flower differentiation, gametogenesis, fertilization, etc.). The result of this further selection was a list of eleven genes whose products were involved, more or less directly, in processes responsible for pollen formation, pollen maturation, pollen tube elongation, pollen tube recognition and eruption as well as flower development and in particular stamens and ovaries.

The molecular mechanisms taking place at open flower stage and that might, potentially, cause the formation of misshapen 'crumbly' like fruit could be involved in:

- (AT3G01640.1 and AT1G63180.1) - altered synthesis of cell wall component, especially in pollen tube during its protrusion inside the stigma with potential premature arrest of pollen tube elongation and consequent missed fertilization
- (AT3G56950.1 and AT2G18840.1) - impaired transport of cell wall components, respectively from endoplasmic reticulum (ER) and Golgi apparatus, especially in pollen tube during its protrusion inside the stigma with potential premature arrest of pollen tube elongation and consequent missed fertilization
- (AT3G51550.1) - compromised interaction/recognition between pollen tube and embryo sac with consequent lack of fertilization of egg cell by sperm cell
- (AT3G12110.1) - delayed pollen germination due to alteration of the mechanisms involved in cytoskeleton rearrangement important for both pollen germination and pollen tube elongation.

The molecular processes taking place at an earlier stage (closed bud) and that might be responsible for the formation of 'crumbly' fruit might be:

- (AT4G24210.1) - impaired modulation of DELLA proteins (gibberellins (GAs) repressors at transcriptional level) with consequent GAs disequilibrium affecting normal flower formation and pollen maturation
- (AT3G12280.1) - impaired mega-gametogenesis causing the formation of an embryo sac without egg and central cells

- (AT1G76490.1) - formation of unviable pollen lacking essential components of cell membrane systems
- (AT1G14830.1) - altered formation of intine layer (the internal layer of the pollen grain wall) and plasma membranes with consequent formation of flawed pollen
- (AT1G42470.1) - impaired sterol biosynthesis with consequent production of poor viable pollen

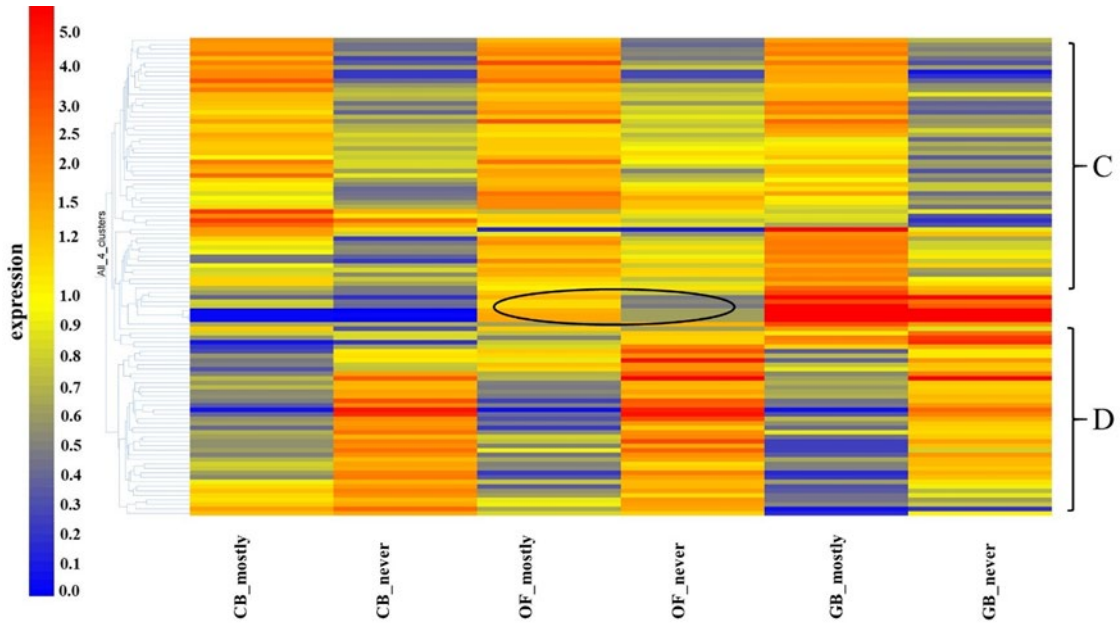

Figure 2. Tree cluster heatmap of the microarray probes differently expressed and matching genes with gene ontology annotation related to flower development, hormones, pollen and transport. The 107 microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and for the three different development stages (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**). Two main clusters of differently expressed probes were highlighted. They were named **C** and **D** and they contained respectively 38 and 37 probes (75 in total) which matched as many *Arabidopsis thaliana* ortholog genes. The probes of cluster **C** were upregulated in the mostly 'crumbly' compared to never one while the opposite for the probes within the cluster **D**. In the centre highlighted with a black circle shows for the sole open flower stage a small cluster of probes upregulated in the mostly 'crumbly samples.' High expression levels are indicated in red colour while low expression levels in blue colour as per scale bar presented on the left side of the heatmap.

The majority of these eleven key genes, from a 'crumbly' fruit perspective, are related to pollen and in particular to its formation and on its functioning. Impaired molecular mechanisms in which those genes are involved might cause formation of poor viable pollen and pollen tube with compromised ability to grow enough into the style and then fertilised the ovule. The final

and overall consequence of these impaired processes could be the formation of fruit with lower number of drupelets that might turn to be misshapen and with no commercial value.

'Crumbly' QTLs identified

The 'crumbly' field scores used here were those recorded over a period of seven fruiting seasons (Graham et al. 2015). The scorings were recorded to investigate the segregation of 'crumbly' fruit in a Latham x Glen Moy cross and the results were two 'crumbly' phenotypes. None of the progenies were scored as 'crumbly' in every season, the maximum was about 75% of the scoring times. However, some individuals never exhibited the 'crumbly' phenotype in any of the seven seasons scored. These same 'crumbly' scores were re-analysed in this study using the enhanced high density (GbS) linkage map to which 2348 new high confidence single nucleotide polymorphism (SNP) markers were added. This linkage map with higher coverage of markers allowed identification of a new QTL as well as a more accurate genetic position of loci detected with the previous linkage map (Hackett et al. 2018).

From the 'crumbly' fruit perspective, the results of the analysis of the 'crumbly' scorings with the new GbS linkage map were the confirmation of the two QTLs on LG1 and LG3 previously identified by Graham et al. (2015) with **RUB256e** at around 8.3 centi-Morgan (cM), previously identified as most representative marker, being replaced by **s3407_p12510_R23** at 7.1 cM. For the 'crumbly' QTL identified on LG3, the previous most significant marker **ERubLR_SQ05.3_D11AOC** at about 109.83 cM was replaced by **s182_p91185_R6** at 106.4 cM. An interesting finding was the discovery of a new QTL on LG3 not previously identified and having as most representative marker **Rub2a1** at 64.2 cM. All the three markers have four alleles (ab in Latham and cd in Glen Moy) and the significant markers always segregated in the Latham parent (the mean proportion of 'crumbly' fruit was higher in the genotypes carrying the Latham 'b' allele) for the two previously identified QTLs while for the newly identified 'crumbly' QTL on LG3, the allele combination 'bc' gave higher occurrence of 'crumbly' phenotype than the others, suggesting that both alleles were dominant and then the most significant markers segregate in both parents. In Figure 3 were reported the two linkage groups (i.e. LG1 and LG3) with the corresponding 'crumbly' QTLs.

The point estimation of the QTL location, the position along the map with the highest LOD for that specific QTL is the first step in the process of identification of a QTL. It is important to estimate the QTL precision by means of a confidence interval (supporting interval) and here, the 'LOD drop-off method' (Hackett, 2002) was applied.

In their work, Graham et al. (2015) found correlations between the 'crumbly' scores and those related to fruit ripening. The largest correlations were found with the time to reach fruit set and that to reach green fruit. These correlations were positive, meaning that the proportion of

‘crumbly’ fruit increased with the time plants took to reach these two development stages. The longer the fruits take to get to fruit set and green fruit, the more likely they tend to be ‘crumbly’ (Graham et al. 2015). Ripening is associated with many markers and many different QTLs were found (Graham et al. 2009). The ripening score consisted of repeated observations carried out on a finite number of occasions. The scores on different dates were then combined using principal coordinates (PCOs) to extract the principal sources of variation in ripening across the mapping population (Graham et al. 2009). The PCOs, being uncorrelated to each other, allowed an easy allocation of the QTLs but the interpretation of the QTLs in respect to the original ripening stage was very hard. It was very difficult to interpret the principal coordinates in terms of the effect of marker genotypes on the ripening process. For such reason a second approach was adopted by interpolating the time to reach each stage of ripening. This strategy allowed the natural interpretation of the QTL effect by simply detecting the presence or absence of a marker in the number of days to reach a certain ripening stage (Graham et al. 2009).

The same ripening scores were re-analysed with the new GbS linkage map and 12 of these QTLs were confirmed (Hackett et al. 2018). In terms of association between ‘crumbliness’ and fruit ripening, of particular interest seemed to be both the two ‘crumbly’ QTLs identified on LG3. The QTL, identified in this work, having **Rub2a 1** as most representative marker and located at 62.3 cM (centimorgan) and the previous QTL with **s182_p91185_R6** as most representative marker and located at 106.4 cM also impact ripening. Focusing on the field trials, the ‘crumbly’ QTL identified in this study on linkage group 3 overlapped two ripening QTLs, representing the time to reach fruit set (76 cM) and that to reach green fruit (49 cM). The ‘crumbly’ QTL previously identified, overlapped three QTLs (i.e. fruit set, green fruit and PCO2), respectively at 102, 105 and 94 cM (Hackett et al. 2018); in Figure 3 were reported the linkage groups with the corresponding QTLs for both ‘crumbly’ and fruit ripening traits. Five microarray probes differently expressed between the two phenotypes (i.e. mostly and never ‘crumbly’) with differences being significative at 99.9% confidence levels were selected since they mapped inside an overlapping region between the ‘crumbly’ QTL previously identified by Graham et al. (2015) and three fruit ripening QTLs (i.e. fruit set, green fruit and PCO2), for further detail see Table 3.

The probe CUST_54460_PI426541283, upregulated at both closed bud and green berry stage, matched the *Arabidopsis thaliana* ortholog gene (AT1G62360.1) and mapped inside the old ‘crumbly’ QTL as well as with the fruit ripening QTLs, fruit set and green fruit. From the ‘crumbly’ fruit perspective, two important gene ontology annotations were found for this gene, GO:0048440 (carpel development) and GO:0009691 (cytokinin biosynthetic process). The second interesting ‘crumbly’ microarray probe, CUST_35866_PI426541283, matched a

gene mapped inside the scaffold65 that was located within the old ‘crumbly’ QTL and again co-located with two fruit ripening QTLs, fruit set and green fruit. The probe was downregulated in the mostly ‘crumbly’ phenotype in all the three stages tested (i.e. closed bud, open flower and green berry). The probe matched the ortholog *Arabidopsis thaliana* gene (AT4G28210.1) that had two interesting gene ontology annotations, GO:0009793 (embryo development ending in seed dormancy) and GO:0009409 (response to cold). The probe CUST_37835_PI426541283 matched scaffold664 and the *Arabidopsis thaliana* ortholog gene AT1G30330.2. The gene was associated with three interesting gene ontology terms, GO:0009734 (auxin activated signalling pathway), GO:0009733 (response to auxin); GO:0009808 (flower development). The probe was upregulated in the mostly ‘crumbly’ samples. The last two probes, CUST_24407_PI426541283 and CUST_12159_PI426541283 were both mapped inside the previously identified ‘crumbly’ QTL (Graham et al. 2015) as well co-locating with two fruit ripening QTLs, fruit set and green fruit. While the first probe matched a gene inside a scaffold with two ripening QTLs (i.e. fruit set and green fruit), the second probe co-located with only one QTL for fruit set. The overlapping genome regions between these QTLs were respectively, scaffold734 for the first probe and scaffold858 for the second one. The probes matched the *Arabidopsis thaliana* ortholog genes, respectively AT5G12210.1 and AT2G19130.1. The two genes both presented gene ontology annotations related to pollen and more in details, for the first gene GO:0009555 (pollen development) and GO:0048544 (recognition of pollen) for the second gene.

Figure 3. Genotype by Sequencing (GbS) linkage map of linkage group 1 (left) and linkage group 3 (right) of Glen Moy x Latham. Maps are ordered by multi-dimensional scaling MDS (Preedy and Hackett 2016), black bars and whiskers (to the left of the chromosome) show one- and two-LOD support intervals for QTL locations.

Table 3. List of 'crumbly' microarray probes mapped in the overlapping regions of the 'crumbly' and fruit ripening QTLs on linkage group 3. For each probe was reported the matched *A. thaliana* ortholog gene and the scaffold were the *Rubus idaeus* equivalent gene is located. The last column reports the gene ontology terms related to the orthologs *Arabidopsis thaliana* genes matching the probes being analysed in this work.

Microarray probe	*A. thaliana* gene ID	[a]cM	[b]LG	scaffold	trait				[d]G.O. terms
					'crumbly'	fruit ripening QTLs (field trials			
					'crumbly'	[c]PCO2	fruit set	green fruit	
CUST_54460_PI426541283	AT1G62360.1	99.4	3	4	X		X	X	[1]GO:0048440 [2]GO:0009691
CUST_35866_PI426541283	AT4G28210.1	103.3	3	65	X		X	X	[3]GO:0009793 [4]GO:0009409
CUST_37835_PI426541283	AT1G30330.2	84.3	3	664	X	X			[5]GO:0009734; [6]GO:0009733; [7]GO:0009808
CUST_24407_PI426541283	AT5G12210.1	102	3	734	X		X	X	GO:0009555 (pollen development)
CUST_12159_PI426541283	AT2G19130.1	105	3	858	X		X		GO:0048544 (recognition of pollen)

[1]carpel development – [2]cytokinin biosynthesis – [3]embryo development ending in seed dormancy – [4]response to cold – [5]auxin activated signalling pathway – [6]response to auxin – [7]flower development – [a]centimorgan – [b]linkage group – [c]principal coordinate 2 – [d]gene ontology

Phytohormones semi-quantification

Targeted LC-MS analysis proved to be successful only for five out of the eighteen target plant phytohormones for which the analytical method was designed and developed. Calibration models were produced to help derive concentration values from the instrument response (i.e. ratio peak areas analyte/corresponding ISTD). When those predicted values were higher than the limit of detection, an estimation of the amount for the specific analyte was performed. The calibration model equations were all quadratic since they fitted the data more precisely than the linear functions. This was as a result of the large difference in the target analyte concentration between the two different stages (i.e. green and red fruit). The calibration curves were run in three different ways, in pure solvents (blank) and in matrix, respectively of drupelet and receptacle.

Unfortunately, quantification could only be partially achieved as only in some cases was the amount of the analyte in the samples above the limit of quantification (LoQ). The only exception was GA_1 whose amount, at least in the receptacle samples, was always above the LoQ. In Figure 4 was reported the bar chart of the relative amount (i.e. ng/50 mg) of sample for the only two phytohormones (i.e. ABA and GA_1) showing differences statistically significant between the two phenotypes (i.e. artificially induced 'crumbly' and control); with ABA being significantly higher, in receptacle of artificially induced 'crumbly' at green stage, while GA_1 significant lower in drupelet at red stage always of the artificially induced 'crumbly' samples. In Figures 5 was reported the bar chart of the relative amount (i.e. ng/50 mg) for all the five target phytohormones of the control samples at green stage for both drupelet and receptacle. In Figure 6 the relative amount (i.e. ng/50 mg) of both drupelet and receptacle at red stage for the control samples. At the green fruit stage, salicylic acid was by far the most abundant target phytohormones in all the samples (Figure 5). Salicylic acid together with ABA, GA_1 and GA_4, were measured at a higher concentration in the receptacle in all the samples (see Figures 5). IAA was the only target compound whose concentration were always higher in the drupelet than the in receptacle as well as at the green stage rather than red fruit stage; all the other target compounds were always higher in receptacle. At the red fruit stage, in the red drupelet ABA was the most abundant phytohormone (Figures 6) while SA was the compound with the highest concentration in the receptacle (Figure 6). The two gibberellins (i.e. GA_1 and GA_4) were the only compounds with no statistically significant differences between the two stages (Table 4). Again there was no statistically significant difference between the two phenotypes (i.e. artificially induced 'crumbly' and control) except for ABA in the receptacle at the green stage and for GA_1 in the drupelet at the red stage (see Figure 4) with ABA concentration being statistically significantly higher in the artificially induced 'crumbly' while that of GA1 being lower.

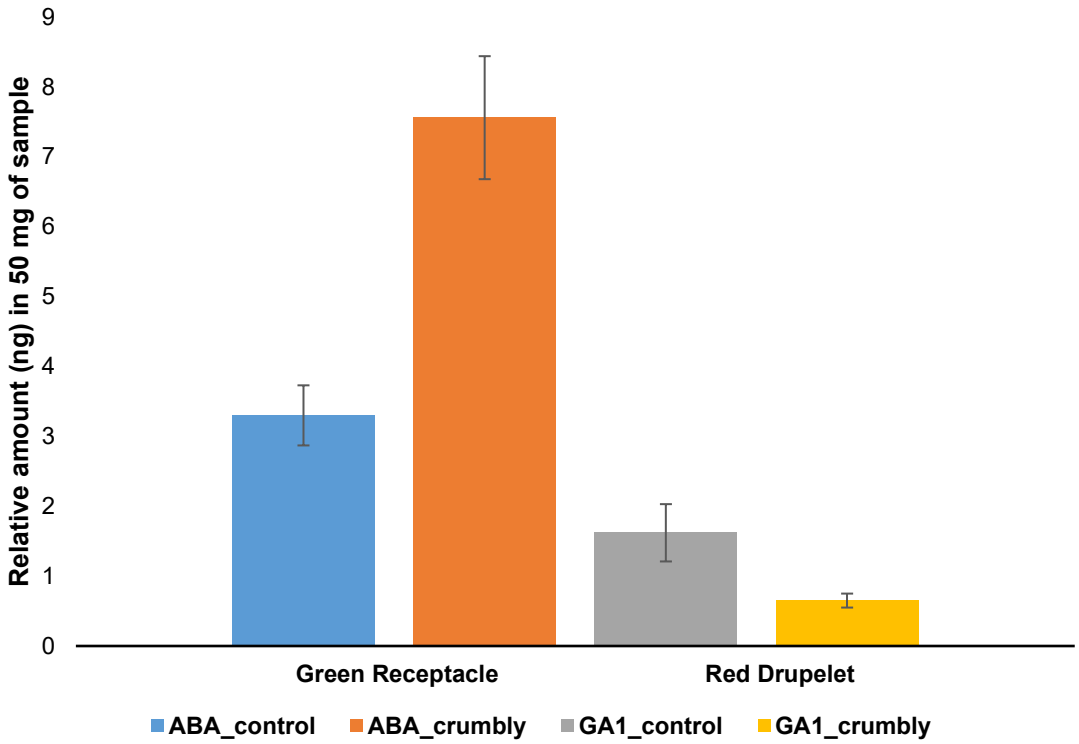

Figure 4. Bar chart of averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of ABA and GA_1 in control and artificially induced 'crumbly' green receptacle (ABA) and control and artificially induced 'crumbly' red drupelet (GA_1); differences being significative at about 95% confidence levels.

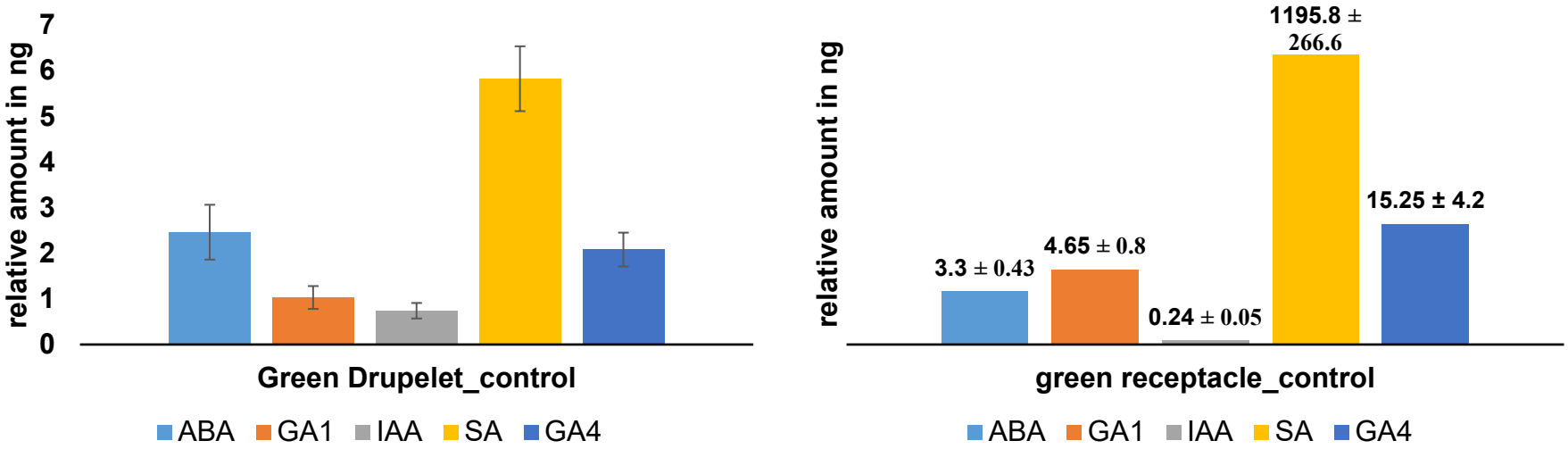

Figure 5. Bar chart of the relative amount (ng) of all the five detected/quantified phytohormones in both drupelet (left chart) and receptacle (right chart) at green stage in control samples. The amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of the five phytohormones in control green stage of drupelet and receptacle. Differences between parts being significant at about 95% of confidence levels. On the right plot, the huge differences, in the amount of salicylic acid (SA) compared to the other phytohormones forced to report data (i.e. amount and standard error) on top of the bars and then remove the y-axis scale.

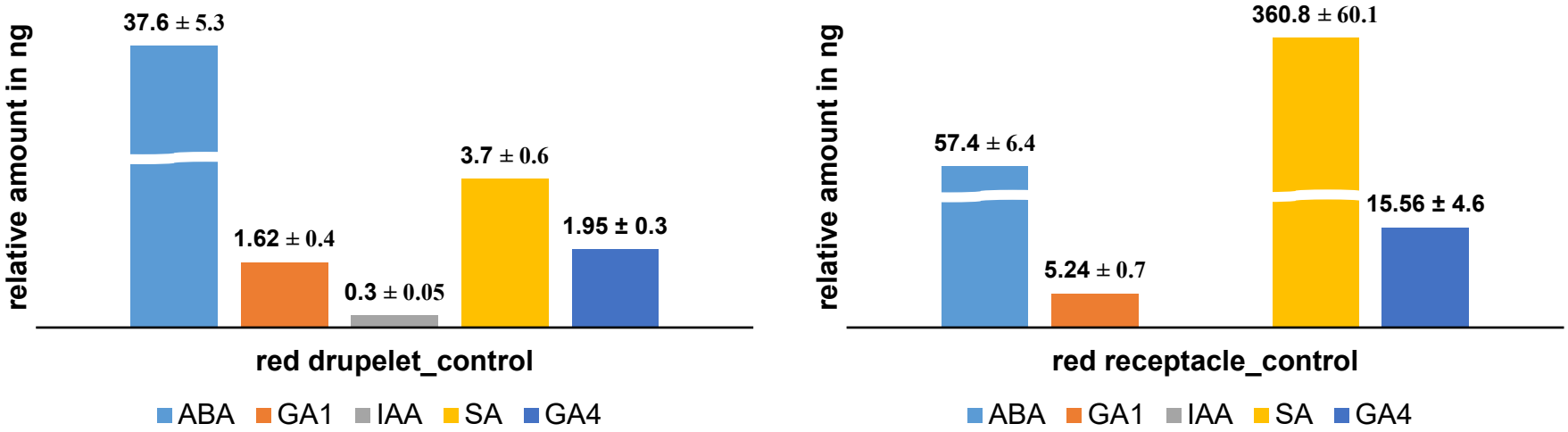

Figure 6. Bar chart of the relative amount (ng) of all the five detected/quantified phytohormones in both drupelet (left chart) and receptacle (right chart) at red stage in control samples. The amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of the five phytohormones in control red stage of drupelet and receptacle. Differences between parts being significative at about 95% of confidence levels. The huge differences, in the amount of abscisic acid (ABA) on the left and of ABA and salicylic acid (SA) on the right, when compared to the other phytohormones forced to report data (i.e. amount and standard error) on top of the bars and then remove the y-axis scale.

Table 4. ANOVA table of means for the stage per part interaction of the five detected/quantified target phytohormones. For each target phytohormones (i.e. ABA, GA_1, IAA, GA_4 and SA) was reported, together with the corresponding standard error (S.E.), the predicted means from the ANOVA analysis for the stage (i.e. green and red fruit) per part (i.e. drupelet and receptacle) interaction for all the 120 samples. The analysis showed that the differences in the predicted means were significative at 95% confidence levels for three out of the five compounds (i.e. ABA, IAA and SA).

Compounds	stage	part		[a]d.f.	[b]rep.	[c]S.E.
		drupelet	receptacle			(stage*part)
ABA	green	0.877*	1.534*	101	30	0.0367
	red	2.100*	2.397*			
GA_1	green	-0.042	0.577	100	30	0.0608
	red	-0.055	0.630			
IAA	green	0.572*	0.01*	70	30	0.0706
	red	0.366*	-0.397*			
SA	green	-0.215*	1.618*	105	30	0.0551
	red	-0.448*	1.189*			
GA_4	green	-1.218	-0.575	81	30	0.0834
	red	-1.237	-0.538			

*differences being statistically significant at 95% of confidence levels
[a]degree of freedom
[b]number of replicates
[c]standard error of differences of means from ANOVA (stage*part)

In conclusion the metabolomic analysis allowed the partial quantification of the target phytohormones with GA_1 being the only exception where full quantification was achieved and only for the receptacle samples. The most abundant compound by far was SA and its amounts in receptacle were only indicative. It could not be estimated properly since the values were much greater than the linear range of the calibration curves; the dilution and subsequent remeasurement of diluted samples could not be performed due to the extremely unstable nature of these target compounds, rapidly degradable in aqueous solutions. The second most abundant target phytohormone was ABA whose amounts were always greater in the receptacle. The only compound found in higher concentrations in drupelet compared to receptacle was IAA (see Figures 5-6).

Extension:

Steps towards the selection of 'crumbly' genetic markers for breeding assisted and diagnostic purposes

In total, eight potential 'crumbly' fruit loci were selected for investigation for marker development. Five of these loci were mapped as Single Nucleotide polymorphisms (SNPs), while the other three were Simple Sequence Repeats (SSRs). The eight gene markers selected were of two different kinds; the so called gene 'tags', located in non-coding regions of the genome, and the target genes, on the other hand, located in coding ones (Collard et al. 2005). **MOY_34151, MOY_35728** and **MOY_36258** were the only genes targets all the rest were genes 'tag'. These three genes were located on LG1 (**MOY_34151**) while the other two were both on LG3 one in the QTL identified here (**MOY_35728**) and the other in the original **'**crumbly' QTL (**MOY_36258**).

Two SNPs markers, **s3407_p12510_R23** and **s182_p91185_R6,** the two new most significant 'crumbly' markers for, respectively, the QTL on LG1 and the QTL on LG3, were identified during this work by re-analysing the 'crumbly' score (Graham et al. 2015) with the new GbS linkage map (Hackett et al. 2018).

The last three genes were all SSRs, **Rub2a1** was the most significant marker for the 'crumbly' QTL identified during this work while **ERubLR_SQ05.3_D11AOC** and **Rub256e** were the original 'crumbly' markers, respectively for the QTL on LG3 and the one on LG1; identified by Graham et al. (2015) before the development of the new GbS linkage map (Hackett et al. 2018). These two SSR markers were also considered for the Genome Wide Association Study (GWAS) because they were inside their corresponding 'crumbly' QTL and they were tightly linked to the new most significative markers. Therefore, the principle followed was, the more markers selected and tested the more chances to find at least one 'crumbly' marker being representative of a wide population of individuals showing 'crumbly' symptoms.

The design of the primers for the marker (**s3407_p12510_R23)** failed due to missing sequences in the **scaffold3407** where it is mapped. A new SNP marker (**s353_p21288_R19**), tightly linked with (**s3407_p12510_R23)** was then used.

The primers for all the eight markers were designed as described in material and methods section. They were used for a Genome Wide Association Study (GWAS) whose aim was to identify a 'crumbly' marker either always or never linked to the 'crumbly' population analysed. Ideally the marker/s should be shared by all the genotypes labelled as 'crumbly' (45 genotypes in this case study) while being absent in all those that never displayed 'crumbly' symptoms (18 genotypes in this case study) or vice versa (see Table A.4.1 for full list of genotypes).

The identification of the five SNPs markers, the two gene tags and the three gene targets, was performed by sequencing the amplified regions containing the markers across the genotypes. The procedure accomplished was described in material and methods of this report by using the primers designed as described always in the material and methods section. The results of the first sequencing analysis, performed only with the forward primers, did not identify any SNP marker that was always (or never) found in all the 45 genotypes, having 'crumbly' phenotype, out of the 63 genotypes (see full list of genotypes in appendix Table A.4.1) tested. Fourteen, out of the sixty-three, genotypes selected for the GWAS did not produce any amplification fragments; suggesting that their genome sequence differed in the primer design regions (see table 5 for further details) therefore these were not informative in these genotypes.

Table 5. Selected genotypes for the GWAS that did not give any amplification fragments for all the five SNPs markers, polymorphic between Glen Moy and Latham. List of fourteen genotypes, out of the 63 selected for the GWAS and not containing the five markers strongly associated with the three 'crumbly' QTLs and then chosen for this study. For each genotype, when available, was reported the pedigree too.

Pedigree				
♀ mother	**x**	**♂ father**	**genotype**	**phenotype**
complex hybrid	**x**	complex hybrid	Autumn Bliss	'crumbly'
Joan Squire	**x**	complex	Brice	'crumbly'
[1]Autumn bliss			Erika	'crumbly'
SCRI 6531/84	**x**	SCRI 6549/1	Glen Prosen	'crumbly'
			Kweli	'crumbly'
Nootka	**x**	Glen Prosen	Tulameen	'crumbly'
			Obbard	'crumbly'
Willamette	**x**	Cuthbert	Meeker	'crumbly'
97134B1	**x**	8510A57	0867E-4	'crumbly'
7326E1	**x**	7412H16	Glen Rosa	'crumbly'
			Imara	'crumbly'
Glen Rosa	**x**	SCRI 8605C-2	Glen Doll	no 'crumbly'
Preussen	**x**	Lloyd George	Malling Minerva	no 'crumbly'
			Malling Leo	no 'crumbly'

[1]open pollinated

The amplified region sequenced, with the procedure described in section 4.2.4, for the five different markers varied slightly between the genotypes. The contig, the set of overlapping DNA sequence for the different genotypes varied in length indicated in base pair (bp): **MOY_34151** (235 bp), **MOY_36258** (186 bp), **MOY_35728** (285 bp), **s182_p91185_R6** (487 bp) and **s353_p21288_R19** (355 bp). The number of SNPs per sequence identified varied

per locus. The contig for **MOY_34151** had four different SNP positions 139 bp, 146 bp, 158 bp and 183 bp. The contig for **MOY_36258** had five different SNP positions but none segregating in the selected population. The contig for **MOY_35728** carried three different SNP positions (i.e. 75 bp, 174 bp and 242 bp). The contig for the locus **s182_p91185_R6** had eleven different SNP positions (i.e. 110 bp, 129 bp, 139 bp, 147 bp, 148 bp 170 bp, 206 bp, 236 bp, 376 bp, 441 bp and 444 bp). The contig of the last marker, **s353_p21288_R19**, carried only three different SNP positions (i.e. 180 bp, 196 bp and 303 bp).

Chi-square (χ^2) tests of independence were used to look for associations between the two 'crumbly' categories (i.e. 'crumbly' and non 'crumbly') and markers. Each marker SNP position was tested to see whether any of the various polymorphic forms detected in the GWAS population was significantly associated with either of the two traits (i.e. 'crumbly' and non 'crumbly'). Ideally a p = 0.05 confidence level would be considered significant but due to the low numbers in the GWAS population higher levels would be considered significant (eg. p=0.01). If this was found, then that SNP position could have been considered linked with the genotypes showing 'crumbly' phenotype and that marker could be then analysed by a generalised mixed mode (Yu et al. 2006) specifically designed to address quantitative traits, such as 'crumbly' fruit and complex levels of relatedness within the analysed population.

None of the segregating SNP position, for the five markers, was highly significantly associated with the 'crumbly' phenotype in the studied population (see Tables 4 and 5 for further details). The use only of the forward primers to sequence the amplified fragment containing the selected markers limited the size of the region in which screen for segregating alleles. The use of both forward and reverse primers and the chromosome walking, along the region were the markers are located, could increase the size of the fragment to be sequenced and thus the number of SNP positions to enhance the chances of identification of segregating SNPs, potentially associated with the 'crumbly' phenotype, in the selected population.

The presence of expected values lower than 5 in any cell of the contingency table would make the probability test calculation unreliable (Steve, 2011). To bypass this issue, a permutation test to calculate the significance probability for a chi-square test of the independence of rows and columns was performed. This procedure converts the usual chi-square test in a nonparametric alternative that, in situations where some values in the contingency table are lower than five, allows to strengthen the results of the analysis (GenStat manual, VSN International, UK). The permutation test simulates the random distribution of table values that may occur in tables that have the same overall distribution of numbers over the columns, and over the rows, as in the original table. Even by means of a more accurate and precise statistical analysis, none of the markers tested proved to be significantly associated with the 'crumbly' phenotype.

For the SSRs markers, while for **Rub2a1** and **ERubLR_SQ05.3_D11AOC** none of the segregating alleles showed significant association with the 'crumbly' phenotype with, respectively, p = 0.28 and p = 0.24 for the permutated chi-square test. The SSR marker Rub256e however showed significant association p = 0.02 for the permeated chi-square test (table 6 and 7).

Table 6. Significance probability for a chi-square test.

marker	allele	*χ^2	†d.f.	p-value
MOY_34151	base_146	1.99	3	0.574
	base_139	2.62	4	0.662
	base_183	0.88	2	0.643
	base_158	2.56	2	0.217
[a]MOY_35	base_75	2.91	4	0.573
	base_174	2.52	2	0.284
	base_242	0.68	1	0.41
s182_p9185_R6	base_110	1.96	3	0.58
	base_129	2.93	2	0.231
	base_139	6.39	4	0.172
	base_147	1.52	2	0.467
	base_148	2.18	3	0.536
	base_170	3.19	2	0.203
	base_206	4.11	3	0.250
	base_236	2.01	2	0.367
	base_376	1.14	2	0.565
	base_441	2.18	3	0.536
	base_444	2.71	2	0.257
[b]s353_p21	base_196	3.69	3	0.297
	base_180	0.94	2	0.625
	base_303	2.9	2	0.235

*Pearson chi-square value - †degree of freedom
[a]**MOY_35728 - [b]s353_p21288_R16**

For each marker was reported the number of alleles, the chi-square test statistic, the probability test (p-value) and the number of degrees of freedom for a chi-square test with two criteria (i.e. 'crumbly' and 'non crumbly') was calculated according the formula (m-1) x (n-1) where m is the number or rows and n the number of column of the contingency table.

Table 7. Significance probability for a for a permutated chi-square test.

marker	allele	*χ^2	†d.f.	p-value
MOY_34151	base_146	1.99	3	0.668
	base_139	2.62	4	0.736
	base_183	0.88	2	0.723
	base_158	2.56	2	0.343
[a]MOY_35	base_75	2.91	4	0.684
	base_174	2.52	2	0.379
	base_242	0.68	1	0.479
s182_p9185_R6	base_110	1.96	3	0.727
	base_129	2.93	2	0.256
	base_139	6.39	4	0.167
	base_147	1.52	2	0.612
	base_148	2.18	3	0.632
	base_170	3.19	2	0.240
	base_206	4.11	3	0.234
	base_236	2.01	2	0.367
	base_376	1.14	2	0.569
	base_441	2.18	3	0.609
	base_444	2.71	2	0.298
[b]s353_p21	base_196	3.69	3	0.346
	base_180	0.94	2	1
	base_303	2.9	2	0.242

*Pearson chi-square value - †degree of freedom
[a]**MOY_35728** - [b]**s353_p21288_R16**

For each marker was reported the number of alleles, the chi-square test statistic, the probability test (p-value) and the number of degrees of freedom for a chi-square test with two criteria (i.e. 'crumbly' and 'non crumbly') was calculated according the formula (m-1) x (n-1) where m is the number or rows and n the number of column of the contingency table.

Discussion

Hormone regulation hypothesis

Although the plant material tested was different between the phytohormone study and the genetic study (QTL and microarray analysis), conclusions can still be drawn. For the microarray experiment, samples were selected from individuals of a mapping population between Latham ('crumbly' donor parent) and Glen Moy. These samples were from three different development stages (i.e. closed bud, open flower and green berry) and of the two different phenotypes (i.e. mostly and never 'crumbly'). The phytohormones quantification was performed on artificially induced 'crumbly' samples from Glen Moy plants grown under optimal condition under protect environment.

The analysis of the microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology annotation related to 'response to ABA' showed that ten probes were upregulated in the mostly 'crumbly' phenotypes. Such ontology terms indicate any kind of process triggered by ABA stimulus that cause change in the state or in the activity of a cell or an organism (i.e. movement, secretion, enzyme production, gene expression, etc.). Such finding, in a certain sense, validated the results of the phytohormones analysis giving further credit to them and confirming the potential role played by ABA, at the green stage in the development of 'crumbly' fruits.

The 'crumbly' microarray analysis showed an interesting result in terms of the hypothesis behind this work, of a hormonal crosstalk between receptacle and fertilised ovaries that drives the synchronised growth of all the pollinated carpels. Many of the 'crumbly' microarray probes that were differentially expressed between the two phenotypes (i.e. mostly and never 'crumbly') matched *R. idaeus* genes whose *A. thaliana* ortholog have gene ontology terms related to 'hormonal signalling pathway', involving seven different class of hormones (i.e. auxins, gibberellins, SA, ABA, ethylene, brassinosteroids (Brs) and jasmonic acid). The hormonal signalling pathway refers to the activation/modulation of molecular signals by the binding of a hormone to a cell receptor resulting in modulation of a molecular process. It might be speculated that this signalling pathway is the means through which the molecular messages might be transmitted between receptacle and fertilised ovaries. The majority of the genes were downregulated in the mostly 'crumbly' phenotype suggesting that this downregulation could indicate a reduction in hormonal crosstalk that might contribute to the formation of 'crumbly' fruit. Currently the data cannot provide an explicit conclusion because in the 'crumbly' microarray experiment, receptacle and drupelet were not separated in the analysis. The microarray still provides clues about the value of the putative receptacle centred hormonal growth regulation system.

The microarray experiment also produced some evidence about the involvement of cytokinins (CKs) in the model describing raspberry fruit growth at early stage (green berry). CKs regulate cell division and in concert with gibberellins even cell enlargement, but in contrast to IAA they are modulated by SA (Honda et al. 2017; Tiwari et al. 2013).

Two probes with ontology terms related to CK biosynthesis/metabolism were upregulated at the green berry stage and of further note, these were mapped inside the two 'crumbly' QTLs identified on linkage group 3. This represent a validation of the results from the microarray experiment, giving further credit to the putative role played by these two genes in 'crumbly' fruit because they reside in an area of the genome that is significantly associated with 'crumbliness'. The up-regulation of these two genes would suggest the involvement of significant higher levels of CKs in the mostly 'crumbly' phenotype, a parallel can be drawn in the artificially induced 'crumbly' fruit samples and we can then speculate about the putative participation of CKs in the hypothetical hormonal model describing fruit growth regulation.

According to this hypothetical model, CKs, in the receptacle, might replace IAA and attenuate ABA levels, because CKs, in contrast to IAA, are not affected by salicylic acid. The putative contribution/involvement of CKs might help to explain the scenario found in the artificial 'crumbly' samples where a significant higher level of ABA was detected/quantified. In fact, these significant higher levels of ABA, in receptacle, could be the consequence of a hormonal disequilibrium caused by lower putative levels of CKs that cannot attenuate ABA levels. Such hormonal imbalance might be responsible for the development of 'crumbly' fruit.

It could be that in the artificially induced 'crumbly' fruits the damage which affects some of the pollination events is similar to the situation that occurs in unpollinated flowers (or in flowers damaged by means such as nectar load, bee foraging etc). In fact, without fertilization of the ovules, the triggering of crucial events such as synthesis and or activation of auxins, cytokinins and gibberellins would not take place (Obroucheva 2014). The levels of ABA would not be attenuated by auxins and the flower gradually senescence (Ozga and Reinecke 2003). In a certain way, it could be speculated that something similar might happen to the receptacle of the damaged flowers of the 'crumbly' induction experiments where it is reasonable to assume a link with the obstructed interplay between ovaries and receptacle. Soon after fertilization, in the damaged receptacle, the activation of the aforementioned events (i.e. synthesis and/or activation of auxins, cytokinins and gibberellins) would take place only in a minimal way. This would explain the significant higher level of ABA in the receptacle and the senescence for the majority of the carpels (see right picture on Figure 7) that, although all hand pollinated, did not progress on compromising the normal growth and development of the fertilised ovaries.

Similar considerations might be taken for the red stage, the targeted phytohormones analysis showed an increase of ABA in both receptacle and drupelet compared to the green stage and this would be compatible with its role as a leading regulatory factor of fruit ripening in non-climacteric plants such as raspberry. The analysis registered a decrease of both IAA and SA in red fruit while gibberellins were again the only compounds whose difference between the two different stages were not significant. According to the results of the phytohormones analysis, in normal fruit growth the interplay between ABA, IAA and SA seems to play a crucial part in fruit ripening with SA that modulating IAA and then indirectly increasing ABA that in turn drives the ripening of the fruit.

Figure 7. Artificially induced 'crumbly' fruit at red and green stage. Examples of artificially induced 'crumbly' fruit produced by pinning the tip and side of the receptacle with a needle soon after emasculation. Two days after the treatment when the stigmas of the flowers become receptive, all the carpels were hand pollinated to make sure all the carpels were pollinated. On the left a fruit at red stage and on the right one at green stage. These are typical of those found in nature.

From the 'crumbly' fruit perspective, the most important result of the target phytohormones analysis was the significantly lower levels of GA_1 in the drupelets of artificially induced 'crumbly' fruits compared to the control. As for the green stage, the hormonal model with five components (i.e. ABA, GA_1, IAA, GA_4 and SA) could not easily explain the data related to the significantly higher level of GA_1 in the drupelet samples of the artificially induced 'crumbly' samples. Once again, the contribution of cytokinins (CKs) was speculated as a putative model of hormonal regulation of fruit growth at late stage (i.e. red berry). The above-mentioned CKs

would cause a hormonal disequilibrium with the result of reducing the levels of GA_1 in red drupelets and of delaying the fruit development. This might explain the peculiar growth and ripening of the artificially induced 'crumbly' samples. In fact, they tended to grow slower, with the longer times required by the drupelets to reach, in some cases (see Figure 7), the very extremely large size observed in this type of samples 'crumbly'. The significantly lower levels of GA_1 coupled with putative high levels of CKS might prolong the cell elongation step that in turn cause the abnormal growth of the drupelets compared to those of normal berry.

The phytohormones analysis of the artificially induced 'crumbly' fruits allowed the discovery of some physiological aspects related to the hormonal regulation of fruit growth and development and even to provide evidence that the damage of the flower, and in particular of the receptacle, can cause the formation of 'crumbly' fruit.

Genetic control of developmental process

The analysis of the 'crumbly' microarray experiment highlighted many probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly' identified seven genes) that were implicated in aspects of the flower/fruit growth that could be linked with processes responsible for the formation of misshapen 'crumbly' like fruits. Pollen formation and anomalies in the activity of pollen and specifically of pollen tube, were selected as key processes leading to the formation of 'crumbly' fruit. From the genetic point of view, the analysis of the microarray experiments suggested the importance of processes upstream of the fertilization event as key steps in the growth and development of 'crumbly' fruit.

Impairments in the formation of mature pollen could substantially impact the formation of normal shape fruit because this defective pollen could not complete its function. The *A. thaliana* gene (SLY1) encoding SLEEPY1, of the components of the SCFSLY-SUMO complex is responsible for the ubiquitination of DELLA proteins and then indirectly of their degradation through the 26S proteosome complex (Kim at al. 2015). In the mostly 'crumbly' phenotype, the downregulation of the probes matching the *R. idaeus* gene, ortholog of SLY1, might have affected the response to gibberellins, a class of hormones important for processes such as pollen development and maturation (Daviere and Achard 2013). The impaired modulation of DELLA might affect the response to gibberellins with consequent reduction of fertility due to production of flawed pollen (Kim et al. 2015).

The downregulation in the mostly 'crumbly' phenotype of HMG1 in pollen might have caused the formation of abnormal pollen due to lack of important component of the membranes (i.e. phytosterols) whose synthesis in the tapetum cell of the pollen depends primarily on HMGR1. The *Rubus idaeus* equivalent to the *Arabidopsis thaliana* HMG1 encodes the β-hydroxy-β-

methylglutaryl-CoA reductase (HMGR1) a key enzyme, in the tapetum cells of the pollen, for the cytosolic mevalonate pathway (MVA) important for the biosynthesis of terpens, precursors of sterols and steroids (Suzuki at al. 2009).

The downregulation of a putative Niemann-Pick C1 protein involved in regulation of the sterol pathway, might cause defect in the gametogenesis (Feldman et al. 2015) with consequent formation of poor viable pollen.

These three examples of alteration in the formation of viable pollen were clear examples of anomalous processes taking place in the mostly 'crumbly' phenotype that suggest clear impacts on 'crumbly' fruit because for every flawed pollen reaching the stigma of the carpel, a raspberry fruit would form lacking drupelet(s).

The analysis of the 'crumbly' microarray also identified four genes controlling the synthesis of important components of the cell wall that are indispensable during processes such as pollen germination and pollen tube elongation.

Firstly, GLUCURONOKINASE G, a member GHMP-kinase superfamily whose specific substrate is d-glucuronic acid. The gene is expressed in all plant tissues with a preference for pollen where it supplies the cell wall polymers indispensable for accompanying the expansion of the pollen tube (Pieslinger et al. 2010). Secondly, AT1G63180.1, encodes (UGE3), a UDP-D-glucose 4-epimerase, responsible for the conversion of UDP-galactose to UDP-glucose (UDP-glc). The enzyme is involved in pollen development (source TAIR https://ui.arabidopsis.org/) and it appears to work alongside the Glucuronokinase G. UGE3 produces UDP-glc which is the precursor of UDP-glcA (UDP-glucuronate), the substrate of Glucuronokinase G. Both the enzymes, UGE3 and Glucuronokinase G, are expressed preferentially in pollen and contributes indirectly to supply the required components to the cell wall during the rapid expansion of the pollen tube (Pieslinger et al. 2009). The probes of the 'crumbly' microarray probes matching these two genes were upregulated in the mostly 'crumbly' phenotypes giving an indication that an inappropriate expression level of these genes might be deleterious for the synthesis of components of the cell wall especially in the fast growing pollen tube compromising its ability to expand and reach the ovule for the fertilization.

The pollen tube, the fast growing tissue of the plant is required to elongate and carry the gametocytes to the ovule, needs both an efficient supply of cell wall components to follow through the expansion of the tissue but also requires an efficient machinery to transfer all these compounds to the site of growth of the cells. The actin cytoskeleton constitutes the track through which the cell coordinates the organized movement of all the components that are required to accompany both membrane expansion and cell wall synthesis.

In the mostly ‘crumbly’ phenotype, the third gene AT3G12110.1 encodes ACT11 a reproductive actin, actively expressed in pollen, required for the correct rearrangement of the actin cytoskeleton (Chang and Huang 2015). In the mostly ‘crumbly’ plants, the gene encoding ACT11 was significantly downregulated and its low expression might be responsible for impairment in pollen germination and pollen tube elongation with consequent lack of fertilization that can cause the formation of misshapen fruits with lower number of drupelets such as the ‘crumbly’ ones.

The pollen tube needs to protrude inside the style of the carpel but once it reaches the ovule, inside the ovary, its elongation should terminate and the pollen tube should erupt to release the two sperm cells and accomplish the fertilization (Escobar-Restrepo et al. 2007). Plants have evolved a series of strategies for the successful reception of the pollen tube inside the embryo sac and in particular a recognition system between Receptor Like Kinase (RLK) expressed on the plasma-membrane of the synergid cells and putative ligand produced by the pollen tube. Without this recognition the pollen tube continues its growth inside the embryo sac, does not erupt releasing the sperm cells and thus fertilization does not take place (Escobar-Restrepo et al. 2007). In *A. thaliana*, FERONIA (FER) is the RLK responsible for the correct interaction between pollen tube and synergid cells (Haruta et al. 2018). In the mostly ‘crumbly’ plants the fourth gene, the *R. idaeus* equivalent gene of FERONIA was downregulated suggesting that a potential cause of misshapen fruits with lower number of drupelets might be found in missed fertilization due to incorrect recognition between pollen tube and embryo sac.

Microarray probes residing within QTLs

Some of the microarray probes, differently expressed between the two phenotypes (i.e. mostly and never ‘crumbly’), were mapped inside the three ‘crumbly’ QTLs. The two original ones, previously identified by Graham et al. (2015), and a third one identified during this work. From the ‘crumbly’ fruit perspective and for all the aforementioned properties, these probes and more precisely the genes to which they were matched were important for validation of the microarray data but also as being potential markers in future for breeding.

These genes included flower anatomical and pollen functional disruptions that represent potential factors responsible for the formation of misshapen fruit.

Under-expression of the *Arabidopsis thaliana* gene (AT5G12210.1) which encodes for the β-subunit (RGTB1) of the RGT (Rab geranylgeranyl transferase), negatively affected polar growth of pollen tubes compromising the fertilization of the ovules and causing the formation of anomalous flowers with carpels much longer that the stamens (Hala et al. 2010; Gutkowska

et al. 2015). The ‘crumbly’ microarray probe, matching the *R. idaeus* gene equivalent of the *A. thaliana* (AT5G12210.1), was down-regulated in the mostly ‘crumbly’ plants suggesting, for the reason discussed above, a potential important role in causing the formation of ‘crumbly’ like misshapen fruits. The microarray probe, matching the *R. idaeus*, ortholog of the *A. thaliana* (AT5G12210.1), maps inside the ‘crumbly’ QTL on linkage group three previously identified by Graham et al. (2015)

Aminoacyl-tRNA synthetases (AARSs) are essential enzymes catalysing the reaction responsible for the attachment of amino acids to their corresponding tRNAs. The methionyl-tRNA synthetase called OVA1 (ovule abortion 1) is an important AARSs acting in the mitochondria whose disruption cause ovule abortion as shown in *Arabidopsis thaliana* T-DNA insertion mutants (Berg et al. 2005). The microarray probe, matching the *R. idaeus* gene ortholog of the *A. thaliana* (OVA1), was upregulated in the mostly ‘crumbly’ plants, those displaying misshapen fruits; the probe maps inside the ‘crumbly’ QTL recently identified during this work on linkage group three (LG3). It might be speculated that behind processes such as ovule fertilization and seed formation there is the participation of many regulators acting both up and downstream to OVA1. The over expression of OVA1 could affect the action of these other regulators causing a functional disequilibrium whose consequence could be ovule abortion and/or arrest of seed formation.

Abnormal flowers with reduced number of carpels and consequently of ovules might be regarded as potential factors causing ‘crumbly’ fruit. The less carpels per flower the lower the drupelets per fruit. A differentially expressed microarray probe that mapped within the ‘crumbly’ QTL, previously identified by Graham et al. (2015) on linkage group three, matched a *R. idaeus* gene whose *A. thaliana* ortholog (AT1G62360.1) encodes the KNOX transcription factor SHOOT MERISTEMLESS (STM) which is responsible for promoting carpel development and their associated placental tissues (Scofield et al. 2007). Experiments with RNA interference on the STM gene (STM-RNAi) showed the formation of anatomical defects at the level of the floral meristem (Scofield et al. 2007). The model for the role of STM in carpel development proposed by Scofield et al. (2007) involves another key regulator of flower organ development, the AGAMUS like protein (AG) and its repressors, with BELLRINGER being one of them. Although STM was upregulated in the crumbly fruit AG and BELLRINGER were not differently expressed between the two phenotypes (i.e. mostly and never ‘crumbly’). This would suggest that an appropriate balance of all the three factors is necessary for normal flower development and that the upregulation of STM would be deleterious and potentially responsible for the formation of flowers with anatomical defects.

The analysis of the gene expression showed a clear genetic component behind the ‘crumbly’ fruit phenomenon. Genes involved in processes related to pollen formation, pollen tube

elongation, pollen tube and embryo sac recognition as well as flower development, were differently expressed, with differences being statistically significant. These genes were deleterious for the processes they were involving in and in fact, for the mostly 'crumbly' plants, the impairment of all these processes might translate easily in higher chances to develop misshapen fruits. Lower level of viable pollen rather than anatomically flawed flowers or impaired pollen tube protrusion into the style or even compromised pollen tube and embryo sac recognition are all clear examples of anomalous phenomena whose main effect is the reduction of the fertilization. In extreme synthesis, the less flowers fertilized, the lower drupelets developed in a berry and the higher chance for the plants to form 'crumbly' like misshapen fruits.

'Crumbly' and fruit ripening

Another important aspect, directly linked to the genetic aspects behind 'crumbly' fruit, was the association found between the three 'crumbly' QTLs and fruit ripening QTLs. This linkage was first found by Graham et al. (2015) and was later strengthened by re-analysing the fruit ripening phenotypic scores with the higher density markers, Genotype by Sequencing map (Hackett et al. 2018). Of particular interest was the association between the original 'crumbly' QTL on linkage 3 and three fruit ripening stages/traits (i.e. PCO2, green fruit and fruit set). PCO2 was the principal coordinate two, (principal coordinates (PCOs) was the approach used to combined scores on different dates to extract the principal sources of variation in ripening across the mapping population) (Graham et al. 2009) interpreted as a plant showing slow fruit development in May but with a rapid ripening during late Jun and/or early July (Graham et al. 2009). For the association with fruit set and green fruit, this was positive and could be expressed as the longer the fruit took to reach fruit set and green fruit stage the more likely to be 'crumbly' (Graham et al. 2015). These associations are very interesting and in future can be explored for genetic interpretations. A consideration might be raised for the potential predisposition of one of the two different class of raspberry varieties, floricane (biennial growing cycle) and primocane (annual growing cycle). The 'crumbly' versus fruit ripening QTLs association was performed on floricane plants (siblings of two floricane cultivars, Latham and Glen Moy) with biennial growing cycle and characterised by early fruiting compared to primocane that are even called autumn fruiting varieties. It cannot be excluded, *a priori*, that differences in terms of predisposition towards 'crumbly' fruit exist between these two classes of raspberry varieties. This issue opens up the need for further research on this topic although a little study was conducted during this work (data not shown) to tackle this aspect. An ordinal regression analysis was conducted on a list of 63 different genotypes (see Table A3 in appendix). The ordinal regression assumes that there is an underlying continuous

trait of crumbliness, splits into groups for scoring (i.e. ‘crumbly’ and non ‘crumbly’) and looks at the effect of other variables, in this case the different genotypes, on the distribution across groups. The result of the analysis showed that the difference between floricane and primocane in terms of predisposition towards ‘crumbliness’ were not statistically significant.

Extension:

The results of the work which has been done within the extension of this project showed that the first ‘crumbly’ markers tested for association with specific genotypes from a population of 63 individuals did not show highly significant associations with the ‘crumbly’ fruit phenotype though some of these are promising and should be examined in future on a larger population and also considering other SNPs close by as due to time constraints and Covid-19 the first steps towards marker identification/validation were focused on the sequencing of amplified regions produced with only the forward primers. This produced a limited number of SNPs for analysis and in future the use of both primers (i.e. forward and reverse) and extending sequencing across a wider region might help with the identification of further alleles that might be significantly associated with ‘crumbly’ fruit. Two different kind of loci, for the identification of potential molecular markers, were considered during this study. Single Nucleotide Polymorphisms (SNPs) and Simple Sequence Repeats (SSRs). Eight markers were selected in total. Five were SNPs and were of the different kinds, gene ‘tags’, located in non-coding regions of the genome, and target genes located in coding regions while the remaining three markers were SSRs.

The markers are all significantly associated with crumbliness in the Latham x Glen Moy population and in order to determine their wider applicability we required a validation population. The numbers in this population were smaller than ideal however due to being able to identify varieties that never show crumbly symptoms. A population of 63 different raspberry genotypes, with ¾ being selected as being prone to ‘crumbliness’ while the ¼ as never showing ‘crumbly’ symptoms was developed. DNA was extracted was from each of the 63 different genotypes and by means of the primers specifically designed to amplify the genome regions containing the selected markers.

Analysis of the allelic polymorphisms was conducted for each marker and potentially significant associations between alleles and ‘crumbly’ or not ‘crumbly’ were assessed; the goal was to identify at least one marker significantly linked with one of the two phenotypes through a Genome Wise Association Study (GWAS) using the selected 63 genotypes population as sample. Though not ideal this allows us to examine the allele status of the markers from the LxM population in a wider gene pool.

The selection of markers associated to complex quantitative traits such as ‘crumbly’ fruit is not an easy task and the size and structure of the population taken under exam play an important role in the whole process.

Three ‘crumbly’ SSR markers were also selected and their primers were designed. In this case Rub256e that showed significant association with the ‘crumbly’ phenotype.

Conclusions

This work focused on the study of Malformed Fruit Disorder (MFD) in raspberry by means of studying the genetics/gene expression and hormone profiles. This study into the bases of ‘crumbly’ fruit, produced many interesting findings that increased the insight about the physiology and the genetics behind the ‘crumbly’ fruit phenomenon.

The starting point was the hypothesis that a molecular mechanism, mediated by phytohormones through crosstalk and controlled by the receptacle, regulates and synchronises the growth of the many drupelets (i.e. on average minimum 60) forming the final fruit. This assumption raised a several questions to be addressed. The first about the phytohormones and their interplay, while the second question concerned the role of the receptacle as leading hub for the fruit growth process. Furthermore, the main question regarded the role of phytohormones in ‘crumbly’ fruit.

The analysis of the phytohormones proved that plant hormones are involved in the process of fruit development, a model of hormonal regulated fruit growth was proposed with the engagement of four measured (i.e. ABA, GAs, IAA and SA) and one speculated CKs. The model was coherent with previous findings on fruit growth (Vriezen et al. 2008; Ozga and Reinecke 2003; Obroucheva 2014; Serrani et al. 2008; Kumar et al. 2014), the only exception was SA, which previously has not been reported in soft fruit. SA is an important component of the signal transduction systems (Elwan and El-Hamahmy 2009) and is an extensive modulator of IAA (LeClere et al. 2008). From the ‘crumbly’ perspective the most important result was the significantly higher concentration of ABA found in the receptacle at the green fruit stage and the significantly lower concentration of GA1 in the drupelet at the red fruit stage in the artificially induced ‘crumbly’ samples. The involvement of hormones in fruit growth was demonstrated with the role of specific growth regulators (i.e. ABA and GA1) highlighted.

In the phase of the design of both experiment and analytical method, it was decided to include methylated phytohormones as the compounds to be targeted. Methylated plant hormones could be the ideal candidate for intra and intercellular transport of hormones, allowing plants to easily move hormones, short and long distance, in a specific site (i.e. tissue or organ)

where they can conduct their regulating action (Qin et al. 2005); the hypothesis behind this question was that the signal would start from the receptacle and then moves to the surrounding drupelets to coordinate and synchronize the growth of the fruit. According to this hypothesis and with the ease of which methylated compounds are transported through the cells, they might be expected to play an important role in regulating the fruit growth and high levels of them might be found in the receptacle where they could be synthesized. However, none of the methylated compounds were detected during the analysis of all the 120 samples. It is reasonable to think that these compounds act at extremely low concentration (fmol-pmol/g of fresh weight) and their detection would be arduous even treating them separately (Simura et al. 2018). Obviously the simultaneous measurement of many different compounds each with specific chemical properties requires a compromise, between all the different target analytes, in terms of both sample preparation and methodology of analysis that would impact those analyte naturally present at lower concentration and/or with features that make their analysis more challenging. After all, the adoption of analytical methods (i.e. immunological assay) for a single compound or a specific class of phytohormones (Du et al. 2012a), although conceptually valid becomes impractical in cases in which many different compounds are under analysis simultaneously, as was the case here.

On the basis of the above considerations, it was evident that no conclusions could be drawn from the data of the LC-MS phytohormones analysis on its own. Support was needed from the gene expression analysis (see chapter 3 for full details). Some clues about this hypothetical hormonal crosstalk were determined from the microarray probes, differentially expressed between the two phenotypes (i.e. mostly and never ‘crumbly’). Nine different classes of hormones (i.e. auxins, gibberellins, SA, ABA, ethylene, brassinosteroids (Brs) and jasmonic acid) were represented suggesting the participation of multiple different factors to this transfer of molecular messages. The majority of these probes were downregulated suggesting that the reduction of this hormonal signalling might be associated with ‘crumbliness’ supporting this theory.

The same considerations for the methylated phytohormones applies to cytokinins (CKs) and more precisely to trans-zeatin, the only target CK analysed in this work. It was not detected in the samples but its involvement in the hormonal fruit growth model was speculated because for the mostly ‘crumbly’ plants, both the significantly higher level of ABA in receptacle samples at green stage and significant low level of GA1 in drupelet at red fruit stage. The decision to select the trans-zeatin as an additional hormone to complete the model was supported by both literature data, confirming the important role play by this growth regulator in fruit development (Kumar et al. 2014) and the results of the ‘crumbly’ microarray experiment. In the latter, two probes significantly differentially expressed between the two phenotypes (i.e.

mostly and never ‘crumbly’) matched two *R. idaeus* genes whose *A. thaliana* ortholog have gene ontology terms related to CKs biosynthesis/metabolism. The two probes were significantly upregulated in the mostly crumbly plants at the green berry stage suggesting that higher level of CKs might be needed during the development of ‘crumbly’ fruit at the green stage. Moreover, to give further credit to these two genes, with respect to ‘crumbly’ fruit, was the position along the linkage map. They mapped inside the two ‘crumbly’ QTLs on linkage group 3, strongly associated with ‘crumbly’ fruit.

The analysis of gene expression clearly demonstrated that ‘crumbly’ fruit had a genetic base. Pollen was identified as a key factor and the genes involved in its formation and functioning seemed to play a crucial role in the formation of ‘crumbly’ like misshapen fruits. The mostly ‘crumbly’ fruit were shown to have impairment in the formation of viable pollen or anomalous expression of genes controlling the biosynthesis of important cell wall components due to the lack of structural component and/or impairment of the machinery (i.e. actin cytoskeleton) indispensable to deliver these compounds (i.e. sterols above all) and sustain a very demanding process such as the protrusion of the pollen tube. Defects in the molecular component indispensable for the recognition between pollen tube and embryo sac were other examples of impaired processes, caused by abnormal expression, always in the mostly ‘crumbly’ plants of the gene involved in these events. Lastly it was clear that the final effect of all this impaired process resulted in no fertilization of the ovules which translates into berries with a reduced number of drupelets, that for this reason might give raise to ‘crumbly’ like misshapen fruits.

The development of the Glen Moy x Latham raspberry linkage map using the Genotype by Sequencing (GbS) technology (Hackett et al. 2018), due to the higher markers density, allowed the identification of a new ‘crumbly’ QTL on linkage group 3 (LG3) as well as a more accurate and precise location for those two previously identified on LG1 and LG3. All three QTL proved to be robust across different seasons therefore there is potential to identify molecular markers located within these QTLs. Markers strongly associated with ‘crumbly’ fruit across a wider gene pool would be valuable for molecular markers assisted breeding and for diagnostic purposes. The ‘crumbly’ SSR marker Rub256e were identified in significant association with the ‘crumbly’ phenotype and primers were developed. Rubus 256e can be considered as a marker for crumbly fruit. Other regions have been identified but need further validation. The identification of ‘crumbly’ molecular markers could pave the way for the development of faster diagnostic test to replace the current fruiting test. The same molecular knowledge could be useful to the Raspberry Breeding programme based at James Hutton Limited allowing the selection of new varieties crumbly free or at least more resistant to this condition.

Knowledge and Technology Transfer

02-2020	*Presentation*	Crumbly Fruit in Red Raspberry: New Findings. Scottish Society for Crop Research (SSCR) - soft fruit winter meeting.
11-2019	***Presentation***	Understanding the causes of 'crumbly' fruit in red raspberry – Raspberry Breeding Consortium (RBC).
07-2019	**Press conference**	Understanding the causes of 'crumbly fruit' in red raspberry – The James Hutton Institute, Fruit of the Future, press conference
01-2019	***Presentation***	Understanding the causes of 'crumbly' fruit in red raspberry. The James Hutton Institute – Cell and Molecular Science (CMS) – winter meeting
11-2018	***Presentation***	*Understanding the Genetic and Environmental Controls of Crumbly Fruit Disorder in Raspberry.* AHDB postgraduate meeting.
03-2018	***Presentation***	Understanding the causes of 'crumbly' fruit in raspberry – The James Hutton Institute postgraduate event.
01-2018	***Video***	*Project Legend: Education in Dundee and Angus - https://www.youtube.com/watch?v=7lQ4rMJ9nIA*
11-2017	***Poster***	Scolari, L.M., Graham, J., Steward D. 'Crumbly' fruit induction: mechanical damaging. How the removal of carpels or the spoilage of the receptacle can interfere with and/or compromise ovules fertilization and subsequent ovaries development into drupelets in red raspberry cv. Glen Ample. Agricultural and Horticultural Development Board (AHDB) - postgraduate meeting

References

Andrikopoulos, C.J. and Cane, J.H. (2018) Comparative Pollination Efficacies of Five Bee Species on Raspberry. J Econ Entomol, 111:2513-2519.

Berg, M., Rogers, R., Muralla, R. and Meinke, D. (2005) Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant Journal 44:866-878.

Cao K, Cui LR, Zhou XT, Ye L, Zou ZR, Deng SL (2016) Four Tomato FLOWERING LOCUS T-Like Proteins Act Antagonistically to Regulate Floral Initiation. Front Plant Sci 6.

Chang, M. and Huang, S. J. (2015) Arabidopsis ACT11 modifies actin turnover to promote pollen germination and maintain the normal rate of tube growth. Plant Journal 83:515-527.

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

Daviere, J. M. and Achard, P. (2013) Gibberellin signaling in plants. Development 140:1147-1151.

Daubeny, H. A., Stace-Smith, R. and Freeman, J. A. (1978) The occurrence and some effects of raspberry bush dwarf virus in red raspberry. J. Am. Soc. Horticult. Sci. 103:519-522.

Du, F.Y. Ruan, G.H. and Liu, H.W. (2012) Analytical methods for tracing plant hormones. Analytical and Bioanalytical Chemistry, 403:55-74.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. Plos One 6:(5)

Elwan, M.W.M. and El-Hamahmy, M.A.M. (2009) Improved productivity and quality associated with salicylic acid application in greenhouse pepper. Scientia Horticulturae, 122:521-526.

Escobar-Restrepo, J. M., Huck, N., Kessler, S., Gagliardini, V., Gheyselinck, J., Yand, W. C. and Grossniklaus, U. (2007) The FERONIA receptor-like kinase mediates male-female interactions during pollen tube reception. Science 317:656-60.

Feldman, M. J., Poirier, B. C. and Lange, B. M. (2015) Misexpression of the Niemann-Pick disease type C1 (NPC1)-like protein in Arabidopsis causes sphingolipid accumulation and reproductive defects. Planta 242:921-933.

Foster, T. M., Bassil, N. V., Dossett, M., Worthington, M. L. and Graham, J. (2019) Genetic and genomic resources for Rubus breeding: a roadmap for the future. Hortic Res-England 6. doi:ARTN 11610.1038/s41438-019-0199-2

Goetz, M. Hooper, L.C. Johnson, S.D. Rodrigues, J.C.M. Vivian-Smith, A. and Koltunow, A.M. (2007) Expression of aberrant forms of AUXIN RESPONSE FACTOR8 stimulates parthenocarpy in Arabidopsis and tomato. Plant Physiology, 145:351-366.

Graham, J., Hackett, C. A., Smith, K., Woodhead, M., Hein, I. and McCallum, S. (2009) Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. Theor Appl Genet 118 (6):1143-1155

Graham, J., Smith, K., McCallum, S., Hedley, P. E., Cullen, D. W., Dolan, A., Milne, L., McNicol, J. W. and Hackett, C. A. (2015) Towards an understanding of the control of 'crumbly' fruit in red raspberry. Springerplus, 4:223.

Gutkowska, M., Wnuk, M., Nowarowska, J., Lichocka, M., Stronkowski, M. M. and Swiezewska, E. (2015) Rab geranylgeranyl transferase beta subunit is essential for male fertility and tip growth in Arabidopsis. Journal of Experimental Botany 66:213-224.

Hackett, C. A. (2002). Statistical methods for QTL mapping in cereals. Plant Mol Biol 48:585-99.

Hackett, C. A., McLean, K., Bryan, G.J. (2013) Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. PloS one. https://doi.org/10.1371/journal.pone.0063939

Hackett, C. A., Milne, L., Smith, K., Hedley, P., Morris, J., Simpson, C. G., Preedy, K. and Graham, J. (2018) Enhancement of Glen Moy x Latham raspberry linkage map using GbS to further understand control of developmental processes leading to fruit ripening. Bmc Genetics, 19:59.

Hala, M., Soukupova, H., Synek, L. and Zarsky, V. (2010) Arabidopsis RAB geranylgeranyl transferase beta-subunit mutant is constitutively photomorphogenic, and has shoot growth and gravitropic defects. Plant Journal 62:615-627.

Haruta, M., Gaddameedi, V., Burch, H., Fernandez, D. and Sussman, M. R. (2018) Comparison of the effects of a kinase-dead mutation of FERONIA on ovule fertilization and root growth of Arabidopsis. Febs Letters, 592:2395-2402.

He, J. F., Zhao, X. Q., Laroche, A., Lu, Z. X., Liu, H. K. and Li, Z. Q. (2014) Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Front Plant Sci 5. doi:ARTN 48410.3389/fpls.2014.00484

Honda, I., Matsunaga, H., Kikuchi, K., Matuo, S., Fukuda, M. and Imanishi, S. (2017) Involvement of Cytokinins, 3-Indoleacetic Acid, and Gibberellins in Early Fruit Growth in Pepper (Capsicum annuum L.). Horticulture Journal 86:52-60.

Jennings, D. L. (1967) Balanced Lethals and Polymorphism in Rubus Idaeus. Heredity 22:465-479. doi:DOI 10.1038/hdy.1967.64

Jennings, D. L. (1967b) Observations on some instances of partial sterility in red raspberry cultivars. Horticultural Research 7:116–122

Jennings D. L. (1988) Raspberries and Blackberries: Their Breeding, Diseases and Growth, London, Academic Press Limited.

Jones, C. S., Iannetta, P. P. M., Woodhead, M., Davies, H., McNicol, R. J. and Taylor, M. A. (1997) The isolation of RNA from raspberry (Rubus idaeus) fruit. Molecular Biotechnology 8:219-221.

Kebriyaee, D., Kordrostami, M., Rezadoost, M. H. and Lahiji, H. S. (2012) QTL Analysis of Agronomic Traits in Rice Using SSR and AFLP Markers. Not Sci Biol, 4:116-123. https://doi.org/10.15835/nsb427501

Kim, S. I., Park, B. S., Kim, D. Y., Yeu, S. Y., Song, S. I., Song, J. T. and Seo, H. S. (2015) E3 SUMO ligase AtSIZ1 positively regulates SLY1-mediated GA signalling and plant development. Biochemical Journal, 469:299-314.

Kumar, R. Khurana, A. and Sharma, A. K. (2014) Role of plant hormones and their interplay in development and ripening of fleshy fruits. Journal of Experimental Botany, 65:4561-4575.

LeClere, S. Schmelz, E. A. and Chourey, P. S. (2008) Cell wall invertase-deficient miniature1 kernels have altered phytohormone levels. Phytochemistry, 69:692-699.

Mackay, T. F. (2001) The genetic architecture of quantitative traits. Annu Rev Genet 35:303-339.

McCallum, S., Simpson, C. and Graham, J. (2018) QTL Mapping and Marker Assisted Breeding in Rubus spp. In. London, p 164. doi:10.1007/978-3-319-99031-6

Moose, S. P. and Mumm, R. H. (2008) Molecular plant breeding as the foundation for 21st century crop improvement. Plant Physiol 147 (3):969-977. doi:10.1104/pp.108.118232

Murant, A. F., Chambers, J. and Jones, A. T. (1974) Spread of raspberry bush dwarf virus by pollination, its association with crumbly fruit, and problems of control. Ann. Appl. Biol. 77:271-281.

Nagpal, P. Ellis, C.M. Weber, H. Ploense, S.E. Barkawi, L.S. Guilfoyle, T.J. Hagen, G. Alonso, J.M. Cohen, J.D. Farmer, E.E. Ecker, J.R. and Reed, J.W. (2005) Auxin response factors

ARF6 and ARF8 promote jasmonic acid production and flower maturation. Development, 132:4107-4118.

Obroucheva, N. V. (2014) Hormonal regulation during plant fruit development. *Russian* Journal of Developmental Biology 45:11-21.

Ozga, J. A. and Reinecke, D. M. (2003) Hormonal interactions in fruit development. Journal of Plant Growth Regulation 22:73-81.

Pan, X. Q., Welti, R. and Wang, X. M. (2008) Simultaneous quantification of major phytohormones and related compounds in crude plant extracts by liquid chromatography-electrospray tandem mass spectrometry. Phytochemistry 69:1773-1781.

Pandolfini, T. Molesini, B. and Spena, A. (2007) Molecular dissection of the role of auxin in fruit initiation. Trends in Plant Science, 12:327-329.

Pieslinger, A. M., Hoepflinger, M. C. and Tenhaken, R. (2010) Cloning of Glucuronokinase from Arabidopsis thaliana, the last missing enzyme of the myo-inositol oxygenase pathway to nucleotide sugars. J Biol Chem 285:2902-10.

Preedy, K. F. and. Hackett, C. A. (2016) A rapid marker ordering approach for high-density genetic linkage maps in experimental autotetraploid populations using multidimensional scaling. Theor. Appl. Genet. 129(11):2117-2132

Qin, Y. Leydon, A. R., Manziello, A. Pandey, R. Mount, D. Denic, S. Vasic, B. Johnson, M. A. and Palanievelu, R. (2009) Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. PLoS Genet, 5:e1000621.

Quito-Avela, D. F., Lightle, D. and Martin, R. R. (2014) Effect of Raspberry bushy dwarf virus, Raspberry leaf mottle virus, and Raspberry latent virus on Plant Growth and Fruit Crumbliness in 'Meeker' Red Raspberry. Plant Disease 98:176-183.

Rajon, E. and Plotkin, J. B. (2013) The evolution of genetic architectures underlying quantitative traits. Proc Biol Sci 280 (1769):20131552. doi:10.1098/rspb.2013.1552

Santner, A., Calderon-Villalobos, L. I., and Estelle, M. (2009) Plant hormones are versatile chemical regulators of plant growth. Nat Chem Biol 5 (5):301-307. doi:10.1038/nchembio.165

Scofield, S., Dewitte, W. and Murray, J. A. H. (2007) The KNOX gene SHOOT MERISTEMLESS is required for the development of reproductive meristematic tissues in Arabidopsis. Plant Journal 50:767-781.

Serrani, J.C. Ruiz-Rivero, O. Fos, M. and Garcia-Martinez, J.L. (2008) Auxin-induced fruit-set in tomato is mediated in part by gibberellins. Plant Journal, 56:922-934.

Simura, J., Antoniadi, I., Siroka, J., Tarkowska, D., Strnad, M., Ljung, K., and NOVAK, O. (2018) Plant Hormonomics: Multiple Phytohormone Profiling by Targeted Metabolomics. Plant Physiol, 177:476-489.

Slater, A. Scott, N. W. and Fowler, M. R. (2008) Plant Biotechnology: The Genetic Manipulation of Plants, Oxford, Oxford University Press.

Steve, L. (2011) How to use statistics, Gosport.

Suzuki, M., Nakagawa, S., Kamide, Y., Kobayashi, K., Ohyama, K., Hashinokuchi, H., Kiuchi, R., Saito, K., Muranaka, T. and Nagata, N. (2009) Complete blockage of the mevalonate pathway results in male gametophyte lethality. Journal of Experimental Botany 60:2055-2064.

Tiwari, A., Vivian-Smith, A., Ljung, K., Offringa, R. and Heuvelink, E. (2013) Physiological and morphological changes during early and later stages of fruit growth in Capsicum annuum. Physiologia Plantarum 147:396-406.

Van Ooijen, J. W. (2004) MapQTL ® 5 Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV. Wageningen.

Vanneste, S. and Friml, J. (2009) Auxin: a trigger for change in plant development. Cell 136 (6):1005-1016.

Vanstraelen, M. and Benkova, E. (2012) Hormonal Interactions in the Regulation of Plant Development. Annual Review of Cell and Developmental Biology 28**:**463-487.

Vriezen, W. H. Feron, R., Maretto, F., Keijman, J. and Mariani, C. (2008) Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. New Phytologist 177:60-76.

Ward, J. A,, Bhangoo, J., Fernandez-Fernandez, F., Moore, P., Swanson, J. D., Viola, R., Velasco, R., Bassil, N., Weber, C. A. and Sargent, D. J. (2013) Saturated linkage map construction in Rubus idaeus using genotyping by sequencing and genome-independent imputation. Bmc Genomics 14. doi:Artn 210.1186/1471-2164-14-2

Waters, D.L. Holton, T. A., Ablett, E. M., Lee, L. S. and Henry, R. J. (2005) cDNA microarray analysis of developing grape (Vitis vinifera cv. Shiraz) berry skin. Funct Integr Genomics 5:40-58.

Wasternack, C. and Kombrink, E. (2010) Jasmonates: Structural Requirements for Lipid-Derived Signals Active in Plant Stress Responses and Development. Acs Chem Biol 5 (1):63-77.

Yu, J., Pressoir, G., Briggs, W. H., Vroh BI, I., Yamasaki, M. et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet, 38, 203-8.

Zwanenburg, B., Pospisil, T. and Zeljkovic, S. C. (2016) Strigolactones: new plant hormones in action. Planta 243 (6):1311-1326.

Appendices

Table A1: Mass spectrometry optimized parameters for the multiple reactions monitoring (MRM) method (part I).

Compound	[1]prec. ion	[2]prod. ion	dwell	[3]Fragm.	[4]c. e.	polarity	[5]tR
d-GA_1 [$^2H_2$]	351.2	333.2	10	105	4	positive	11.68
d-GA_1 [$^2H_2$]	351.2	305.2	10	105	4	positive	11.68
GA_1	349.2	331.1	10	95	20	positive	11.89
GA_1	349.2	285.1	10	95	16	positive	11.89
d-GA_3 [$^2H_2$]	347.1	241.2	10	150	12	negative	11.2
d-GA_3 [$^2H_2$]	347.1	143.1	10	150	32	negative	11.2
GA_3	345.1	239.2	10	155	12	negative	11.44
GA_3	345.1	143.1	10	155	24	negative	11.44
GA_4	331.1	287.2	10	220	16	negative	24.28
GA_4	331.1	257.1	10	220	20	negative	24.28
OPDA	291.2	247.2	10	175	12	negative	25.43
OPDA	291.2	165.1	10	175	16	negative	25.43
d-ABA [$^2H_6$]	269.2	225.2	10	190	8	negative	15.65
d-ABA [$^2H_6$]	269.2	159.1	10	190	8	negative	15.65
ABA	263.1	219.1	10	95	8	negative	16.04
ABA	263.1	153.1	10	95	4	negative	16.04
MeJA	225.1	151.1	10	85	8	positive	23.95
MeJA	225.1	133.1	10	85	12	positive	23.95
[*]d-Z. [$^2H_5$]	225.1	137	10	100	16	positive	7.2
[*]d-Z. [$^2H_5$]	225.1	135.9	10	100	16	positive	7.2
Zeatin	220.1	202.1	10	90	8	positive	7.32
Zeatin	220.1	136	10	90	16	positive	7.32
d-JA [$^2H_5$]	214.2	62.1	10	130	8	negative	18.95
d-JA [$^2H_5$]	214.2	42.1	10	130	48	negative	18.95
JA	209.1	109	10	113	16	negative	19.36
JA	209.1	59.1	10	113	8	negative	19.36
IBA	204.1	186.1	10	90	8	positive	18.95
IBA	204.1	130.1	10	90	28	positive	18.95
[ł]d-Me. [$^2H_5$]	195.1	135	10	90	16	positive	16.6
[ł]d-Me. [$^2H_5$]	195.1	134.1	10	90	12	positive	16.6
MeIAA	190.1	130	10	75	12	positive	17.14
MeIAA	190.1	77.1	10	75	50	positive	17.14

[1]precursor ion - [2]product ion - [3]fragmentator - [4]collision energy - [5]retention time (min)
[*]d-Zeatin [2H5]
[ł]d-MeIAA [$^2H_5$]

For each of the 18 target compounds and the 9 isotopically labelled internal standards the best parameters for their detection were reported in the table. These parameters were, the mass of the precursor ion and of its product after fragmentation reaction (product ion), the

dwell time (the duration in which each m/z (mass to charge ratio) ion signal is collected) of 10 ms. The fragmentation voltage that controls the speed at which the ion moves from the electrospray chamber to the mass spectrometer and the collision energy that controls the rate of acceleration as the ions enter the quadrupole 2 (Q2) and then regulate fragment ion intensity. The best ionization mode (i.e. positive or negative) and the retention time at which each compound is eluted through the chromatographic column were reported too.

Table A2: Mass spectrometry optimized parameters for the multiple reactions monitoring (MRM) method (part II).

Compound	[1]prec. ion	[2]prod. ion	dwell	[3]Fragm.	[4]c. e.	polarity	[5]tR
d-IAA [$^2H_2$]	178.1	132	10	90	12	positive	12.6
d-IAA [$^2H_2$]	178.1	78.1	10	90	48	positive	12.6
IAA	176.1	130.1	10	85	12	positive	12.89
IAA	176.1	77.1	10	85	50	positive	12.89
MeCA	163.1	131	10	70	8	positive	23.62
MeCA	163.1	103.1	10	70	20	positive	23.62
ICA	162.1	144	10	80	12	positive	12.47
ICA	162.1	118.1	10	80	12	positive	12.47
MeSA	153.1	121	10	85	12	positive	21.48
MeSA	153.1	65.1	10	85	32	positive	21.48
d-CA [$^2H_7$]	154.1	110.1	10	115	12	negative	16.97
MeBA	137.1	100	10	85	4	positive	18.41
MeBA	137.1	77.1	10	85	28	positive	18.41
d-SA [$^2H_4$]	141.1	97	10	90	16	negative	15.1
d-SA [$^2H_4$]	141.1	69.1	10	90	31	negative	15.1
SA	137	93.1	10	75	16	negative	15.49
SA	137	65.1	10	75	31	negative	15.49
BA	121	77.1	10	105	8	negative	12.98

[1]precursor ion - [2]product ion - [3]fragmentator - [4]collision energy - [5]retention time (min)

*d-Zeatin [2H5]

ǂd-MeIAA [$^2H_5$]

For each of the 18 target compounds and the 9 isotopically labelled internal standards the best parameters for their detection were reported in the table. These parameters were, the mass of the precursor ion and of its product after fragmentation reaction (product ion), the dwell time (the duration in which each m/z (mass to charge ratio) ion signal is collected) of 10 ms. The fragmentation voltage that controls the speed at which the ion moves from the electrospray chamber to the mass spectrometer and the collision energy that controls the rate of acceleration as the ions enter the quadrupole 2 (Q2) and then regulate fragment ion intensity. The best ionization mode (i.e. positive or negative) and the retention time at which each compound is eluted through the chromatographic column were reported too.

Table A3. List of 63 different genotypes of which 45 specifically selected for their tendency to show 'crumbly' symptoms and 18 on the contrary for having never displayed the symptoms.

Pedigree					
♀ mother	x	♂ father	genotype		phenotype
DJ1185	x	8510A21	0560E11	FC	crumbly
Tulameen	x	8510A28	0565F3	FC	crumbly
Malling Minerva	x	8510A41	0663RE3	FC	crumbly
R4A1	x	Glen Fyne	0946/4	FC	crumbly
7826C1	x	8627RE7	9059D-2	FC	crumbly
9421A4	x	9434B-1	99105RC-2	FC	crumbly
9426C-5	x	9429E-2	99116E-4	FC	crumbly
SCRI 8631D-1	x	SCRI 8605C-2	Glen Fyne	FC	crumbly
9059D-2	x	8510A73	0019E2	FC	crumbly
9455F-2	x	8510A5	0304F6	FC	crumbly
0096RF-4	x	8510A6	04101A5	FC	crumbly
0003RB1	x	8510A9	0433F2	FC	crumbly
9046RA2	x	8510A22	0511F1	FC	crumbly
Glen Fyne	x	8510A29	0573B5	FC	crumbly
0312E3	x	8510A35	0658C5	FC	crumbly
0312E3	x	8510A32	0658E-1	FC	crumbly
0312E3	x	8510A37	0658F-7	FC	crumbly
Glen Ample	x	8510A43	0671D-4	FC	crumbly
97134B1	x	8510A57	0867E-4	FC	crumbly
R4A1	x	Glen Fyne	0946/19	FC	crumbly
8735J-7	x	8626RJ-2	9050RD3	FC	crumbly
8820E3	x	88K-7	9238D5	FC	crumbly
EM5961/1	x	7826C1	9350F3	FC	crumbly
WSU1068	x	ORUS 2078	97134B1	FC	crumbly
9349F5	x	9349A4	9764F-3	FC	crumbly
9351D-3	x	9350E1	9769RD1	FC	crumbly
	x		Autumn Bliss	PF	crumbly
Autumn Bliss[1]	x		Erika	PF	crumbly
7326E1	x	7412H6	Glen Ample	FC	crumbly
00123A5	x	0019B11	Glen Dee	FC	crumbly
SCRI 7331/1	x	SCRI 7256/1	Glen Lyon	FC	crumbly
SCRI 688/12	x	SCRI 6815/113	Glen Moy	FC	crumbly
SCRI 6531/84	x	SCRI 6549/1	Glen Prosen	FC	crumbly
7326E1	x	7412H16	Glen Rosa	FC	crumbly

Pedigree					
♀ mother	x	♂ father	genotype		phenotype
7741D4	x	7919B11	Glen Shee	FC	crumbly
	x		Imara	PF	crumbly
Autumn Bliss	x		Joan Squire	PF	crumbly
	x		Kweli	PF	crumbly
King	x	Louden	Latham	FC	crumbly
Willamette	x	Cuthbert	Meeker	FC	crumbly
(unknown)	x		Obbard	PF	crumbly
Nootka	x	Glen Prosen	Tulameen	FC	crumbly
9431G-8	x	8510A71	Sanibelle	FC	no crumbly
0015D3	x	8510A14	0453C4	FC	no crumbly
	x		00123A7	FC	no crumbly
Glen Rosa	x	8510A20	0550E4	FC	no crumbly
0304F6	x	Autumn Treasure	0925B4	PF	no crumbly
0304F6	x	Autumn Treasure	0925B8	PF	no crumbly
0304F6	x	Autumn Treasure	0925D-15	PF	no crumbly
8003G10	x	8003C1	8605C-2	FC	no crumbly
8020E8	x	8631D-1	9025A1	FC	no crumbly
complex	x		Autumn Britten	PF	no crumbly
EM6304/36	x	EM6330/96	Autumn Treasure	PF	no crumbly
SCRI 9422C-4	x	SCRI 9434B-1	Glen Cally	FC	no crumbly
0030E-12	x	8510A16	Glen Carron	FC	no crumbly
Glen Rosa	x	SCRI 8605C-2	Glen Doll	FC	no crumbly
SCRI 9422C-4	x	SCRI 9434B-1	Glen Ericht	FC	no crumbly
	x		Malling Leo	FC	no crumbly
Joan Squire	x	complex	Brice	PF	crumbly
	x		Chief	FC	no crumbly
Joan J	x	complex	Joan Irene	PF	crumbly
EM3689	x	Gaia	Malling Hiesta	FC	crumbly
EM selection	x	SCRI selection	Malling Minerva	FC	no crumbly
Lloyd George	x	Preussen	Schoenemann	FC	crumbly

[1]open pollinated - FC (floricane) – PF (primocaine)