Cloning and Characterization of the dominant *Inhibitor of Wax 1* (*Iw1*) gene in polyploid wheat

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Faculty of Biological Sciences

University of East Anglia

Submitted by Nikolai M. Adamski

Norwich

30th September 2013

To my family

Table of Contents

Figures

Tables

Abbreviations

AFLP	Amplified fragment length polymorphism
amu	atomic mass unit
BAC	Bacterial artificial chromosome
BC	Back cross
BES	BAC end sequence
bp	base pair
cM	centiMorgan
ddH_2O	Double distilled water
DH	Doubled-haploid
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxy-nucleotidetriphosphate
EMS	Ethyl methanesulfonate
ER	Endoplasmic reticulum
EST	Expressed sequence tag
FAE	Fatty acid elongase
FAS	Fatty acid synthase
GC-MS	Gas chromatography-mass spectrometry
InDel	Insertion / Deletion
KASPar	KBiosciences competitive allele specific PCR genotyping system
kb	kilo base pair
KCS	Type III polyketide synthases
LTP	Lipid transfer protein
Mb	Mega base pair
M-MLV	Moloney murine leukemia virus
NIL	Near isogenic line
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
PAGE	Polyacrylamide gel electrophoresis
PAR	Photosynthetically active radiation
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
RNA-seq	RNA sequencing
rpm	rounds per minute
RSL65	Recombinant substitution line 65
RT	Room temperature
RT-PCR	Reverse transcriptase - polymerase chain reaction
SEM	Scanning electron microscope
sem	standard error of the mean
SSCP	Single-strand conformation polymorphism
SSR	Simple sequence repeat
TE	Transposable element

TEM	Transmission electron microscope
TIC	Total ion chromatogram
TLC	Thin layer chromatography
TTD	*Triticum turgidum* ssp. *dicoccoides*

Abstract

The *Inhibitor of Wax 1* (*Iw1*) is a dominant inhibitor of glaucousness, the whitish waxy bloom present on the aerial surfaces of a wheat plant: The presence of *Iw1* leads to non-glaucousness. In previous work a doubled-haploid population segregating for the presence of the *Iw1* locus was created. The non-glaucous doubled-haploid lines of this population showed increases in yield and green-canopy duration under UK conditions compared to their glaucous counterparts by on average 4.15% and 1.5 days, respectively.

The aim of this study was to identify *Iw1* via a positional cloning approach and to characterize its effects on yield and green-canopy duration in a field-grown set of glaucous and non-glaucous Near Isogenic Lines (NILs). In addition a number of physiological experiments were carried out on these NILs to determine the effects of non-glaucousness on light reflectance and transmission as well as on water-use efficiency (WUE). Finally, the composition of surface waxes in glaucous and non-glaucous NILs was elucidated using a combination of electron microscopy and biochemical methods.

Here, we have fine-mapped *Iw1* to a 0.42-cM interval on the short arm of chromosome 2B and we have constructed a physical map, which is currently 1,200 kb in size. Gene models were predicted *in silico* and we have begun to test candidate genes using allelic diversity and expression analysis. The results of our physiological experiments clearly show a reduction in light reflectance and a possible increase in light transmission through the canopy leaves in non-glaucous NILs. We could not detect a negative effect on WUE in field-grown NILs nor did we identify significant increases in yield. A consistent extension in green-canopy duration was associated with the *Iw1* region, although not significant in all years. Our analysis of the composition of surface waxes has shown that only a discrete type of wax, the β–diketone aliphatics, is being inhibited by *Iw1*.

General Introduction

Origin of wheat

The invention of agriculture marks one of the biggest turning points in the history of mankind. Approximately 10,000 years ago humans started to cultivate cereals for their seeds, thus changing their lifestyle from that of hunter-gatherers to sedentary farmers. Einkorn wheat (*Triticum monococcum*, genomes A^mA^m) and barley (*Hordeum vulgare*, genomes HH) were the first cereals to be domesticated (Nesbitt and Samuel 1996) in southeastern Turkey (Heun et al. 1997) and northern Israel, respectively (Badr et al. 2000).

The cultivation of tetraploid emmer wheat (*Triticum turgidum* ssp. *dicoccon*, tetraploid, genomes AABB) from wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*, genomes AABB) (Luo et al. 2007) marked a significant advance in wheat agriculture. Wild emmer wheat was formed via the hybridization of two diploid grasses; the A-genome was donated from *Triticum urartu* (genomes A^uA^u), while the B-genome donor is an unknown member of the *Sitopsis* section closely related to *Aegilops speltoides* (genomes SS) (Feldman et al. 1995, Huang et al. 2002). The tetraploid emmer wheat proved to be more vigorous, high-yielding and more adaptable than its diploid progenitors and was the source for many subspecies that spread around the world and were cultivated for thousands of years. One of these subspecies, *Triticum turgidum* ssp. *durum* (genomes AABB) gave rise to our modern day pasta wheat cultivars *Triticum durum* (genomes AABB) (Feldman 2001, Gepts 2003).

The second major step in wheat agriculture occurred ~10,000 years ago when emmer wheat spread northeast towards the Caspian Sea where it hybridized with *Aegilops tauschii* (genomes DD) (Dvorak et al. 1998) to form hexaploid common wheat (*Triticum aestivum*, genomes AABBDD) (Kihara 1944), from which modern bread wheat (*Triticum aestivum* ssp. *aestivum*) cultivars are derived. As was with emmer wheat, hexaploid bread wheat was superior in terms of viability, yield and adaptability to its progenitor species. The addition of the D-genome also modified the end use characteristics of wheat by the introduction of genes controlling endosperm texture (*Pina* and *Pinb*) (Feldman et al. 1995). This increased bread making quality of hexaploid wheat compared to the very hard endosperm of *T. durum* (Morris 2002).

Different grades of grain hardness were subsequently selected for by farmers in addition to many other traits before and afterwards, all of which led to a clear differentiation between the cultivated forms of wheat and their wild ancestors, also referred to as the domestication syndrome (Hammer 1984). The key domestication event for wheat, barley and other cereals was the reduction of spike-

shattering, which facilitated grain harvest and prevented spikelets (and thus grains) to be scattered by the wind, rain and/or animals. In wheat, this trait was shown to be controlled by the *brittle rachis* (*Br*) loci on chromosomes 3A and 3B (Nalam et al. 2006). Other domestication traits in wheat include the change from hulled wheat to free-threshing wheat by softening of the glumes holding the grain, which reduced the amount of labour required to separate the grain from the spikelets. This trait is predominantly affected by the *tenacious glume* (*Tg*) locus on chromosome 2D and to a lesser extent by the *Q* gene on chromosome 5A (Jantasuriyarat et al. 2004). In addition, the *Q* gene, which encodes an *AP2*-like transcription factor, pleiotropically affects other domestication traits like rachis fragility, spike length or the square spike phenotype (Simons et al. 2006). Interestingly, the mutation that caused the *Q*-allele is identical between free-threshing tetra- and hexaploid wheat, indicating that it arose only once (Simons et al. 2006). Other domestication traits involve an increase in seed size, a reduction in tiller number and a more erect growth habit. Increases in seed and spike sizes in wheat over the past millennia have greatly boosted yields but at the cost of a higher risk of lodging. The introduction of semi-dwarfing genes from Japanese cultivars, also called the 'Green Revolution Genes', reduced the height of wheat plants, which increased the Harvest Index (the ratio of grain yield to the aboveground tissue at maturity (Huehn 1993)) and also allowed the application of higher amounts of nitrogen fertilizer to the crop without the risk of lodging, thus greatly boosting wheat yields around the globe (Hedden 2003, Borojevic and Borojevic 2005).

World Production

Wheat is a major provider of the calories and protein consumed worldwide, having supplied as many calories (kcal/capita/day) and one and a half times as much protein (g/capita/day) as rice in 2009 (FAOSTAT, Figure 1). But wheat production, and food production in general, need to rise by as much as 70% in the next 40 years, as the world population is predicted to increase from currently seven billion to nine billion by the year 2050 (FAO 2009). In order to meet these demands the annual increase in wheat yield needs to rise from currently <1% to approximately 1.7% (Initiative 2013). A second Green Revolution is needed to solve these problems.

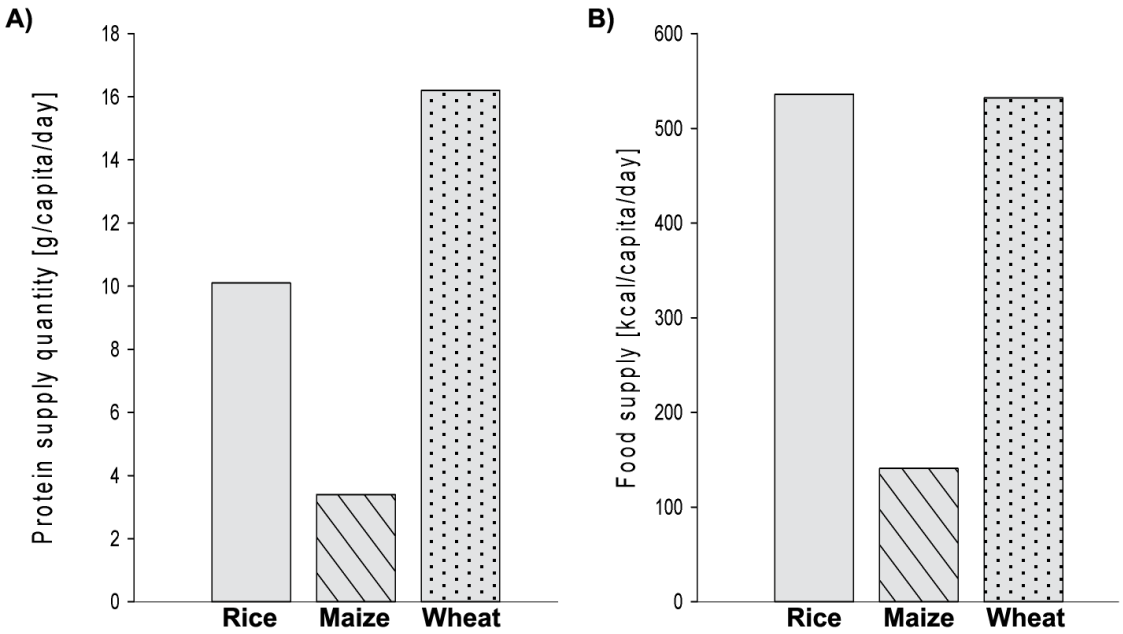


Figure 1: Calories and protein provided worldwide by rice, maize and wheat
These graphs display the amount of calories and protein provided worldwide by rice, maize and wheat in 2009. (A) Wheat supplies as many calories per capita per day as rice and five times as many as maize. (B) Wheat supplies one-and-a-half times more protein per capita per day as rice and five times as much as maize.

Domestication and the use of wild progenitor and alien species

Domestication relies on selecting individuals with desirable traits. This, however, creates diversity bottlenecks and reduces the genetic variability of cultivated plants in comparison to their wild ancestors (Tanksley and McCouch 1997, Buckler et al. 2001). Wheat breeding has mostly relied on developing new allele combinations within the pool of domesticated germplasm to achieve increases in production. But in order to meet the food security challenges of the future, new and useful alleles need to be identified so they can be deployed in adapted germplasm, just like Norman Borlaug did ~60 years ago.

There are several possible sources to introduce new alleles into elite wheat germplasm. Over the last millennia farmers have grown and selected wheat varieties that performed best in their fields, thus creating genetically diverse and dynamic populations adapted to specific environments. These landraces have been largely abandoned by farmers in the developed world with the advent of high-yielding wheat varieties and the adoption of high input wheat farming systems. However, landraces represent a vast resource of new allelic variation that can be easily transferred into modern varieties (Dreisigacker et al. 2005, Ghimire et al. 2005, Zhang et al. 2006, Wingen et al. 2012).

New genetic variation can also be found in the birthplace of agriculture. Modern wheat is the result of allopolyploidisation of a few wild and cultivated individuals, which comprises only part of the genetic diversity that is present in the wild populations. A nucleotide diversity analysis of 21 genetic loci revealed an estimated loss of diversity of ~75% for *T. durum*, ~55% for the A- and B- genome of *T. aestivum* and a staggering ~90% for the D-genome of *T. aestivum* (Haudry et al. 2007). This highlights the possibilities offered by wild ancestral species, especially for the wheat D-genome. This was realized by scientists at CIMMYT and at NIAB, where *Aegilops tauschii* accessions were crossed to a wide selection of tetraploid wheat lines to recreate the formation of hexaploid wheat, thus creating so-called synthetic wheat. These synthetics lack a good agronomic phenotype, but they can be incorporated into pre-breeding programs to boost diversity and create new allele combinations. The CIMMYT synthetics have been used successfully in Chinese breeding programs, increasing kernel size and spike weight and providing new resistance alleles against stripe rust (CIMMYT 2004).

Alternatively, wild wheat progenitors can be crossed directly to modern hexaploid wheat. This method has been used successfully in the past to introduce new traits like increased photosynthetic activity (Carver et al. 1989), an increased grain protein content (Gerechter-Amitai and Stubbs 1970), an increased kernel weight (Kushnir and Halloran 1984) or increased disease resistance against leaf rust (Nevo et al. 1986, Fahima et al. 1998), stem rust (Nevo et al. 1991) and powdery mildew (Nevo et al. 1985) from wild emmer wheat into modern varieties.

Alien grass species can also be used to introduce new traits into wheat, but this requires the use of mutant wheat line. The *Ph1* locus is located on wheat chromosome 5B and prevents the pairing and thus recombination of non-homologous chromosomes (Griffiths et al. 2006). Thus, *ph1* mutant wheat lines are needed to introgress desirable traits from alien grass species such as the *Lr19* disease resistance gene from *Lophopyrum ponticum* (Zhang et al. 2005). However, these crosses often form a low frequency of chromosome pairing which can cause deleterious characters to be transferred together with the desired genes (Islam and Shepherd 1991). These deleterious chromosome segments, the so-called linkage drag, often require many back-crossing steps to be completely broken up. Molecular markers can help to monitor this process though (Fedak 1999).

Mapping desirable traits in polyploid wheat

Genetic maps are essential in identifying regions in the genome that contribute towards specific, desirable phenotypic features, like early/late flowering, plant height or yield: These genomic regions are called quantitative trait loci (QTL). There are also less complex phenotypic features that are

controlled by single genetic loci, like major disease resistance loci. DNA markers are used to map genetic loci to certain intervals. Genetic maps have been also used in marker-assisted selection for breeding, the development of comparative maps as well as high resolution screening and map-based cloning. Approximately 80 interactive genetic maps are available on the GrainGenes website, but research groups around the world have likely produced hundreds of genetic maps in wheat so far, using a variety of different markers.

Early genetic maps have been based on restriction fragment length polymorphisms (RFLPs)(Chao et al. 1989, Devos et al. 1992). The advent of PCR-based markers like random amplified polymorphic DNAs (RAPDs)(Williams et al. 1990, Devos and Gale 1992) or amplified fragment length polymorphisms (AFLPs)(Vos et al. 1995), but especially of simple sequence repeats (SSRs)(Röder et al. 1998, Gupta et al. 2002) revolutionized the speed at which genetic maps were created and also their accuracy. Over the last decade single nucleotide polymorphisms (SNPs) were described as a new source for markers (Gupta et al. 2001, Rostoks et al. 2005). These SNP based maps can be easily compared between different varieties as the SNPs are usually discovered in genic sequence, although their bi-allelic nature makes them less-informative than SSR. Combined with new low-cost high-throughput genotyping systems (KASPar) the creation of high-density genetic maps takes only a few weeks.

Synteny between grass species

Grass species have diverged over a period of 60 million years from a common ancestor via multiple chromosome rearrangements, chromosome deletions and whole genome duplications (WGD) (Gale and Devos 1998). As a result grass genomes differ in haploid chromosome number from for example five chromosomes in *Brachypodium distachyon* to twelve chromosomes in *Oryza sativa*. Likewise, the genome size can range from for example 270 Mb in *Brachypodium distachyon* to 5,000 Mb in the various Triticeae genomes (Salse et al. 2009). Despite these differences the gene content and order among the grass species has remained remarkably conserved (Moore et al. 1995); this concept is referred to as synteny (Bennetzen and Freeling 1997).

This conservation can be used for comparative genomics between cereal or grass species; with the switch to SNP based genetic maps, these comparisons have become a lot simpler and more powerful. The comparison between genomes can help to enhance the map-density in specific regions for map-based cloning approaches (Kilian et al. 1997) and it also allows the establishment of genetic maps in less well studied species (Van Deynze et al. 1998), as for example sugarcane (Asnaghi et al. 2000).

Map-based cloning in polyploid wheat

The polyploid wheat genome represents a challenge for any map-based cloning approaches due to its vast size (~5 Gb per genome), closely related genomes (~95% similar in genic regions) and high content of transposable elements (TEs, ~80%); despite these challenges genes have been successfully cloned in wheat. Map-based cloning rests on the creation of a physical map of the genetic region of interest. While this is a slow process compared to the candidate gene approach (Pflieger et al. 2001) it is also an unbiased method in that no assumptions about the locus of interest are made prior to its identification. Thus, the approach is very much in keeping with the scientific method.

A map-based cloning approach can be divided into three steps. The first step is the genetic mapping of the locus of interest. Most successful cloning projects have screened a large number of gametes (>5,000) for recombinants to obtain a small genetic map interval (Yan et al. 2003, Uauy et al. 2006) although this is not always necessary (~1,000 gametes screened in (Faris et al. 2003, Huang et al. 2003)).

The second step is the creation of a physical map, which is dependent on the genetic mapping phase. In general, the more gametes that have been screened, the smaller the genetic interval becomes and the smaller effort that is needed to create the physical map. This step requires the creation of a genomic library, usually either bacterial artificial chromosome (BAC) or yeast artificial chromosome (YAC) libraries. These libraries are used to create a physical contig of the genetic interval, which can vary in size from 0.36 Mb/cM (Faris et al. 2003) to more than 13 Mb/cM (Yan et al. 2003).

The third step is the identification and validation of candidate genes inside the physical map. Usually a mix of several methods, like allelic diversity, expression analysis and mutant analysis are used to validate candidate genes, as no one single method guarantees success. Genetic complementation is another effective method to test candidate genes and is often used for single effect loci like resistance genes (Huang et al. 2003).

Current publicly available genomic resources

Over the last few years the number of genomic resources available to wheat geneticists has increased dramatically. Although wheat is still lacking a complete and fully assembled genomic sequence, individual wheat chromosome arms have been separated by flow-cytometry (Vrána et al. 2000) and sequenced individually to create a chromosome arm specific genomic sequence database (International Wheat Genome Sequencing Consortium 2011). The sequences have been assembled

into contigs of varying size and researchers can perform BLASTN queries to obtain genome- and chromosome arm-specific sequence. This tool was used extensively in the present study for marker design and for the physical mapping stage of the positional cloning approach.

Recently more than 100,000 SNPs in hexaploid wheat were detected using the NimbleGen Capture Array, 10-20% of which are estimated to be co-dominant based on the experimental validation of a subset of the SNPs (Allen et al. 2011, Winfield et al. 2012, Allen et al. 2013). The presence of SNPs has been studied in 169 wheat varieties, enabling researchers to design variety-specific markers or use a set of existing genome- and variety-specific KASPar markers. All this information is integrated in a freely accessible database (Wilkinson et al. 2012). This resource is extremely helpful for map-based cloning projects as it offers a set of high-density markers as well as information on the level of polymorphism between wheat varieties.

The transcriptome of diploid and tetraploid wheat was recently sequenced (Krasileva et al. 2013). This data can be used as a reference in RNA-seq studies, replacing the incomplete and collated wheat UniGene list, but also as a reference for proteomics studies in wheat, as the predicted open reading frames (ORFs) have been translated and functionally annotated. We have used this data in the present study as a reference for a RNA-seq experiment to assign reads to the different wheat homoeologous.

Targeting induced local lesions in genomes (TILLING) populations of tetraploid and hexaploid wheat have been generated using EMS-mediated mutagenesis (Uauy et al. 2009). This reverse genetics platform allows wheat researchers to easily obtain mutant lines for their studies. The power of a mutant collection for the identification of new genetic alleles and to perform functional genetics in wheat cannot be overstated.

These new resources represent a tremendous step forward in wheat genetics and will undoubtedly prove to be invaluable tools for the isolation, validation and characterization of genes and gene networks in the future.

Project background

Glaucousness is the production of a waxy bloom on the aerial surfaces of a plant, which gives plants a whitish appearance. There is little variation in terms of glaucousness in modern UK varieties, although some Recommended List varieties differ in their relative visual glaucousness. In addition, some varieties, including Shamrock and derivatives (Gulliver and Crusoe) have a bright green appearance because of a complete lack of this waxy bloom. Shamrock was derived from a cross between a *Triticum turgidum* ssp. *dicoccoides* (wild emmer wheat) derivative and a UK breeding line (CWW 4899/25—Moulin x Monopol). A recent study has examined the performance of a doubled-haploid (DH) population made between the glaucous cultivar Shango and the non-glaucous cultivar Shamrock in several locations over consecutive years in the field (Simmonds et al. 2008). The study concluded that the non-glaucous DH lines outperformed their glaucous counterparts in yield on average by 4.15% over a three year trial period. Likewise, the green-canopy duration of the non-glaucous DH lines was extended by 1.5 days compared to their glaucous siblings. The non-glaucous phenotype, the increase in yield as well as the increase in green-canopy duration were mapped to the *T. dicoccoides* introgression on the distal end of the short arm of chromosome 2B. The non-glaucous phenotype was attributed to a single locus named *Viridescence* (*Vir*), but it was unclear whether the increases in yield and green-canopy duration were pleiotropic effects of *Vir* or not (Simmonds et al. 2008).

We hypothesize that *Vir* is identical to the dominant inhibitor of glaucousness *Inhibitor of Wax 1* (*Iw1*). The aim of the present study was to identify *Iw1* via positional cloning and validate it. The effects of *Iw1* on yield and green-canopy duration were to be tested in a set of Near Isogenic Lines (NILs) in the field. In addition, several physiological parameters like light reflectance, light transmission and water-use efficiency were to be measured in these NILs as well. Furthermore, the effect of *Iw1* on the composition of surface waxes was to be elucidated using a combination of microscopy and biochemical techniques.

Chapter 1 The effects of *lw1* on cuticular wax layer composition and other plant physiological properties

Introduction

Plants have evolved in an aquatic environment and needed to adapt their morphology before being able to succeed in a terrestrial environment. One of the key adaptations of plants that colonized dry land is the cuticle, which forms a thin and continuous membrane around the outer surface of aerial plant organs (Edwards et al. 1998). It consists of a polymer matrix (cutin), polysaccharides, as well as monomeric, and thus solvent-soluble, lipids (cuticular waxes)(Holloway 1982, Jeffree 1996). While cutin mainly consists of short C_{16} and C_{18} fatty acids (Heredia 2003) the cuticular wax is a complex mixture of C_{20} to C_{60} aliphatics including secondary metabolites such as flavonoids, triterpenoids and phenylpropanoids (Jetter et al. 2007).

Cuticular waxes are present throughout the entire width of the cuticle (Jeffree 1996) and can be divided into two classes based on their position in the cuticle. Intracuticular waxes are interspersed within the cutin polymer matrix while epicuticular waxes are present on the outer surface of the cutin matrix. Intracuticular waxes, and the cuticle itself, have been associated with preventing non-stomatal water loss, thus forming a key adaption of land-living plants (Edwards et al. 1982, Kerstiens 1996), while no counterpart of cutin is known in algae (Gray and Boucot 1977).

The epicuticular wax layer has been shown to have profound effects on the interaction of a plant with its environment. Depending on its composition it influences plant-insect interactions (Ni et al. 1998, Morris et al. 2000, Cervantes et al. 2002), the germination and appressorium formation of biotrophic fungi like powdery mildew (Hegde and Kolattukudy 1997, Tsuba et al. 2002) or the reflection of solar radiation by forming a waxy bloom on the plant surface (Johnson et al. 1983). This waxy bloom is also referred to as glaucousness and it gives plants a whitish or bluish appearance (see Figure 1.1).

Glaucousness causes a whitish or bluish appearance of plant surfaces. Most UK and European wheat varieties are glaucous like the plot on the right in Figure 1.1, while non-glaucous plants appear green like the left plot in Figure 1.1.

The epicuticular wax layer can vary in composition and quantity of individual compounds not only between individual plants, but also between plant organs and developmental stages. In wheat and barley, glaucousness first becomes visible to the naked eye during stem elongation when the flag leaf starts to emerge. In contrast, maize leaves are glaucous at the juvenile stage, but after the switch to

the reproductive stage new leaves are glossy or non-glaucous. In various plant species, like wheat, barley, eucalyptus and blue tussock, glaucousness was suggested to be the result of the presence of β-diketones in the epicuticular wax (Hall et al. 1965, Barber and Netting 1968, von Wettstein-Knowles 1972).

Figure 1.1: Field plots of wheat exhibiting non-glaucous (front left) and glaucous (front right) phenotypes.

The lipid components of the epicuticular wax layer, including β-diketones, have several basic steps in their biosynthetic pathway in common. Fatty acid synthase (FAS) multienzyme complexes located in the plastids form carbon chains by adding activated units of malonyl-CoA to a carbon acceptor, initially C_2. Through reiteration of this process, fatty acids of chain length C_{16} or longer are formed (von Wettstein-Knowles 1976, Lodwig et al. 2005). These precursor molecules are then exported, by an unknown mechanism, from the plastids into the endoplasmic reticulum (ER), where they become, targets for fatty acid elongase (FAE) multienzyme complexes or Type III polyketide synthases (KCSs). The first complex produces very long chain acyl-CoAs, predominantly C_{30}, which can be further modified to become alkanes, aldehydes, esters, ketones, primary or secondary alcohols (see Figure 1.2A) (Von Wettstein-Knowles 1974). The latter complex elongates the fatty acid primers from both the FAS and FAE pathway, but takes 'shortcuts' in doing so, i.e. keto- and hydroxy-groups, as well as double bonds, are retained during elongation. The resulting carbon chains act as precursors for a three-branch pathway, which either leads to alkylresorcinols, alkan-2-ols and alkan-2-ol-esters, or to β- and hydroxy-β-diketones (see Figure 1.2B) (Von Wettstein-Knowles 1974).

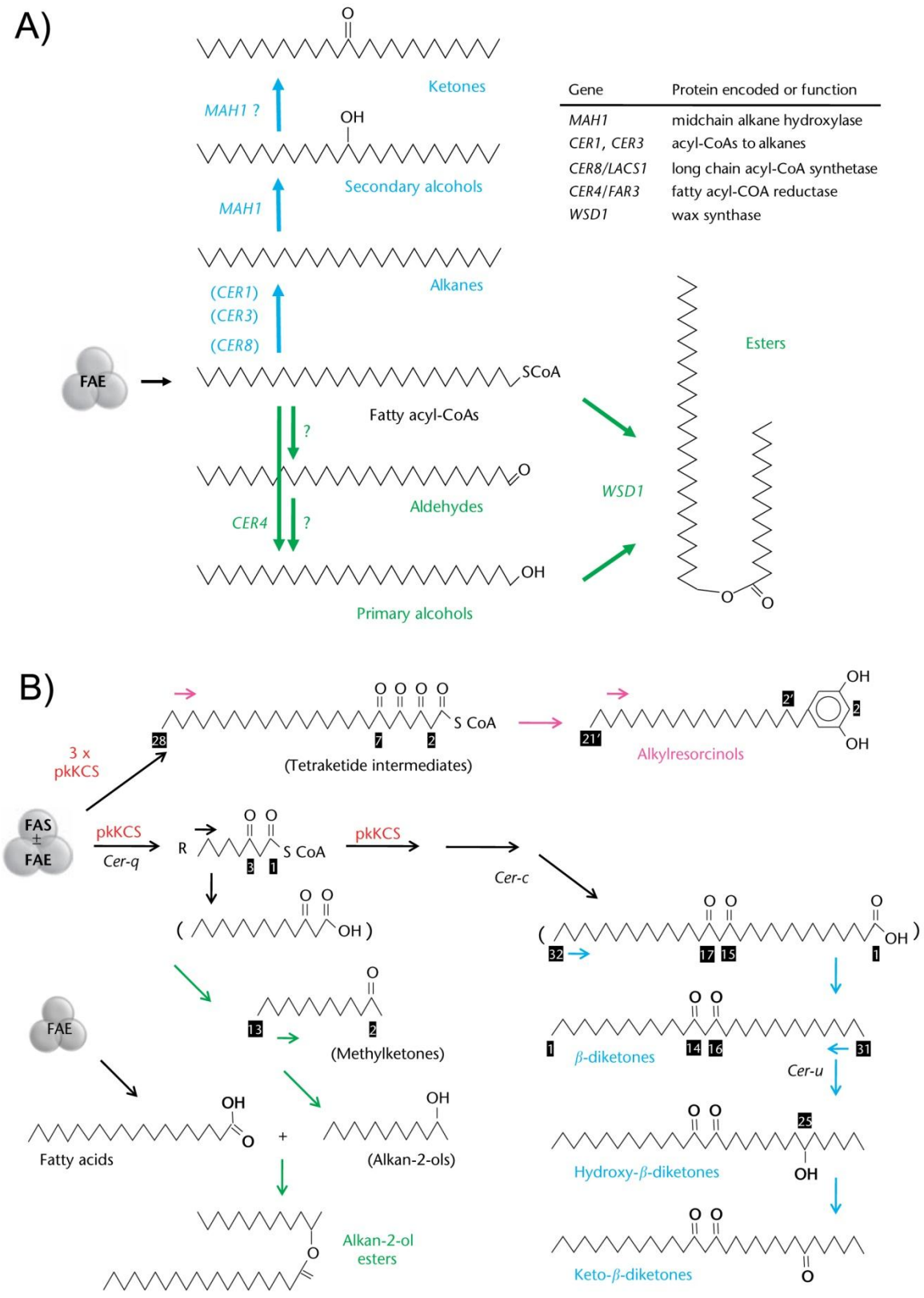


Figure 1.2: Biosynthetic pathways of lipid components in plants

Acyl chains from the FAS pathway get exported to the ER where they become targets for various types of FAE complexes. The classic pathway (A) is common to all plant species, whereas the β- and hydroxy-β-diketone pathway depicted in panel B is not present in all plant species. Figures taken from von Wettstein-Knowles (2012).

The components of the epicuticular wax layer are produced in the epidermal cell layer only. After completion of synthesis, the wax compounds are delivered to the plasma membrane (PM), although there is little evidence on how this is accomplished; transfer at the site of contact between the ER and the PM as well as vesicle mediated trafficking have been proposed, but none has been yet conclusively proven (Figure 1.3) (von Wettstein-Knowles 2012). Once the wax compounds have reached the PM, they are exported to the cuticle, probably via a number of transport proteins including ABC transporters (Kunst and Samuels 2009), although again it is not clear how exactly the wax molecules are transported through the cell wall onto the cuticle. The current hypothesis is that Lipid Transfer Proteins (LTPs), which are small proteins with hydrophobic pockets that bind long-chain fatty acids, are responsible for the transport of the wax molecules onto the cuticle (Samuels et al. 2008). However, this has not been conclusively proven yet. Once upon the cuticle the wax compounds self-assemble into wax crystals(Koch and Ensikat 2008).

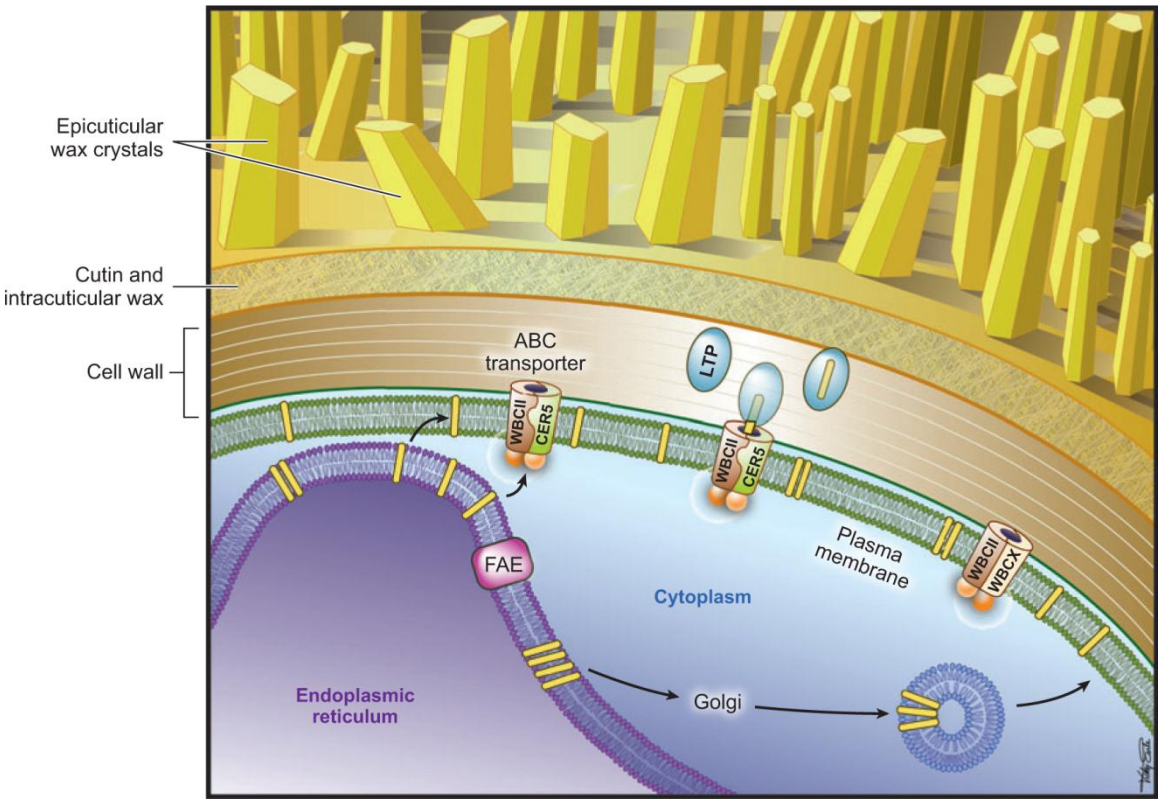


Figure 1.3: Transport of lipid compounds from the cell onto the cuticle

After synthesis in the endoplasmic reticulum the lipid compounds are transported to the plasma membrane. From there, they are thought to be exported from the cells via ABC transport proteins and further transported through the cell wall onto the cuticle surface by LTPs. Figure taken from (Samuels et al. 2008).

Analyses of leaf surfaces using electron microscopy have associated rod-like crystals with glaucousness and platelet shaped crystals with non-glaucousness (Hall et al. 1965, Barber and Netting 1968). This suggests that the rod-like tubes contain β-diketones, while the platelet shaped crystals probably are mainly made up of primary alcohols (Netting and Wettstein-Knowles 1973).

Glaucousness is thought to be an adaption to water-limited environments, hence many studies have tried to understand the relationship between glaucousness and drought stress or radiation intensity (Chatterton et al. 1975, Reicosky and Hanover 1978, Johnson et al. 1983, Richards 1984, Blum 1988, Febrero et al. 1998, Merah et al. 2000, Monneveux et al. 2004). Most of the above-mentioned studies on abiotic interactions agree that glaucousness increases the reflectance of solar radiation and that glaucous leaves are cooler than non-glaucous ones, but they also show that the transpiration efficiency (net photosynthesis to transpiration ratio) is lower in glaucous plants. An increased yield was reported in glaucous wheat, durum wheat and barley lines compared to their non-glaucous counterparts in a Mediterranean climate (Chatterton et al. 1975, Johnson et al. 1983, Richards et al. 1986, Merah et al. 2000, Monneveux et al. 2004).

All of these studies have been performed under Mediterranean conditions and for some of these studies the yield was evaluated in the glasshouse (Richards et al. 1986), gas exchange and cuticular conductance were unchanged or higher in non-glaucous lines (Johnson et al. 1983, Febrero et al. 1998, Merah et al. 2000) and glaucousness did not seem to have an effect in a well-watered or extremely dry environments (Johnson et al. 1983, Febrero et al. 1998). Also the relationship between glaucousness and cuticular wax load was inconsistent (Johnson et al. 1983, Febrero et al. 1998).

In contrast a study performed in the UK at several locations over consecutive years using a UK adapted doubled haploid (DH) population has shown an increase in yield (~4.15% on average) in non-glaucous lines over their glaucous counterparts, as well as a delay in senescence, thus extending the green canopy duration of the non-glaucous material by ~1.5 days (Simmonds et al. 2008). This suggests that non-glaucousness might be a beneficial trait in a UK environment, which has more rainfall and less sunshine than Mediterranean environments.

However, the molecular mechanism underlying the non-glaucous phenotype, what type(s) of compound(s) are being affected by *Iw1*, and its effects across developmental stages remain unknown. Based on previous studies (Hall et al. 1965, Barber and Netting 1968, Netting and Wettstein-Knowles 1973, Johnson et al. 1983), where non-glaucous plants were shown to have waxes on their surfaces we hypothesise that *Iw1* abolishes the rod-like tubes found on plant surfaces,

but that at least some other wax compounds remain largely unaffected. We further hypothesise that the non-glaucous lines of the DH population have an increased yield because they (i) reflect less light than their glaucous counterparts, thus using the “sparse” solar radiation to its full potential and (ii) senesce later and thus increase the grain filling period, allowing them to channel more nutrients to the grain. Alternatively, other factors (genes) linked to the non-glaucous trait affect yield positively either by themselves or in combination with the reflectance and senescence effects.

The present chapter aims to test the hypotheses regarding the effects of *Iw1* on epicuticular wax composition and its effects on physiological properties (cuticular conductance, reflectance, water use efficiency) as well as yield and grain-filling period.

1.1 Materials and Methods

1.1.1 Plant material

Shamrock is a non-glaucous hexaploid UK winter wheat variety that originates from a cross between a hexaploid breeding line (CWW 4899/25—Moulin x Monopol) and a tetraploid *Triticum turgidum* ssp. *dicoccoides* (TTD) accession (Comp Tig 323-1-3 M). It was shown that the non-glaucous trait originates from the TTD introgression (Simmonds et al. 2008). Shango is a hexaploid UK winter wheat variety that has a strong glaucous phenotype. A doubled-haploid (DH) population of the cross Shango x Shamrock (87 lines) has been described before (Simmonds et al. 2008). Here we focus on four of these DH lines: DH21 and DH119 are glaucous lines while DH81 and DH93 are non-glaucous lines. The non-glaucous phenotype of the DH population was mapped between SSR marker *Xgwm614* and DArT marker *wPt-4453* (Simmonds et al. 2008).

To study the effects of *Iw1* in UK adapted germplasm six glaucous hexaploid UK varieties (Alchemy, Einstein, Hereward, Malacca, Robigus, Xi19) were crossed to a non-glaucous Shango x Shamrock DH line (consists to ~70% of alleles from Shango). Plants heterozygous across the *Iw1* interval were selected at each generation and backcrossed to the respective recurrent parent. After the second backcross, plants were self-pollinated and homozygous BC_2F_2 near isogenic lines (NILs) were selected. Four streams of NILs were generated and used for field experiments, but only one stream from each variety (the streams did not differ in phenotype) was used for the bulk of experiments in this chapter (AS4, ES1, HS2a, MS1, RS1, XS1). The nomenclature for the NILs is based on their parents: AS4 is stream 4 derived from a cross between Alchemy x Shamrock-like (AS); ES1 is stream 1 derived from a cross between Einstein x Shamrock-like (ES), etc. The BC_2 material was bulked up in Hege-90 plots (H90; 1 m^2) in the field in 2009/2010, giving us enough seed to be used for Hege-80 yield plots (H80; 6 m^2) at Church Farm (Bawburgh, Norfolk) and at several locations of the breeding companies Limagrain and RAGT (Materials and Methods 1.1.15). Two further backcrosses were performed in parallel and the resulting lines were self-pollinated to yield homozygous BC_4F_2 NILs. These were bulked up in the field in H90 plots in 2011/2012 and used in H80 yield plots in 2012/2013.

1.1.2 Scanning electron microscopy (SEM)

Flag leaf blades and the exposed part of peduncles of Shango, Shamrock, and the six pairs of *Iw1* BC_2F_3 isogenic lines were sampled from the 2009-2010 field grown plants in July 2010 9 days after anthesis (visible wax deposition in glaucous recurrent parents). The collected plant material was quickly dissected by hand and samples were stuck to the aluminium SEM stub using O.C.T. compound (BDH Laboratory Supplies, Poole, England). The stub was then immediately plunged into liquid nitrogen slush at approximately -210°C to cryo-preserve the material. The sample was transferred,

onto the cryostage of a CT1500HF cryo-transfer system (Gatan, Oxford, England) attached to a Philips XL30 FEG scanning electron microscope (FEI, Eindhoven, The Netherlands). Sublimation of surface frost was performed at -95°C for three minutes before sputter coating the sample with platinum for 135 sec at 10 mA, at colder than -110°C. After sputter-coating, the sample was moved onto the cryo-stage in the main chamber of the microscope, held at approximately -140°C. The sample was viewed at 3 kV with the secondary electron detector and digital TIFF files were stored.

1.1.3 Freeze fractionation of SEM samples

Samples of two to three millimeter in height were fixed in horizontal orientation on an aluminum stub using O.C.T. compound (BDH Laboratory Supplies, Poole, England). The stub was then immediately plunged into liquid nitrogen slush at approximately -210°C to cryo-preserve the material. The sample was transferred onto the cryostage of a CT1500HF cryo-transfer system (Gatan, Oxford, England) attached to a Philips XL30 FEG scanning electron microscope (FEI, Eindhoven, The Netherlands). The sample stage was then heated to -100°C at which the tissue was hit with a dissecting knife integrated into the sample chamber, causing the sample to fracture. Sublimation of surface frost was performed at -95°C for three minutes before sputter coating the sample with platinum for 135 sec at 10 mA, at colder than -110°C. After sputter-coating, the sample was moved onto the cryo-stage in the main chamber of the microscope, held at approximately -140°C. The sample was viewed at 3 kV with the secondary electron detector and digital TIFF files were stored.

1.1.4 Isolation of waxes

To study the effect of *Iw1* on cuticular waxes, flag leaves and the exposed part of peduncles from five independent biological replications were collected for Shango, Shamrock, the four DH lines (DH21, DH81, DH93, DH119), and five BC_2 NIL pairs (Alchemy, Einstein, Hereward, Malacca, Robigus) from the 2009-2010 field plots. For the time course analyses, three flag leaves from independent replications were collected from four BC_2 NIL pairs (Alchemy, Malacca, Robigus, Xi19) grown in 2010-2011. Samples were collected in the field, placed in pre-weighed 15 mL polypropylene tubes and frozen on dry ice. Tubes were reweighed to determine the wet tissue weight before freezing in liquid nitrogen and storing at -80°C.

Waxes were extracted using 5 mL chloroform (Merck, analytical grade; as are all other solvents below) in glass tubes with screw-cap polytetrafluoroethylene lids containing chloroform and triacontane (C_{30} alkane, Sigma 263842, Poole, UK) as an internal standard (35 µg/mL for leaves and 10 µg/mL chloroform for peduncles); samples were immersed for 10 min at room temperature and shaken three times for 10 sec. The extracts were transferred to new glass tubes and dried down in a Vortex Evaporator (3-2201, Buchler Instruments Inc., NJ, USA). Each wax sample was re-suspended in

1 mL of chloroform and transferred to a pre-weighed Agilent glass vial, dried under nitrogen and then re-weighed to determine the total amount of wax extracted.

1.1.5 Thin layer chromatography (TLC)

Thin layer silica-glass Uniplates (Analtech Inc, Newark, USA; 21011) were run in a hexane-diethyl ether solvent (9:1 v/v) and waxes visualized by spraying the plate with 0.05% primuline (Sigma, 206865) in acetone-water (8:2 v/v) and imaging under UV light in a BioDoc-It™ transilluminator (UVP, Cambridge, UK). The edges of preparative plates were sprayed to reveal the locations of the different lipids and the unsprayed silica in the same relative positions scraped off, extracted in chloroform and processed for GC-MS analysis as described in Experimental Procedures. To visualize alkylresorcinols and methyl alkyresorcinols, TLC plates run in either hexane-diethyl ether (9:1 v/v) or chloroform-ethyl acetate (8.5:1.5 v/v) were sprayed with 0.05% aqueous Fast Blue B (Sigma, D9805), reacting with the hydroxyl groups at positions 1 and 3 of the benzol ring of the MARs (Kozubek and Tyman 1995).

1.1.6 Spectrophotometry

The maximum amount of β-diketone aliphatics in a wax sample was estimated from optical density measurements at 273 nm (von Wettstein-Knowles 1976) using a calibration curve made from purified barley β-diketones. The relative amounts of hentriacontane-14,16-dione (β-diketone) and 25-hydroxyhentriacontane-14,16-dione (OH-β-diketone) of the total β-diketone lipids were determined similarly after extracting them from preparative TLC plates (R_f circa 0.87 and 0.27, respectively).

1.1.7 Gas chromatography-mass spectrometry

Wax samples were derivatised in a 100 µL aliquot of a pyridine and TMS-BSTFA (Sigma 15238) mixture (1:1 v/v) at 75 °C for 1 h; samples were vortexed every 15 min. Commercial standards [C_{30} alkane (Sigma 263842), a mix of C_7-C_{40} *n*-alkanes (Sigma 49452), a mix of 100 µg each of 1-tetracosanol (Sigma L350), 1-hexacosanol (Sigma H2139), 1-octacosanol (Sigma O3379) and 1-triacontanol (Sigma T3777)] and samples of β-diketones (96% hentriacontane-14,16-dione) and hydroxy-β-diketones (97% 25-hydroxyhentriacontane-14,16-dione) isolated from barley (*Hordeum vulgare* L. cv. Bonus) spikes (von Wettstein-Knowles 1976) were derivatised similarly. The derivatised fraction was analysed on an Agilent GC 6890N gas chromatograph (Agilent Technologies, Wilmington, Delaware, USA) coupled to a 5973 Inert Mass Selective Detector. Automated splitless 3 µL injections were made using an Agilent 7683 automatic sampler. Conditions of chromatography were: inlet temperature 250°C, He carrier gas at a flow rate of 0.8 mL/min, nominal inlet pressure of 9.27 psi, the oven temperature program was: from 140°C (1 min) to 380°C (at 10°C/min), then held for 5min.

The column used was a ZB-5HT Inferno (Zebron; 7HG-G015-02, Phenomenex, Torrance, CA, USA) 30 m x 0.25 mm x 0.1 μm with a 5 m guard column fitted to the front end. The Retention Time Locking feature was used and the method locked to the retention time of the triacontane internal standard (16.3 min). The mass spectrometer parameters using electron ionisation in positive mode (70eV), with a source temperature of 230°C and a quad temperature of 150°C were set to the manufacturer's recommended defaults. Total ion scans were made from 50-500 amu; all data were processed via the Agilent GC Chemstation software (D.03.00) in conjunction with the NIST Mass Spectral Library, V8.0 (National Institute of Standards and Technology, Gaithersburg, Maryland, USA).

1.1.8 Quantification of wax compounds

Subtracting the percentages of the β-diketone aliphatics from the total wax gives the percentage attributable to the other components. Relative abundances for these compounds were calculated from GC-MS TIC peaks by automatic integration using the Custom Report function in ChemStation. Where compounds such as the C_{26} FA/C_{28} aldehyde and C_{28} FA/C_{30} aldehyde eluted closely together, so that individual total ion chromatogram (TIC) peaks could not be integrated separately, the major ion for each compound was searched and integrated separately for the relevant retention time. The derivatised β-diketone peak was integrated manually and then the characteristic ion for the obscured C_{30} FA peak was integrated and subtracted from the β-diketone peak. The same approach was used to estimate the C_{23}MAR peak hidden within the hydroxy-β-diketone peak. While the data presented do not take into account that not all wax aliphatics are silylated nor the differential responses of the chemical groups to flame ionization (Sternberg et al. 1962), they give a reproducible approximation of the quantities of the wax aliphatics. All together 53 components were identified, with 26 being studied in more detail as they account for >95% of the total wax load in both Shango and Shamrock flag leaves and peduncles.

1.1.9 Chemical modification of waxes

Some wheat wax samples dissolved in hexane were passed through columns containing inert silica ± NaOH powder extracted with chloroform (Von Wettstein-Knowles 1974). The eluates lack free fatty acids which are retained on the column. Control and treated eluates, together with wax standards, were then assessed by TLC and GC-MS.

To further assist in identifying some GC-MS peaks, selected wax samples were treated either with $NaBH_4$ to reduce the carbonyl group of ketones and aldehydes to secondary and primary alcohols, respectively, or with $K_2Cr_2O_7$ to oxidize the hydroxyl group of secondary alcohols to the carbonyl group of ketones (von Wettstein-Knowles 1976, Mikkelsen and von Wettstein-Knowles 1978).

1.1.10 Transmission electron microscopy

Field grown (2009-2010) Shamrock and Shango (264 days old; 3 biological replicates) and glasshouse grown Xi19 BC_2 NILs (3 biological replicates) were dissected by hand. Plant material was fixed in 2.5% (v/v) glutaraldehyde, 0.05 M sodium cacodylate, pH 7.3, and embedded in LR White resin (London Resin) (Lodwig et al. 2005). The embedded material was cut into ultra-thin sections of approximately 90 nm using a diamond knife on a Leica UC6 ultramicrotome (Leica Microsystems; Milton Keynes, UK) and placed on 200 mesh copper grids which had been pyroxylin- and carbon-coated. The sections were contrast-stained with 2% (w/v) uranyl acetate and 1% (w/v) lead citrate before imaging in a FEI Tecnai 20 TEM (FEI, Eindhoven, Netherlands) at 200 kV. Digital TIFF images were acquired on the microscope using a Deben AMT XR60B CCD camera (Deben, Bury St Edmunds, UK). Cuticle thickness was calculated from 20-40 pictures taken for each biological sample and the average cuticle thickness was calculated for statistical analyses.

1.1.11 Analysis of cuticular conductance

Flag leaf tissue of six pairs of field-grown BC_2F_3 NILs (15 replicates each) from two different UK locations (Woolpit and Docking) was collected and imbibed in water in the dark over night at 4°C. The tissue was weighed on a balance (Sartorius analytical balance BP61S) and then dried in the dark at 25°C and ~50% relative humidity. Tissue weight was recorded in intervals of 20 min for 120 min total, thus obtaining seven data points. The leaves were then dried at 70°C for 24 h and the dry weight of the leaf was subtracted from the individual measurements.

1.1.12 Bulk $\delta^{13}C$ measurements and calculation of ^{13}C discrimination

Carbon isotope discrimination was determined for tissues from field grown plants in 2011 and 2012. Flag leaves and spikes were collected at anthesis and grains at maturity. Freeze-dried samples were milled to powder using mortar and pestle. Approximately 0.5 mg of powder was measured for $\delta^{13}C$ using a Thermo Finnigan Deltaplus XP isotope ratio mass spectrometer interfaced to a Costech Elemental Combustion System CHNS-O 4010. Isodat Version 2.0 was used for data processing of $\delta^{13}C$ measurements. Carbon isotope composition is expressed relative to Vienna PeeDee belemnite Standard (VPDB) based on an in-house reference gas. The accuracy of $\delta^{13}C$ measurements was ±0.2‰ or better based on an in-house collagen standard. ^{13}C discrimination was calculated using the following formula $\Delta = (\delta a - \delta p)/(1 + \delta p)$ (Farquhar et al. 1989) and expressed in per mL on the VPDB scale, where ambient CO_2 value δa, was assumed to be -8‰. δp represents the $\delta^{13}C$ values for flag leaf, spike, and grain samples measured using the above methodology.

1.1.13 PAR reflectance measurements

PAR reflectance was measured in the field in consecutive years (2010-2011 and 2011-2012). The sensor system consists of two PAR Quantum sensors (SKP215, Skye Instruments) attached to a pole (SKL 910, Skye Instruments) with one hemispherical sensor facing upwards registering all incoming light and a second sensor with a collar, which limits the sensor's field of view to 25°, facing downwards collecting reflected light from a ~50 cm^2 piece of a H80 plot. The sensors were connected to a Spectrosense+ data logger (SKP 215LQ/SS2, Skye Instruments) and a circular level bubble ensured that the sensors were level over the plot. Approximately 30 measurements were taken at three to four different positions of each plot and combined into one average value. For each genotype five plots were measured. The ratio of incoming to reflected light remained constant independent of light intensity, cloud cover or time of day. From this ratio the difference in reflectance (in percent) was calculated.

1.1.14 Integrating sphere measurements

Field-grown Alchemy and Malacca BC_2F_3 NILs (five replicates each) were dug up from Church Farm field in 2012, transferred into 2 L pots and left to acclimatise in the lab for one week, before transporting them by car to the University of Colchester, where their flag leaf blades were analysed using an integrating sphere. Transmission and reflectance between 300 and 800 nm was measured for both the abaxial and adaxial side of the flag leaf blades and normalized using blank measurements (no leaf tissue in the path of light).

The reflectance values across each individual time point were quite variable, thus making single point analysis difficult. To compensate this effect, trendlines were added to the graphs to smoothen the reflectance values: A single point of the trendline consists of the sliding window average of 15 datapoints, e.g. from 400 – 414 nm. The next point of the trendline consists of datapoints from 401 – 415 nm and so on. This results in a much smoother and much more informative graph than before.

1.1.15 Measurements of green-canopy duration and yield

The grain filling period is the time after flowering when plants channel photo-assimilates into their grain. Since grain filling period is not easily measured for large amounts of samples, the green canopy duration, the time between flowering and full plant senescence, is generally used as a surrogate measurement. Flowering date was assessed as the day in which ¾ of the plot had reached ¾ of inflorescences emerged, and full plant senescence as the day where a loss of chlorophyll was observed on all leaves and peduncles of the main shoots for ¾ of the plot. The green canopy duration was then calculated for each plot by subtracting the number of days from drilling to flowering from

the number of days to full plant senescence. Yield was measured as weight of grains per plot and then normalized to 15% moisture content (plot size was verified to be uniform across the field). The yield of glaucous lines was then used as a baseline (100%), which the yield of non-glaucous lines could be compared to. A value below 100% represents a reduction in yield in the non-glaucous line, while values above 100% represent an increased yield in the non-glaucous line. At Church Farm all streams for every BC_2 NIL pair were used, while only one stream of each NIL pair was sent to the breeders. In 2010 and 2011, these were the same streams that were used for most of the biochemical and physiological experiments: AS4, ES1, HS2a, MS1, RS1 and XS1 (see Materials and Methods 1.1.1). Two streams were generated for the BC_4 NILs and both of them were used for the field trials at JIC and at the breeder's trial sites. The UK locations at which the field trials were performed are: Church Farm (Bawburgh, Norfolk, NR9), Docking (Norfolk, PE31), Drinkstone (Suffolk, IP30), Dukes (Cambridgeshire, CB10), Ickleton (Cambridgeshire, CB10), Pampisford (Cambridgeshire, CB2), Whittlesford (Cambridgeshire, CB22), Wolferton (Norfolk, PE31) and Woolpit (Suffolk, IP30).

1.1.16 Statistical analysis of results

We have used Genstat version 15.2.0.8821 (VSN International Ltd) throughout for the statistical analysis of our data. We have performed analysis of variance (ANOVA) for multiple mean comparisons.

1.2 Results

1.2.1 Scanning electron microscopy analysis of glaucous and non-glaucous material

As was mentioned before, a glaucous phenotype is caused by the deposition of wax crystals onto the cuticle surface of aerial plant organs. This so-called epicuticular wax layer scatters incoming light, thus causing the plant surface to appear whitish or bluish, rather than bright green. To analyse the structure of this epicuticular wax layer, field-grown tissue samples from both glaucous and non-glaucous lines were collected and examined using cryogenic scanning electron microscopy (cryo-SEM). Since the tissue to be analysed is frozen in a slush of liquid nitrogen, the tissue surface is perfectly conserved when using cryo-SEM, thus allowing for a detailed and accurate analysis.

Three biological replicates of flag leaf blades and the exposed part of peduncles were used for each genotype and tissue. Samples from both glaucous and non-glaucous lines were always loaded on the same sample stage and thus underwent the same pre-treatment (see Materials and Methods 1.1.2).

1.2.1.1 SEM of the abaxial side of the flag leaf

The abaxial side of the flag leaf, which is the outer side of the flag leaf that faces away from the stem, showed a dense network of tubular shaped wax crystals covering the entire surface of the flag leaf in Shango sections (Figure 1.4A, B). Stomata were sometimes barely visible underneath this thick layer (see arrow in Figure 1.4A). In stark contrast, the abaxial flag leaf surface of Shamrock completely lacks these tubular shaped wax crystals (Figure 1.4C, D). The surface is not completely bare though, as other types of wax crystals are visible. These compounds are probably also present on the Shango leaf samples, but not visible due to the dense layer of tubular shaped wax crystals covering the leaf surface.

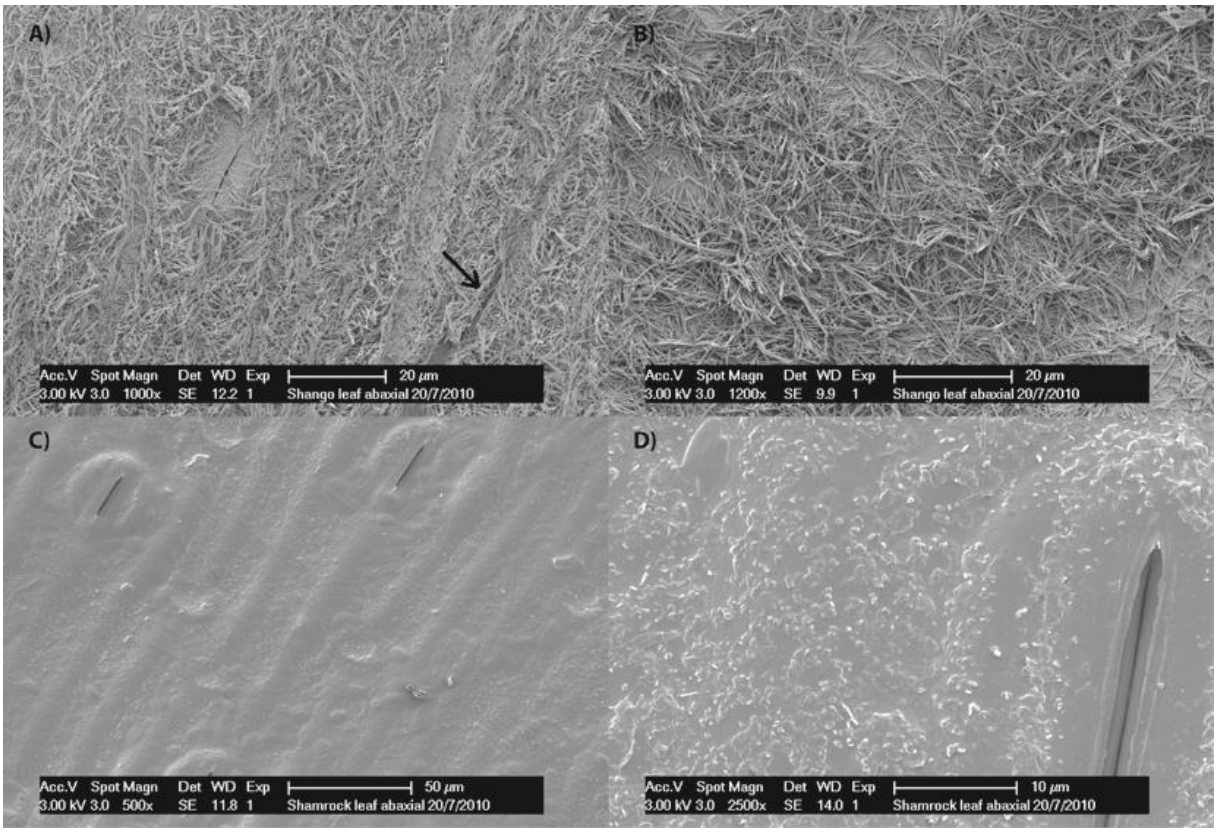


Figure 1.4: Epicuticular wax pattern on abaxial side of Shango and Shamrock flag leaves
Images from flag leaf blades of two different plants from both Shango and Shamrock are shown. The abaxial side of the flag leaf in Shango is completely covered with a tubular shaped wax (A, B), while the corresponding surface in Shamrock is almost bare, with only low amounts of other wax crystals visible (C, D). The arrow in (A) points shows a stomata that is almost completely covered by wax.

1.2.1.2 SEM of the adaxial side of the flag leaf

The adaxial side of the flag leaf, which is the inner side of the flag leaf that faces towards the stem, revealed a more complex pattern of wax crystals. Unlike the abaxial side, a pattern of light and dark strips is visible in leaf sections of both Shamrock and Shango when using a low magnification level (Figure 1.5A, B). The dark strips seem to be separated from the light strips by rows of stomata and leaf hairs. Zooming in unto the dark strips reveals that their surface is densely covered by a platelet shaped wax crystal in both Shamrock and Shango (Figure 1.5C, D). These platelet shaped crystals are also present on the surface of the light strips in both varieties, but at a much lower density than on the dark strips (Figure 1.5E, F). Similarly to the abaxial surface of the flag leaf, Shango samples showed presence of a tubular shaped wax in addition to the other type of wax crystal (Figure 1.5D, F). However the density of the tubular shaped wax is markedly reduced on the adaxial surface compared to the abaxial side.

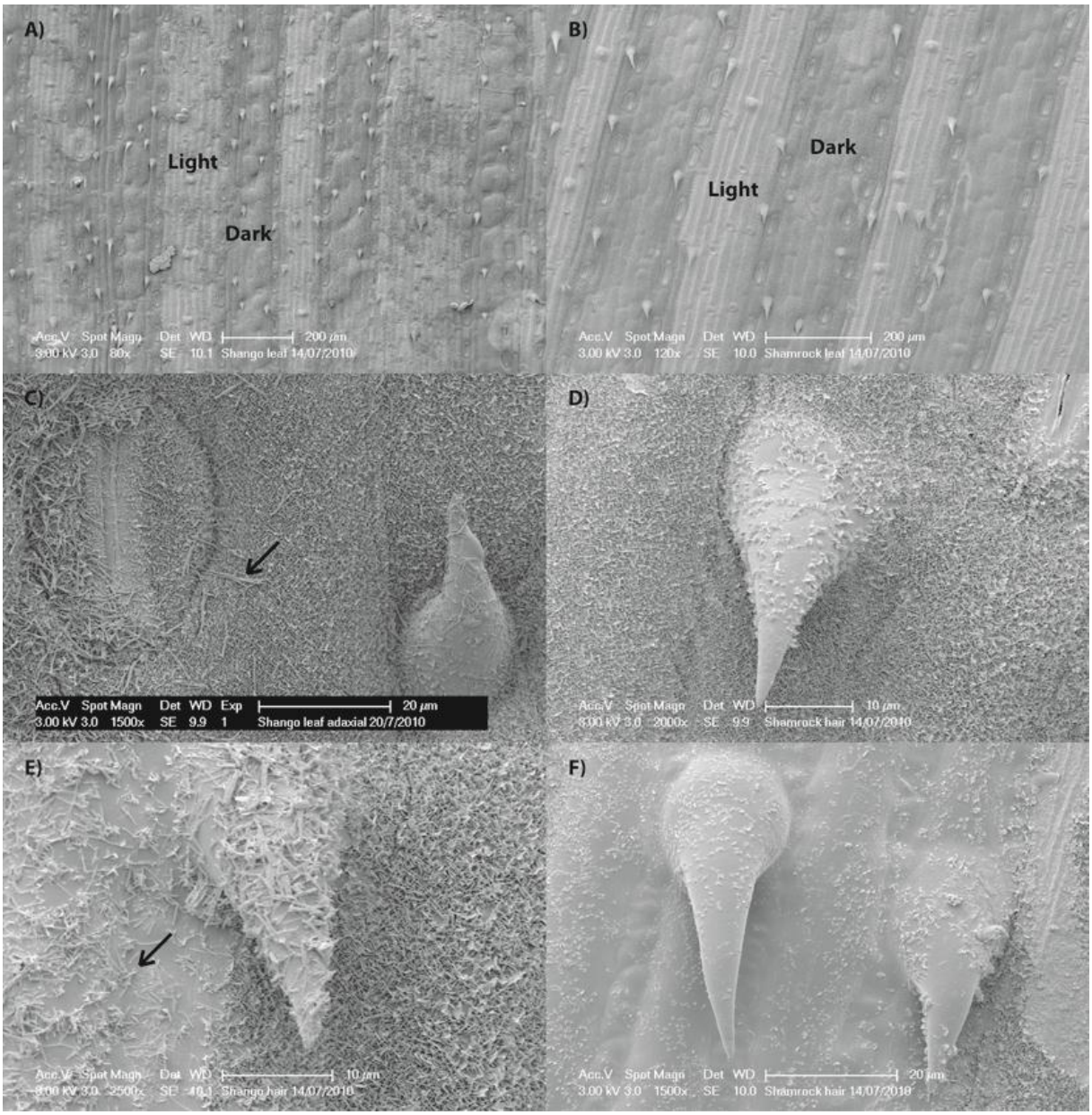


Figure 1.5: Epicuticular wax pattern on adaxial side of Shango and Shamrock flag leaves

The adaxial side of flag leaves shows a pattern of light and dark strips in both Shango (A) and Shamrock (B). A close-up of the dark strips reveals a high abundance of platelet shaped wax crystals in both Shango (C) and Shamrock (D), whereas this type of wax crystal is much rarer in the light strips (Shango E, Shamrock F). In addition, both light and dark strips in Shango show presence of the tubular shaped wax (C, E), which is completely absent from Shamrock (D, F). Arrows point out tubular shaped wax.

1.2.1.3 SEM of the surface of peduncles

Sections of the exposed part of peduncles from both Shamrock and Shango look very similar to abaxial flag leaf sections, Shango samples having a high density network of the tubular shaped wax, which covers the entire peduncle (Figure 1.6A, B). In contrast, Shamrock samples appear bare, with only small amounts of the platelet shaped wax visible (see Figure 1.6C, D). This platelet shaped wax is

also present on peduncle sections of Shango, but in a much lower abundance than the tubular shaped wax.

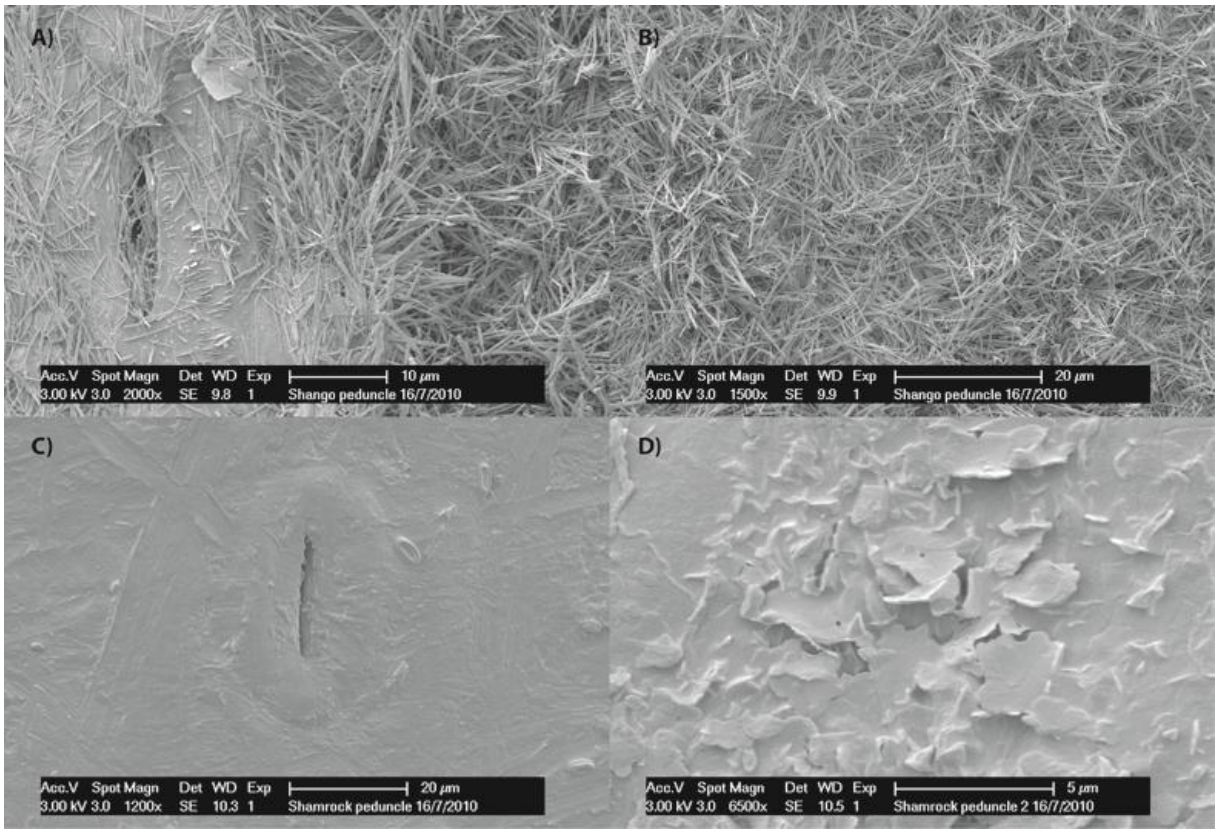


Figure 1.6: Epicuticular wax pattern on Shango and Shamrock peduncles

Images from the exposed part of peduncles of two different plants from both Shango and Shamrock are shown. The glaucous Shango peduncle is covered by a thick layer of the tubular shaped wax (A, B), while the non-glaucous Shamrock peduncle is completely devoid of this type of wax crystals (C, D). Both varieties have the platelet shaped wax crystals present on the peduncle surface (A, D).

1.2.1.4 Freeze fractionation of flag leaf blades

To get a qualitative overview of the cuticle and the epicuticular waxes, flag leaf blades of Shango and Shamrock were freeze fractionated (Materials and Methods 1.1.3). This essentially produces a cross-section of the tissue that can be analysed using the cryo-SEM. Similar to the standard cryo-SEM analysis Shango samples contained a tubular shaped wax on the surface of the cuticle (Figure 1.7A, B), which was absent from Shamrock samples (see Figure 1.7C, D). From the images obtained it appears that the wax crystals form a single relatively thin layer on top of the cuticle. However, this method proved to be too difficult and inaccurate an approach to accurately quantify the thickness of the epicuticular wax layer.

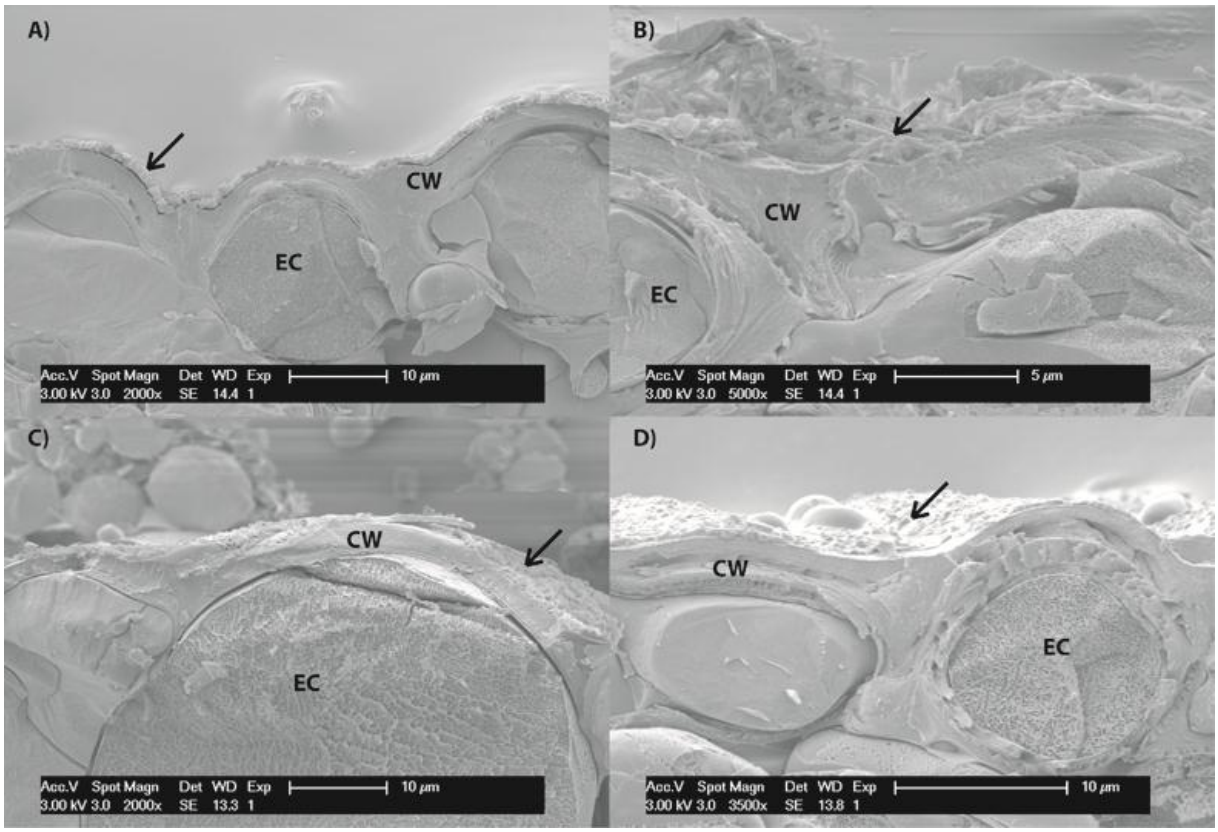


Figure 1.7: Freeze Fractionation of Shango and Shamrock Flag Leaf Blades

Images from flag leaf blades of two different plants from both Shango and Shamrock are shown. Cross section through flag leaf blades of Shango (A, B) and Shamrock (C, D) made using freeze fractionation. Epidermal cells (EC) and cell walls (CW) are clearly visible, as are epicuticular wax crystals (arrows).

In addition to Shango and Shamrock, flag leaves and the exposed part of peduncles of the six pairs of BC_2F_2 near isogenic lines (NILs) and their parents were also analysed using cryo-SEM. All NILs carrying the *Iw1* introgression mimicked the Shamrock phenotype and lacked the tubular shaped wax on all surfaces, while all NILs without the introgression had the tubular wax crystals present, similar to Shango. However, the density of these tubular wax crystals varied perceptibly between the different varieties (Appendix, A1.1-A1.6).

The results from this cryo-SEM analysis show a very consistent pattern regarding the function of the introgressed segment containing *Iw1*. In its absence, all three plant organ surfaces were covered with a tubular shaped wax to varying degrees in addition to other wax crystals. However, upon introduction of this physical segment, the tubular shaped wax crystals completely disappeared, while other types of wax crystals remained present. This suggests that the function of the introgressed segment containing *Iw1* is to alter the biosynthesis or transport to the surface of the tubular shaped wax crystals.

1.2.2 Biochemical analysis of cuticular wax compounds in glaucous and non-glaucous wheat

The results in this section (1.2.2) were performed in close collaboration with Prof. Penny von Wettstein-Knowles from the University of Copenhagen. Prof. von Wettstein-Knowles has been working on wax composition in barley for several decades and was paramount in the following work by supplying both chemical standards and her extensive knowledge. The biochemical experiments described in this section (1.2.2) were performed by Dr. Sarah G. Mugford and Dr. Max Bush. My role in these experiments was to supply tissue for analysis, helping with the experimental design and the ultimate analysis of the results. This analysis is highly relevant in order to understand the big picture of *Iw1* and non-glaucousness and the results were published in The Plant Journal (Adamski et al. 2013) and some figures and tables from this publication are used here.

1.2.2.1 TLC and spectrophotometric analysis of cuticular waxes

The results from section 1.2.1 showed that a tubular shaped type of wax is present in all glaucous samples and absent from all non-glaucous samples, suggesting that this type of wax is the causal agent that scatters incoming light and thus creates a glaucous phenotype.

The tubular shaped wax has been reported in previous studies on barley, wheat, *Poa colensoi* (blue tussock) and several *Eucalyptus* spp. to correlate with the amount of β-diketone aliphatics, suggesting that these compounds cause glaucousness (Netting and Wettstein-Knowles 1973). Indeed, a thin-layer chromatography (TLC) analysis of wax extracts from Shango and Shamrock flag leaf blades clearly shows an absence of β-diketones from the Shamrock lane, further strengthening the correlation between tubular wax crystals and β-diketones (Materials and Methods 1.1.4 and 1.1.5; Figure 1.8). In addition, one band in the Shango extract (band 6, labeled “unknown”) was absent from Shamrock as well and it also did not correspond with the hydroxy-β-diketone band from the eucalyptus standard (band 5), as was expected. A mass-spectrometric analysis of this band did not yield any diagnostic ions, thus leaving the identity of the compounds in this band unknown (Figure 1.8).

The total and relative amount of β-diketones and their hydroxy derivatives was determined by OD_{273} measurements on flag leaf blade extracts from Shango and Shamrock (Materials and Methods 1.1.6; Table 1.1). Shango had more than twice as much total wax in the extracts than Shamrock, most of which (67%) consisted of β- and hydroxy-β-diketones, while in Shamrock only 7.5% of waxes consisted of β-diketone aliphatics. In both varieties the hydroxy-β-diketones constitute only a small part of the total β-diketone aliphatics (0.5 and 2% respectively). Both genotypes also had a number of compounds in common that have been previously reported to be part of cuticular waxes in

different wheat varieties, for example n-alkanes, primary alcohols, aldehydes and free fatty acids (Netting and Wettstein-Knowles 1973, Tulloch and Hoffman 1973, Koch et al. 2006).

Table 1.1: **Spectrophotometric OD_{273} measurements of Shango and Shamrock wax extracts**
The relative abundance of β-diketones in Shango and Shamrock flag leaf epicuticular waxes from field grown plants 20 DPA in 2011 was assessed using spectrophotometric OD_{273} measurements. The values are the average of two biological replicates; Dik = β-diketones plus hydroxy-β-diketones; β = β-diketones; OH-β = hydroxy-β-diketones.

Variety	Wax [mg[§]]	Leaf [mg[¶]]	Total Dik [mg]	% of Dik[†]		µg Dik/mg leaf	
				β	OH-β	β	OH-β
Shango	6.39	959	4.28	99.5	0.54	4.67	0.03
Shamrock	2.81	919	0.21	97.9	2.12	0.23	tr

[§] Dried weight of the leaf extract.
[†] Estimations based on OD_{273} measurements.
[¶] Fresh wet weight
tr = < 0.01µg/mg leaf

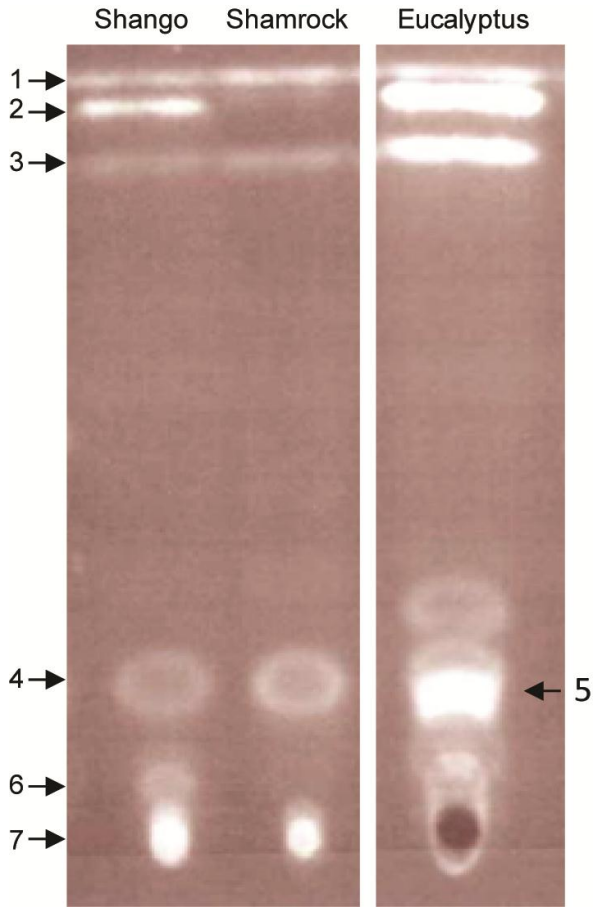


Figure 1.8: Thin layer chromatography (TLC) of Shango, Shamrock and eucalyptus wax extracts
Wax lipids in flag leaf blades of Shango (*iw1*), Shamrock (*Iw1*), and eucalyptus visualized after thin layer chromatography in hexane:ether (9:1 v/v) using primuline. The numbered arrows correspond to n-alkanes (1); β-diketone aliphatics (2); aldehydes (3); primary alcohols (4); hydroxy-β-diketones (5); unknown (6); alkylresorcinols, methylalkylresorcinols and free fatty acids (7). All samples were run on the same TLC plate.

1.2.2.2 GC-MS analysis of flag leaf blades and peduncles

The accumulation of results suggests that the presence of β-diketone aliphatics is responsible for a glaucous phenotype and that the amount of these compounds is at least strongly reduced, if not even completely abolished in an *Iw1* background. To systematically analyse the cuticular wax of Shango and Shamrock flag leaf blades and the exposed part of peduncles, an extensive gas chromatography/mass spectrometry (GC-MS) analysis was performed (Materials and Methods 1.1.7).

Gas chromatography in combination with mass spectrometry is a cost-effective way to analyse large numbers of complex samples with high efficiency. The high reproducibility of fragmentation patterns in electron impact (EI) ionization mass spectra combined with readily available mass spectra libraries make GC-MS an accurate method of identifying chemical compounds. One limitation to the approach is the restriction to volatile analytes, which requires chemical derivatization of metabolites, like

silylation, to increase their volatility, stability and peak symmetry (Orata 2012). In the resulting total ion chromatogram (TIC) traces, non-silylated aliphatics, for example n-alkanes and aldehydes, are less prominent than silylated primary alcohols (POHs), and silylated β-diketone aliphatics are underestimated versus silylated POHs (Tulloch and Hogge 1978). Peaks of individual compounds were resolved straightforwardly, apart from three non-symmetrical peaks (elution time between 18.5 and 20 min) present in Shango, but not Shamrock, wax samples (Figure 1.9). Silylation and subsequent GC-MS analysis of the C_{31} standards hentriacontane-14,16-dione (a β-diketone) and 25-hydroxy-14,16-dione (an hydroxy-β-diketone) resulted in similarly shaped peaks eluting with the same retention time as the three peaks in Shango samples. The mass spectra from the silylated standards suggest that the first peak consists of non-derivatized β-diketone (hentriacontane-14,16-dione), that the second peak consists of silylated isomers of hentriacontane-14,16-dione and that the third peak consists of silylated isomers of 25-hydroxy-14,16-dione (Tulloch and Hogge 1978). As these three peaks are not present in Shamrock TIC traces, it can be suggested that β- and hydroxy-β-diketones are absent from Shamrock waxes. This hypothesis is further supported by the fact that trace components in cuticular waxes absorbing at 273 nm will result in an overestimation of the amount of the β-diketone aliphatics (Table 1.1)(von Wettstein-Knowles 1976) and that β-diketone aliphatics were absent from Shamrock samples on TLC plates. Combining these results with the SEM analysis and previous reports on glaucousness (Netting and Wettstein-Knowles 1973), it is evident that the causal agent of glaucousness in Shango is a tubular shaped wax formed by β-diketone aliphatics.

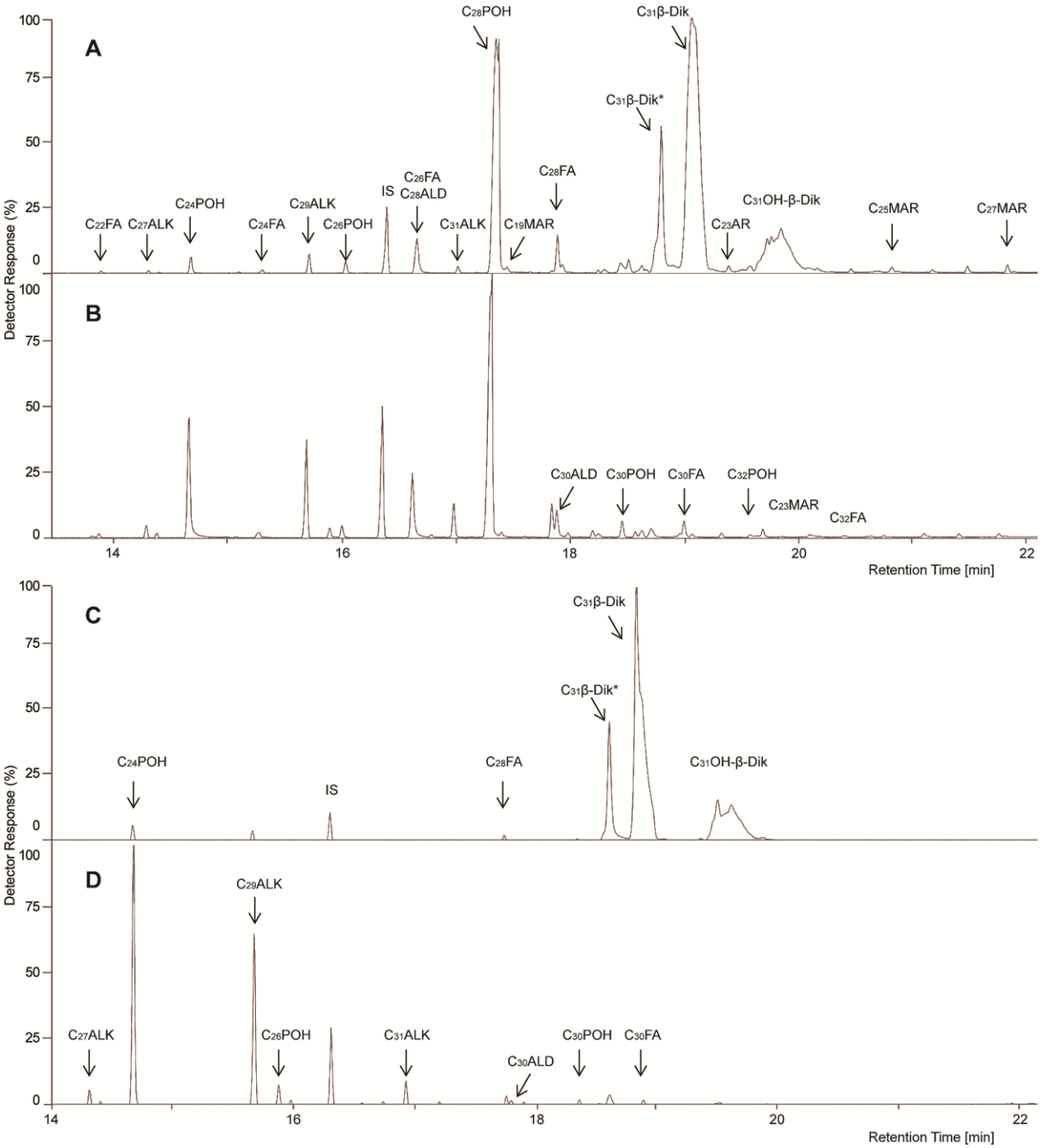


Figure 1.9: GC-MS analysis of Shango and Shamrock flag leaf blades and peduncles

Gas chromatography-mass spectrometry total ion chromatograms of Shango (A and C; *iw1*) and Shamrock (B and D; *Iw1*) flag leaf blades (A-B) and the exposed part of peduncles (C-D). Vertical axis is relative abundance; IS, internal standard (C_{30} ALK); * indicates non silylated. Fatty acids (FA), *n*-alkanes (ALK), primary alcohols (POH), aldehydes (ALD), methylalkylresorcinol (MAR), and both β- and hydroxy-β-diketones (β-Dik and OH-β-Dik) are indicated.

Furthermore, the GC-MS analysis identified C_{28}POH as a major constituent of cuticular waxes from flag leaf blades, but not of peduncles, where it is only present in trace amounts (Figure 1.9). Again, this agrees with the SEM analysis, the results of the TLC plates and previous work (Baker et al. 1982). Other highly abundant compounds on both flag leaf blades and peduncles are C_{24}POH, C_{29} and C_{31} *n*-alkanes. All in all, 26 compounds were quantified and a further 27 compounds were detected by this

GC-MS analysis, including n-alkanes from C_{22}-C_{33} and POHs from C_{18}-C_{34} chain lengths as well as less abundant C_{22}-C_{32} free fatty acids and C_{24}-C_{34} aldehydes (Materials and Methods 1.1.8 and 1.1.9; Appendix, A2).

In addition, field-grown flag leaf blades and the exposed part of peduncles of a doubled-haploid population derived from a Shango x Shamrock cross as well as five BC_2F_3 NILs were analysed using GC-MS. The wax load on Shamrock flag leaves was 38% of that on Shango, whereas in peduncles it was only 13% and similar results were obtained for the DH lines. The cuticular waxes of the *Iw1* DH lines contained only low amounts of β-diketone aliphatics, similar to Shamrock (Table 1.2), whereas waxes from the *iw1* DH lines contained high amounts of β-diketone aliphatics. In contrast, the amount of C_{28}POH was not significantly different between flag leaf blades of all lines. However, a significant increase in the amounts of *n*-alkanes and C_{24}POHs was observed in all lines carrying *Iw1* ($p<0.001$). The same patterns were obtained for the BC_2F_3 NILs independent of the amount of glaucousness of the original parent (Table 1.2). Wax samples from *Iw1* containing lines had low or trace amounts of β-diketone aliphatics and increased amounts of *n*-alkanes and C_{24}POHs. Interestingly, there is a lot of variation in the amount of individual wax components in the different varieties. Shango for example has ten times more β-diketones on its flag leaf blades and ~14 times more on its peduncles than Malacca, suggesting that even low amounts of β-diketone aliphatics will result in a glaucous phenotype (Table 1.2). These stark differences in the amount of β-diketone aliphatics can be recognized also by the naked eye, with Shango plants appearing much more glaucous than Malacca *iw1* NILs (MS1-, Figure 1.10).

These results suggest that β-diketone aliphatics (β- and OH-β-diketones) are the causal agent of glaucousness in wheat and that *Iw1* inhibits the production of these compounds.

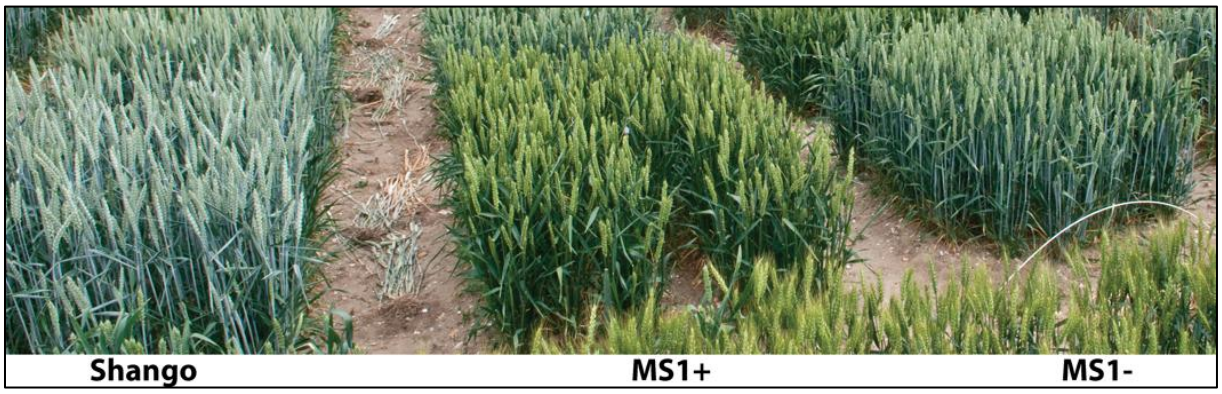


Figure 1.10: Field-grown Shango and Malacca NIL pair in 2010

Field-grown plots of Shango and a NIL pair of Malacca. While Shango and MS1- are both glaucous in comparison to MS1+, Shango is visibly more glaucous than MS1-, possibly reflecting the large difference in β-diketone aliphatics between the two genotypes.

Table 1.2: **Total wax load on Shango, Shamrock, DH and BC_2F_3 lines**

Total wax load (µg wax/mg leaf) on field-grown flag leaf blades and the exposed part of peduncles plus amounts (µg/mg) of the six most important constituents in the parental Shango and Shamrock lines, four DH lines and four BC_2F_3 lines nine days post anthesis in 2010. The sem is indicated for total wax load as well as β- and Hydroxy-β-diketones. tr = < 0.01; - = not detected

Wheat line	*Iw1* +/-	Wax load [µg wax/mg leaf]	*n*-Alkanes		Primary alcohols		β-diketone	Hydroxy-β-diketone
			C_{29}	C_{31}	C_{24}	C_{28}		
Flag leaf								
Shango	-	12.68 ± 1.35	0.1	0.04	0.12	3.29	7.09 ± 0.93	1.59 ± 0.17
Shamrock	+	4.76 ± 0.65	0.33	0.13	0.55	3.23	0.06 ± 0.01	-
DH119	-	10.81 ± 1.27	0.11	0.05	0.06	3.56	5.1 ± 0.85	1.46 ± 0.14
DH93	+	5.32 ± 0.57	0.38	0.13	0.78	3.45	0.06 ± 0.01	-
DH21	-	13.06 ± 1.27	0.11	0.05	0.11	2.85	7.86 ± 0.77	1.67 ± 0.26
DH81	+	4.80 ± 0.47	0.34	0.14	0.41	3.38	0.01 ± 0.001	-
Alchemy	-	7.67 ± 1.59	0.12	0.08	0.13	3.29	2.68 ± 0.59	0.83 ± 0.23
Alchemy	+	4.43 ± 0.57	0.2	0.26	0.39	3.09	-	-
Malacca	-	5.82 ± 1	0.1	0.04	0.15	3.99	0.71 ± 0.16	0.35 ± 0.11
Malacca	+	4.80 ± 1.04	0.21	0.09	0.36	3.68	-	-
Robigus	-	5.64 ± 0.39	0.08	0.04	0.07	3.04	1.15 ± 0.14	0.78 ± 0.3
Robigus	+	4.42 ± 0.84	0.27	0.09	0.23	3.31	-	-
Einstein	-	6.97 ± 1.05	0.09	0.13	0.17	3.57	2.12 ± 0.61	0.42 ± 0.09
Einstein	+	6.23 ± 0.64	0.21	0.34	0.59	4.46	-	-
Hereward	-	6.48 ± 0.76	0.09	0.03	0.19	3.53	1.76 ± 0.36	0.50 ± 0.12
Hereward	+	5.50 ± 0.47	0.34	0.11	0.72	3.75	-	-
Exposed part of peduncle								
Shango	-	6.03 ± 0.38	0.04	0.01	0.07	-	4.54 ± 0.31	1.34 ± 0.06
Shamrock	+	0.81 ± 0.08	0.27	0.05	0.37	-	0.04 ±0.003	0.03 ± 0.002
DH119	-	3.42 ± 0.33	0.07	0.01	0.04	-	2.32 ± 0.29	0.96 ± 0.06
DH93	+	1.24 ± 0.07	0.46	0.05	0.61	tr	0.03 ± 0.001	0.01 ± 0.005
DH21	-	4.26 ± 0.38	0.04	0.01	0.05	-	3.13 ± 0.25	1.01 ± 0.13
DH81	+	0.82 ± 0.11	0.35	0.04	0.38	tr	tr	-
Alchemy	-	1.82 ± 0.35	0.02	0.01	0.03	-	1.23 ± 0.25	0.53 ± 0.09
Alchemy	+	0.27 ± 0.05	0.08	0.06	0.1	-	-	-
Malacca	-	0.91 ± 0.28	0.06	0.01	0.11	-	0.33 ± 0.13	0.4 ± 0.17
Malacca	+	0.44 ± 0.06	0.16	0.03	0.18	tr	-	-
Robigus	-	1.29 ± 0.28	0.01	0.01	0.02	-	0.66 ± 0.14	0.58 ± 0.16
Robigus	+	0.36 ± 0.05	0.17	0.03	0.11	-	-	-
Einstein	-	1.47 ± 0.35	0.04	0.03	0.06	-	0.88 ± 0.26	0.43 ± 0.11
Einstein	+	0.47 ± 0.04	0.12	0.13	0.18	-	-	-
Hereward	-	0.85 ± 0.12	0.03	0.01	0.07	-	0.42 ± 0.08	0.26 ± 0.07
Hereward	+	0.48 ± 0.05	0.15	0.03	0.21	-	tr	tr

1.2.2.3 The effect of *Iw1* throughout plant development

The previous paragraphs have shown the effect of *Iw1* on the composition of cuticular waxes at nine days post anthesis, when glaucousness is clearly visible on all plant organs. But what are the effects of this locus at different points in time during a plant's development? To answer this question, a time course experiment was designed, where leaf samples of four pairs of field-grown BC_2F_3 NILs were collected at discrete points in plant development. These samples were analysed using the previously established GC-MS platform. Wheat developmental stages were assessed using two digit decimal Zadoks scale (Zadoks et al. 1974). The time points chosen were GS31, GS47, GS51, 18 days post anthesis (DPA) and 42 DPA (Figure 1.11). At the early time point GS31 no wax is visible on the surface of a wheat plant, whereas at GS47, the flag leaf and flag leaf sheath are displaying wax, while the ear has not yet emerged. The later time points were chosen to elucidate whether cuticular waxes continue to accumulate over time or not.

At the earliest time point (GS31) the cuticular wax of all samples consisted solely of primary alcohols, specifically of chain length C_{26}, C_{28} and C_{30}. Of these compounds C_{28}POH was overwhelmingly dominant, with its abundance being 100 to 400-fold higher than either of the two other primary alcohols in the different lines. At the second time point (GS47) the complexity of the cuticular wax increased markedly, with *n*-alkanes, aldehydes and fatty acids appearing alongside the primary alcohols. In addition, glaucous lines also contained β-diketone aliphatics, which were absent from non-glaucous lines. Interestingly, although C_{28}POH was still the most abundant constituent of cuticular wax at this stage, the total amount of it had dropped to 2 µg/mg leaf. This is most likely due to the differences in leaf thickness, and hence weight, between early vegetative leaves and the flag leaf that was sampled at boot stage. During the next developmental stages, the general trend for all wax components was to increase in amount over time.

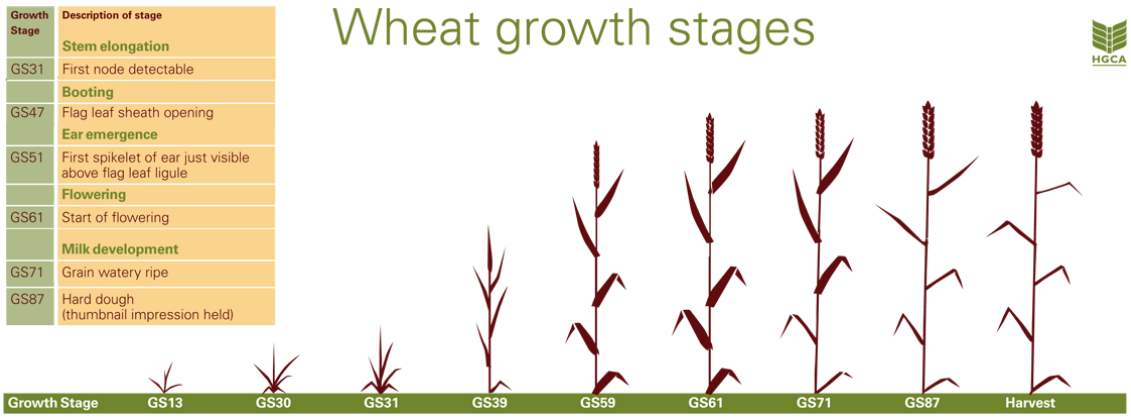


Figure 1.11: Wheat growth stages

Wheat growth stages are shown according to Zadoks scale. Modified after "The Wheat Growth Guide" released by the Home Grown Cereals Association (HGCA 2008).

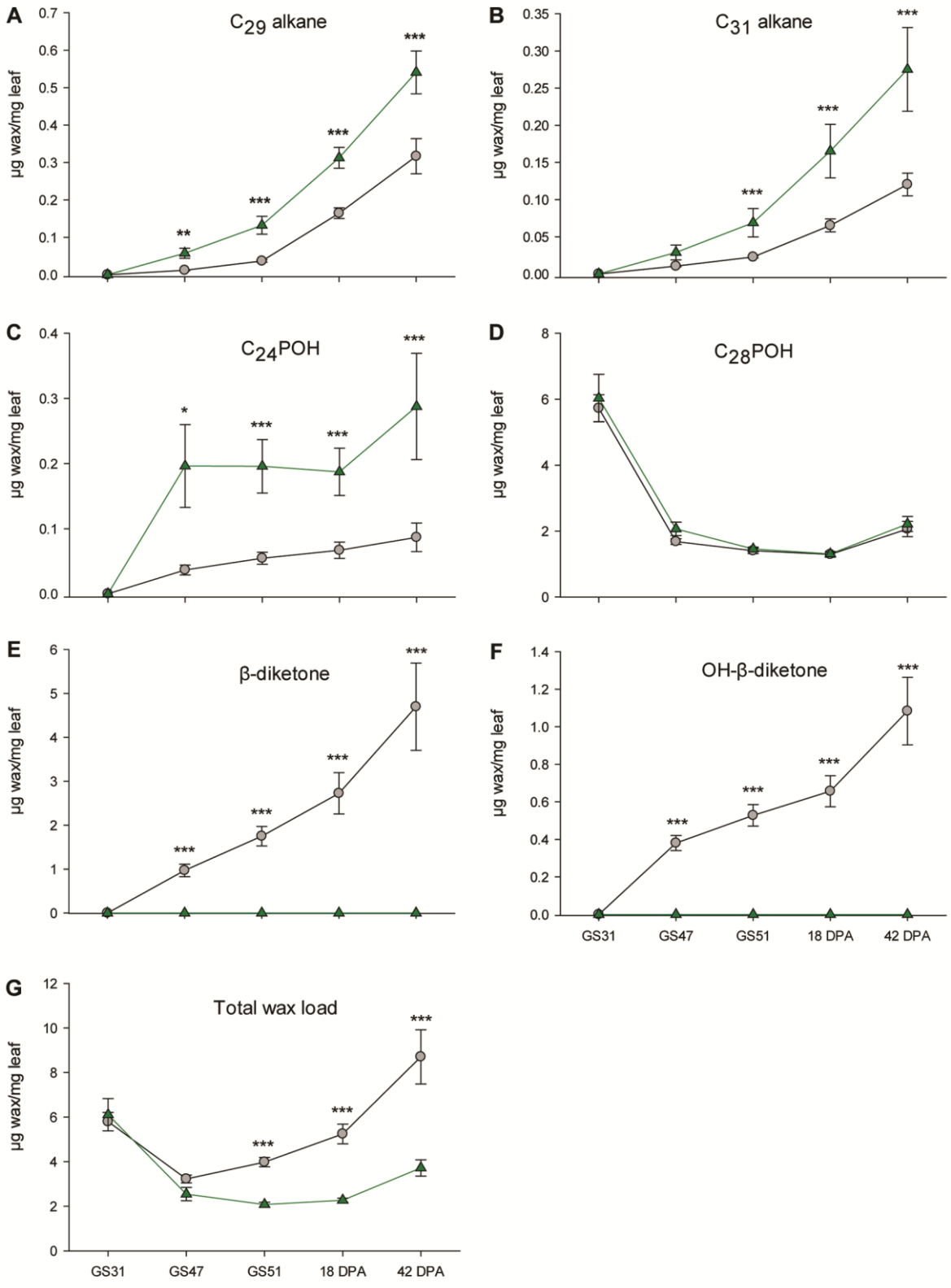


Figure 1.12: **Time course analysis of major cuticular wax components of BC_2F_3 NILs**

Time course analysis of major cuticular wax components of four field-grown BC_2F_3 NILs across five growth stages in 2011 (Zadoks GS31, GS47, GS51, 18 days post anthesis (DPA) and 42 DPA). Panels include C_{29} and C_{31} *n*-alkanes (A-B), C_{24} and C_{28} primary alcohols (C-D), β- and hydroxy-β-diketones (E-F) and total wax load (G). Green lines and triangles represent *Iw1* NILs; grey circles represent *iw1* NILs. Asterisks indicate significance at probability <0.05 (*), <0.01 (**), and <0.001 (***).

The effect of *Iw1* on cuticular wax composition was similar to the effects observed in the previous GC-MS analysis. Non-glaucous lines displayed significantly higher amounts of C_{29}- and C_{31}-alkanes, as well as C_{24}POH than their glaucous counterparts (Figure 1.12A-C). In glaucous lines, β-diketone aliphatics became the most abundant compound from GS51 onwards, making up 50% of all cuticular wax components, while in non-glaucous lines no β-diketone aliphatics were detected and hence C_{28}POH, amounting to 70% of all cuticular waxes, remained the dominant constituent of cuticular wax (Figure 1.12D-F). The total wax load was not significantly different between glaucous and non-glaucous lines until GS47, but with the ever increasing amounts of β-diketone aliphatics from GS51 onwards, the wax load of glaucous lines was becoming significantly higher (Figure 1.12G).

In summary, the biochemical analyses of section 1.2.2 consistently showed an inhibition in production of β- and hydroxy-β-diketones in *Iw1* germplasm coupled with an increase in the production of n-alkanes and primary alcohols. Interestingly, an increase in the amount of C_{24}POH and *n*-alkanes appeared to be linked with the decrease of β-diketone aliphatics in the non-glaucous *Iw1* lines.

1.2.3 Transmission electron microscopy analysis of glaucous and non-glaucous material

As was mentioned in the introduction the cuticle prevents excessive transpiration by forming a physical barrier that prevents water vapour from escaping. To determine the cuticle thickness, field-grown flag leaf blades and peduncles from both Shamrock and Shango were analysed using transmission electron microscopy (TEM; Materials and Methods 1.1.10). The TEM pictures were taken by Dr Kim Findlay.

The thickness of the cuticle was noted for all samples and analysed (Figure 1.13). The cuticles of Shango flag leaf blades had an average size of 0.204 ± 0.009 µm (adaxial) and 0.189 ± 0.005 µm (abaxial), while Shamrock flag leaf blade cuticles had an average size of 0.176 ± 0.016 µm (adaxial) and 0.172 ± 0.018 µm (abaxial). A single factor analysis of variance (ANOVA) did not show any significant differences in cuticle thickness between flag leaf blades of Shamrock and Shango, neither on the abaxial (p-value: 0.19) nor the adaxial (p-value: 0.40) side. Likewise, the cuticle thickness in peduncles of Shango (0.244 ± 0.024 µm) and Shamrock (0.239 ± 0.05 µm) was not significantly different from each other (p-value: 0.86). However, cuticle thickness was significantly different between peduncles and flag leaf blades (p-value: 0.02) irrespective of genotype.

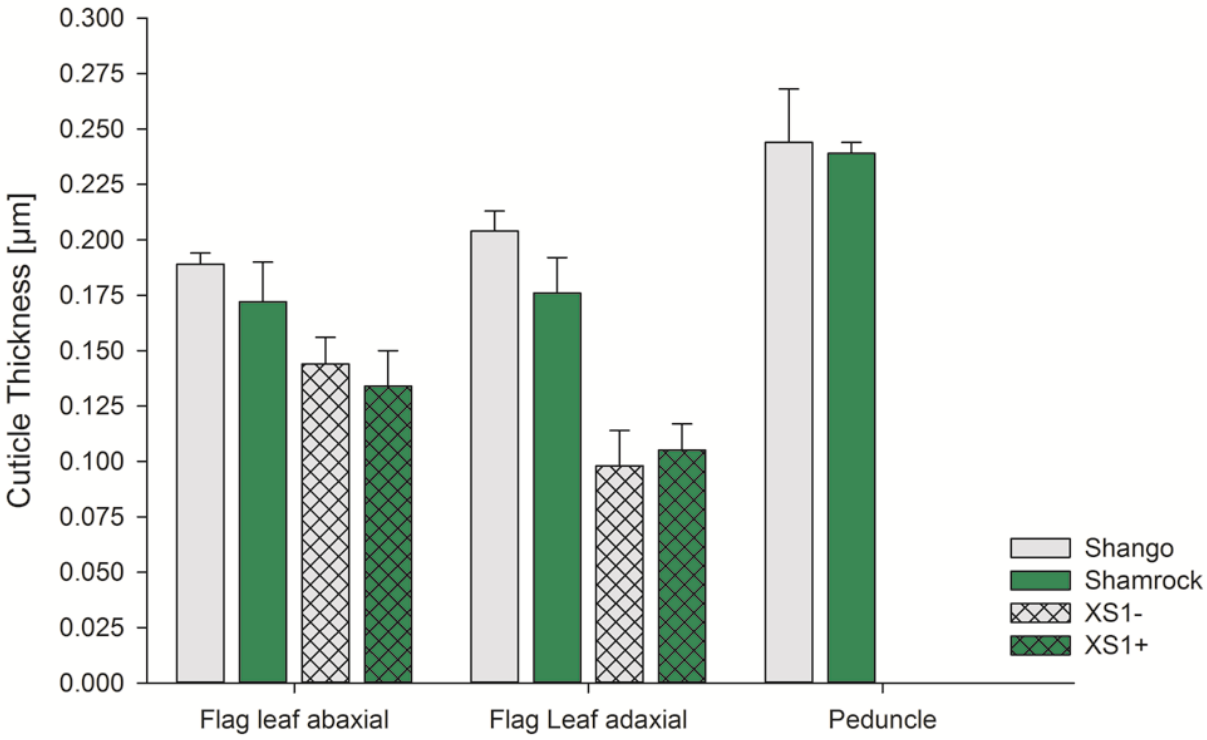


Figure 1.13: Cuticle thickness between glaucous and non-glaucous lines
The cuticle thickness between glaucous (grey bars) and non-glaucous (green bars) lines was analysed using TEM. No significant differences were detected between the two varieties Shango and Shamrock or between the two NILs XS1- and XS1+ in any of the analysed tissues. NS = not significant

In addition, glasshouse-grown flag leaf blades from a BC_2F_3 NIL pair, one with the *Iw1* introgression (XS1+) and one without it (XS1-), were also collected, prepared and analysed using TEM. The cuticle thickness was not significantly different between XS1- and XS1+ for the abaxial side (0.098 ± 0.016 and 0.105 ± 0.012 respectively; p-value: 0.76) or the adaxial side (0.144 ± 0.012 and 0.134 ± 0.016 respectively; p-value: 0.63) (Figure 1.13).

Figure 1.14 shows cuticles from glaucous and non-glaucous lines next to each other highlighting the lack of difference in cuticle thickness between glaucous and non-glaucous NILs. ANOVA showed that the cuticles of Xi19 (both parental and NIL pair) flag leaf blades were significantly smaller ($p<0.001$) than those from Shango or Shamrock, while tissue (p-value: 0.174) and genotype-tissue interaction (p-value: 0.221) were non-significant factors.

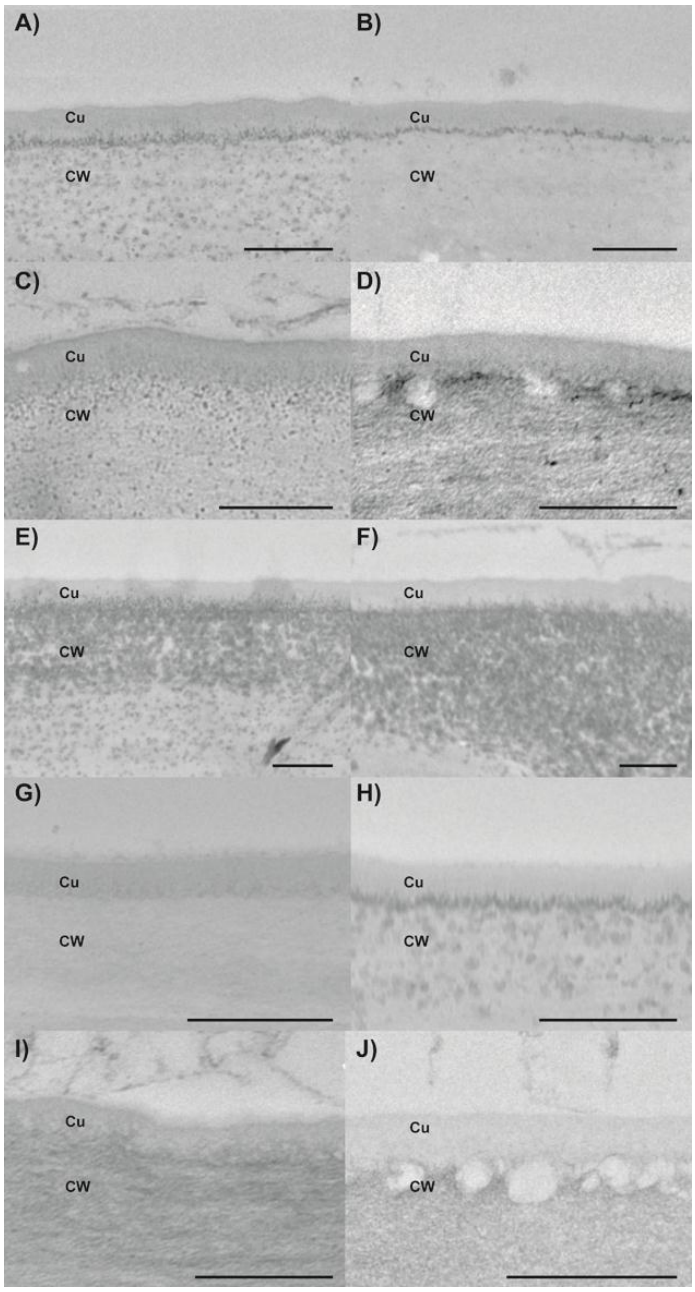


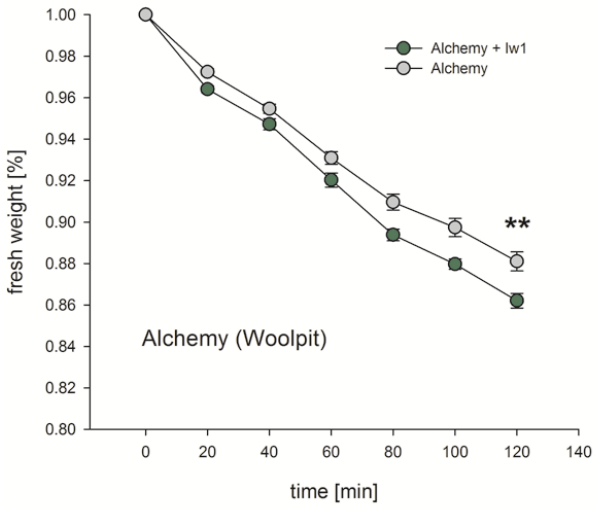
Figure 1.14: Transmission electron microscope images
The cuticle thickness between glaucous and non-glaucous lines was analysed using TEM images. The abaxial and adaxial side of flag leaf blades of Shango (A, C), Shamrock (B, D), XS1- (G, I) and XS1+ (H, J) was analysed in addition to peduncles of Shango (E) and Shamrock (F). No significant differences in cuticle thickness were detected between Shango and Shamrock in any tissue. Likewise, no significant differences in cuticle thickness were detected between the Xi19 NIL pair. The bar in each picture represents a size of 500µm. Cu = Cuticle; CW = Cell Wall

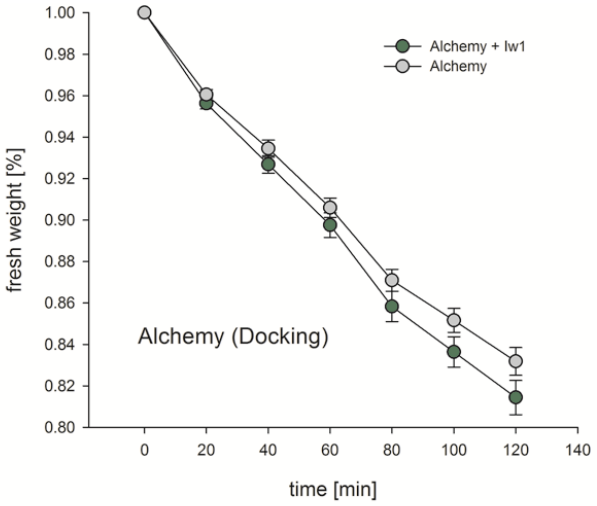
The results from this analysis clearly demonstrate that cuticle size is not affected by the introgressed genetic segment containing *Iw1*. By comparing cuticle thickness in Shamrock and Shango as well as a NIL pair, where the recurrent parent has a different cuticle size than the other two varieties, it is evident that this lack of difference in cuticle thickness is not unique to the Shamrock genotype.

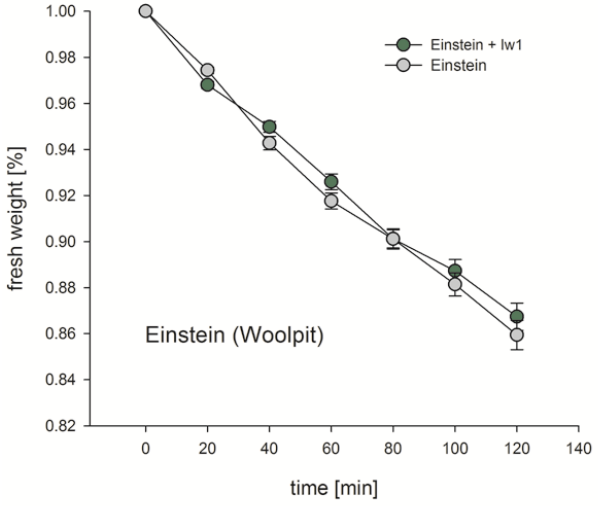
1.2.4 Analysis of cuticular conductance

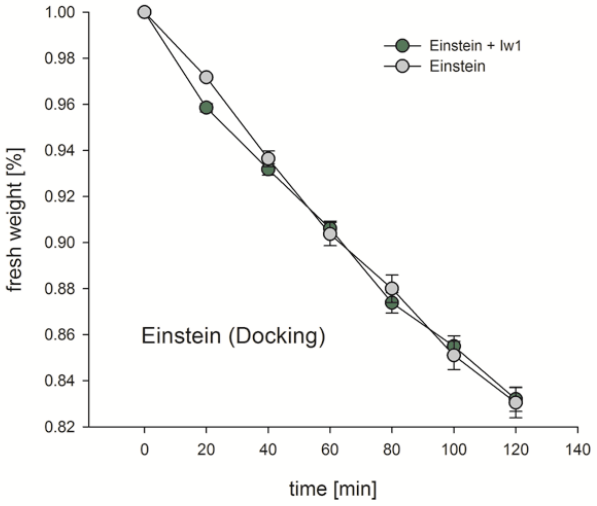
As was mentioned before, one of the major functions of the cuticle is to reduce water loss by forming a physical barrier preventing water vapour from escaping. Analysis of the cuticle by TEM did not show any significant changes in cuticle thickness upon introgression of the genetic segment containing *Iw1*. However, it is possible that while cuticle thickness remained unaltered, its composition and structure might have been affected. One way to test this hypothesis was to measure cuticular conductance. Cuticular conductance is a measure of how "watertight" a plant organ is, as it determines the amount of water that is being lost from a tissue over time while the stomata are closed.

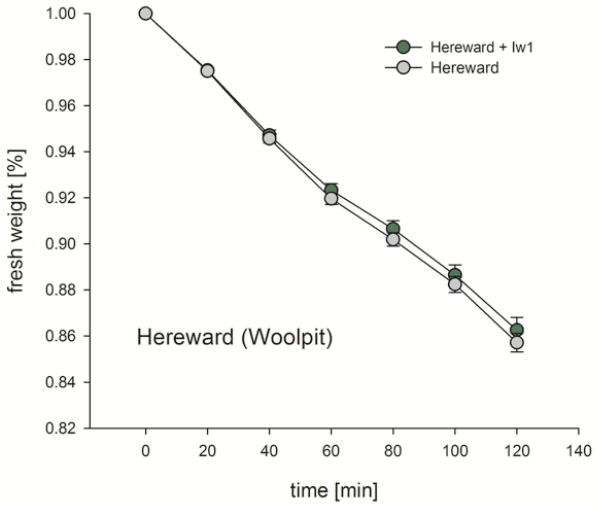
The method used here is based on Febrero et al. (1998) (Materials and Methods 1.1.11). Flag leaves of field-grown BC_2F_3 NILs (15 replicates per NIL) were collected from two locations in the UK (Woolpit and Docking), imbibed with water over night and then dried down, with the weight of the leaves being recorded in regular intervals. Plotting this data onto graphs (Figure 1.15) revealed that it was linear within the times studied. The percentage of fresh weight left at the last time point was recorded for all 15 replicates and compared between NIL pairs using a single factor ANOVA. Significant differences were observed between the AS4± (p-value: 0.0065) and XS1± (p-value: 0.033) NIL pairs at Woolpit and the RS1± (p-value: 0.037) NIL pair at Docking. In these NIL pairs the non-glaucous Alchemy, Xi19 and Robigus lines had lost 2.3%, 2.7% and 3.4% more water than their glaucous counterparts respectively, implying that these non-glaucous NILs have a higher cuticular conductance. However, this effect was only observed in one location for any of the three NIL pairs, while at the other location these NILs did not significantly differ from each other. The Einstein, Hereward and Malacca NILs did not exhibit this effect in any of the two locations either. Thus the enhanced cuticular conductance observed in some of the non-glaucous NILs is not caused by the introgression of *Iw1* but instead is most likely due to environmental factors and/or other genetic interactions.

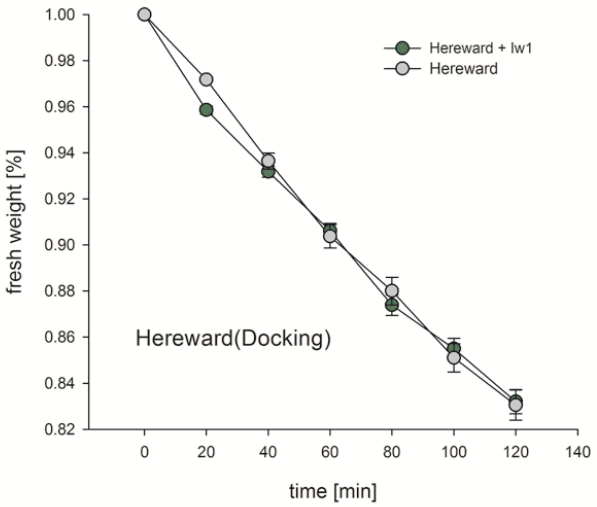

Alchemy + lw1
Alchemy
**
Alchemy (Woolpit)
fresh weight [%]
time [min]


Alchemy + lw1
Alchemy
Alchemy (Docking)
fresh weight [%]
time [min]


Einstein + lw1
Einstein
Einstein (Woolpit)
fresh weight [%]
time [min]


Einstein + lw1
Einstein
Einstein (Docking)
fresh weight [%]
time [min]


Hereward + lw1
Hereward
Hereward (Woolpit)
fresh weight [%]
time [min]


Hereward + lw1
Hereward
Hereward(Docking)
fresh weight [%]
time [min]

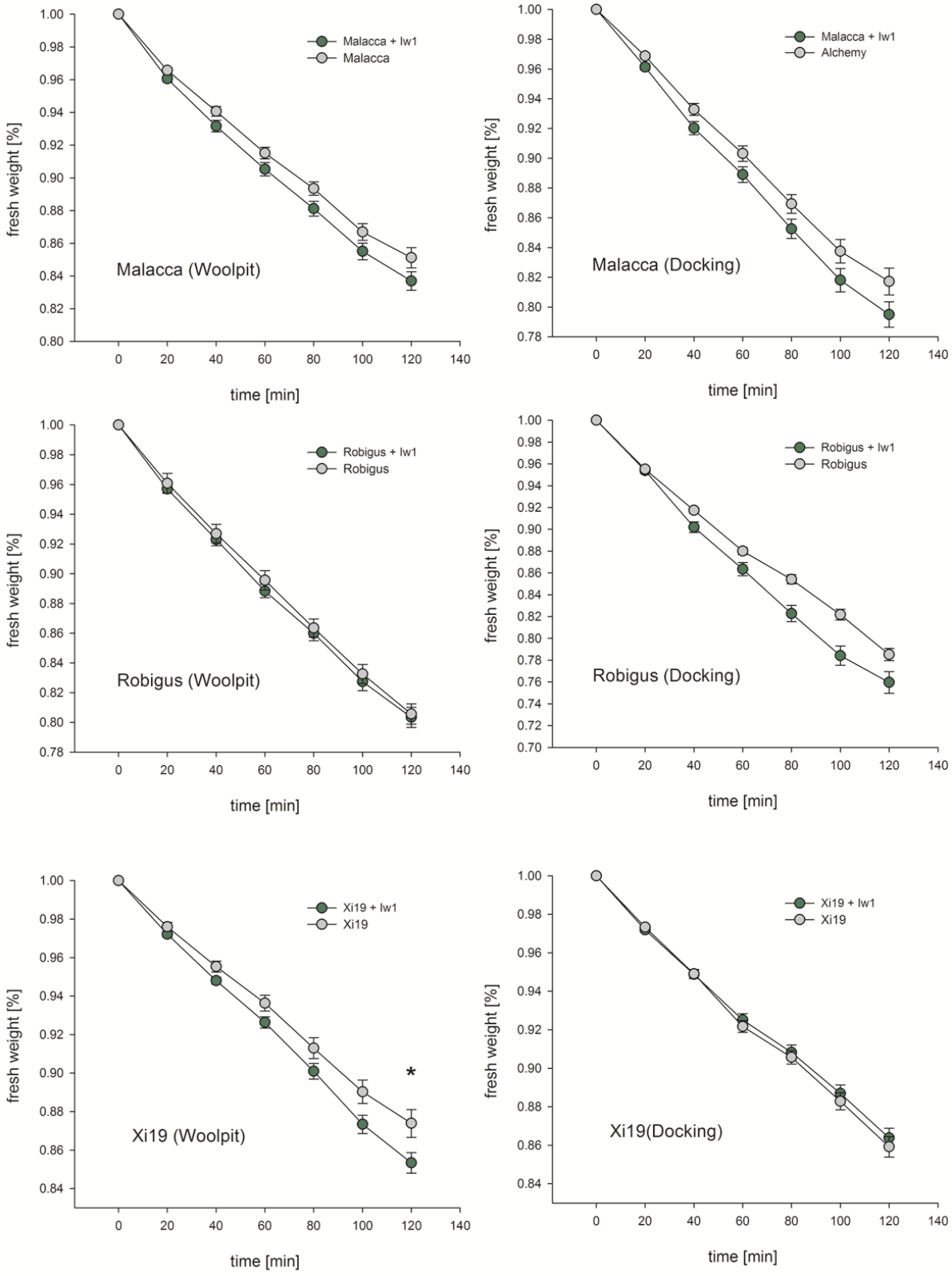


Figure 1.15: Cuticular conductance between NIL pairs

Field-grown flag leaves were collected from pairs of BC_2F_3 NILs at two locations in the UK (Woolpit and Docking) in 2011. Flag leaf blades were collected from all NILs in both locations and cuticular conductance was assessed. Significant differences were observed in some genotypes in a single location, but no overall effect could be seen in all genotypes and all locations. Green dots represent non-glaucous *Iw1* NILs, whereas grey dots represent glaucous *iw1* NILs. Asterisks indicate significance at probability <0.05 (*) and <0.01 (**).

This preliminary study suggests that the structure and composition of the cuticle is not drastically altered in germplasm containing the *Iw1* introgression as the six pairs of NILs seem to lose water at equal rates with their stomata closed and thus have the same cuticular conductance. However, these results need to be replicated in several locations over consecutive years before a definite conclusion can be reached.

1.2.5 Carbon isotope discrimination

Most of the carbon on earth (98.9%) is made up of C^{12}, but there is another naturally occurring stable isotope, C^{13}, which is much less abundant (1.1%). When plants fix carbon via photosynthesis they discriminate between CO_2 made with C^{12} or C^{13} carbon, preferentially using C^{12}. However, if stomata aperture is restricted, plants cannot afford to discriminate for C^{12} thus reducing the ratio of C^{13} to C^{12} relative to the atmospheric composition. This carbon isotope discrimination can be used as a surrogate to measure water-use efficiency (WUE) (Farquhar et al. 1989). Instead of using the isotope effect α, it was proposed to use the deviation of α from unity (Δ) as the measure of carbon isotope discrimination.

To determine whether the introduction of *Iw1* into UK germplasm and the associated change in epicuticular wax load and composition affect WUE we assessed the carbon isotope discrimination for tissues from field grown plants in 2011 and 2012 (Materials and Methods 1.1.12). Flag leaves and spikes of two BC_2F_3 NILs were collected at anthesis and later at maturity grain samples were also taken. The carbon isotope discrimination was performed by Dr. Nikolai Pedentchouk from the University of East Anglia. No significant effect was detected for *Iw1* in either 2011 or 2012. Carbon isotope discrimination in both NIL pairs was not significantly different in any of the three tissues examined (Figure 1.16). The Δ of flag leaves and grains, however, was significantly greater ($p<0.001$) in 2012 than in 2011 (by ~1.5 to 2.5‰, Figure 1.15A-B), reflecting the huge impact that environmental factors have on this type of measurement. These results indicate that glaucousness is not detrimental for WUE in a UK climate.

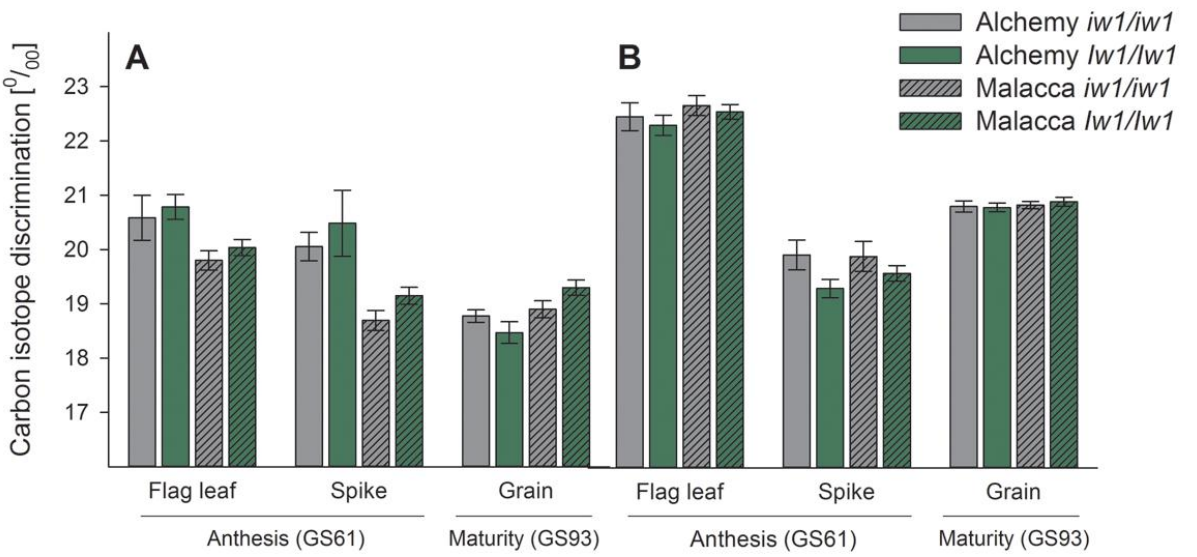


Figure 1.16: Carbon isotope discrimination of two BC_2F_3 NILs in 2011 and 2012
Carbon isotope discrimination of Alchemy and Malacca BC_2F_3 NILs in 2011 (A) and 2012 (B). Green bars represent non-glaucous *Iw1* NILs, whereas grey bars represent glaucous *iw1* NILs. No significant differences were detected between NIL pairs across tissues or years. GS: Zadok's growth stage.

1.2.6 The effects of *Iw1* on light reflectance and transmission

1.2.6.1 PAR reflectance of field-grown NILs

As was mentioned before, glaucousness is the appearance of a whitish bloom on the surface of plant organs, caused by a light-scattering effect of epicuticular crystals. This light-scattering effect decreases the amount of radiation absorbed in the visible and near-infrared spectrum, thus protecting the photosynthetic machinery and reducing leaf surface temperature (Blum 1975a, Blum 1975b).

To measure the difference in light reflection caused by *Iw1* field-grown pairs of BC_2F_3 NILs were assessed for their reflectance rate in the photosynthetic active radiation (PAR) spectrum (400-700 nm). Incident and reflected light was measured and the ratio between the two calculated for each line (Figure 1.17, Materials and Methods 1.1.13). Pair-wise comparison and ANOVA analysis showed that glaucous NILs reflect a higher amount of the incident PAR (13.73% more over all lines, locations and years; Table 1.3). This effect is independent of the genetic background, the location and the year. Note that in 2012 the Robigus and Xi19 NILs could not be measured due to excessive infection with *Septoria tritici*. These results agree with previous reports where it was shown that non-glaucous plants reflect less PAR than glaucous lines (Johnson et al. 1983, Febrero et al. 1998).

Figure 1.17 Measurement of PAR reflectance

PAR reflectance was measured using two sensors. One hemispherical sensor was measuring all incoming light in a 180° angle (red arrow), while a second sensor with a 25° collar attached to it measured the light reflected by the canopy (blue arrow). Three to four measurements were taken at different positions in a single plot, as described in Materials and Methods 1.1.13.

Table 1.3: PAR Reflectance measurements of field-grown BC_2F_3 NILs

The difference in the ratio of incident to reflected light between glaucous and non-glaucous BC_2F_3 NIL pairs is shown. Differences are significant in all NILs, locations and years except for Alchemy (AS4) at Docking in 2011 (see p-values).

	Docking 2011		Woolpit 2011		Churchfarm 2012	
NIL pair	**difference**	**p-value**	**difference**	**p-value**	**difference**	**p-value**
Alchemy	2.50%	2.22E-01	21.50%	2.92E-06	14.70%	8.17E-04
Einstein	4.19%	7.07E-05	9.22%	2.19E-02	18.03%	2.38E-02
Hereward	16.21%	5.16E-65	20.67%	1.91E-06	16.20%	2.64E-04
Malacca	15.84%	2.44E-10	16.51%	1.33E-05	13.55%	5.37E-03
Robigus	17.64%	8.46E-35	8.29%	5.44E-03	-	-
Xi19	14.23%	1.83E-23	21.45%	1.50E-07	-	-

1.2.6.2 Reflectance and transmission measurements using an integrating sphere

The results of the above mentioned PAR measurements are conclusive, but they represent a combined value over the entire PAR spectrum and do not allow us to distinguish the rate of reflectance for specific spectra of light (i.e. red or blue).

To elucidate whether the reduction in reflectance in non-glaucous plants applies equally over the entire PAR range or whether only certain spectra of light are actually affected by *Iw1*, flag leaf blades of two field-grown (Church Farm 2012) pairs of BC_2F_3 NILs were analysed using an integrating sphere. This device emits radiation of a specific wavelength and measures how much of the emitted light is being transmitted or reflected by the surface of a leaf (Materials and Methods 1.1.14).

The transmission and reflectance properties of both abaxial and adaxial sides of flag leaf blades from Alchemy and Malacca NIL pairs were measured. Abaxial and adaxial flag leaf surfaces between NIL pairs showed a remarkable similarity in the average transmission and reflectance properties. The values for transmission of radiation produced smooth graphs and no difference could be seen between the glaucous and non-glaucous Malacca NILs (Figure 1.18A).

In contrast, flag leaves of the non-glaucous Alchemy NIL transmit a greater percentage of radiation than the glaucous leaf irrespective of which side of the leaf is analysed (Figure 1.18B). The difference is largest in the green spectrum of light with the non-glaucous leaf transmitting an average of 1.166% more radiation than the glaucous leaf from 517 – 574 nm, peaking at 542 nm with 1.43%. On average the non-glaucous AS4 NIL transmits 0.78% and 0.74% more radiation through the abaxial and adaxial leaf surface than the glaucous AS4 NIL, respectively. However, a single factor ANOVA showed no significant difference between AS4 NIL pairs for either side of the flag leaf blade.

The glaucous and non-glaucous Malacca NILs display a similar reflectance profile, differing at most by 0.7% in the amount of radiation reflected by the leaf surface. On average, the non-glaucous Malacca NIL reflects 0.01% more radiation than the glaucous NIL (Figure 1.19A). In contrast, the non-glaucous NIL reflects 0.6% less PAR on average than the glaucous NIL from both sides of the flag leaf with the lowest reflection at 427 nm (2% and 2.5% less reflection than glaucous NIL on abaxial and adaxial side of the flag leaf respectively; Figure 1.19B). A single factor ANOVA though showed no significant difference in reflectance between AS4 NIL pairs on either side of the flag leaf blade.

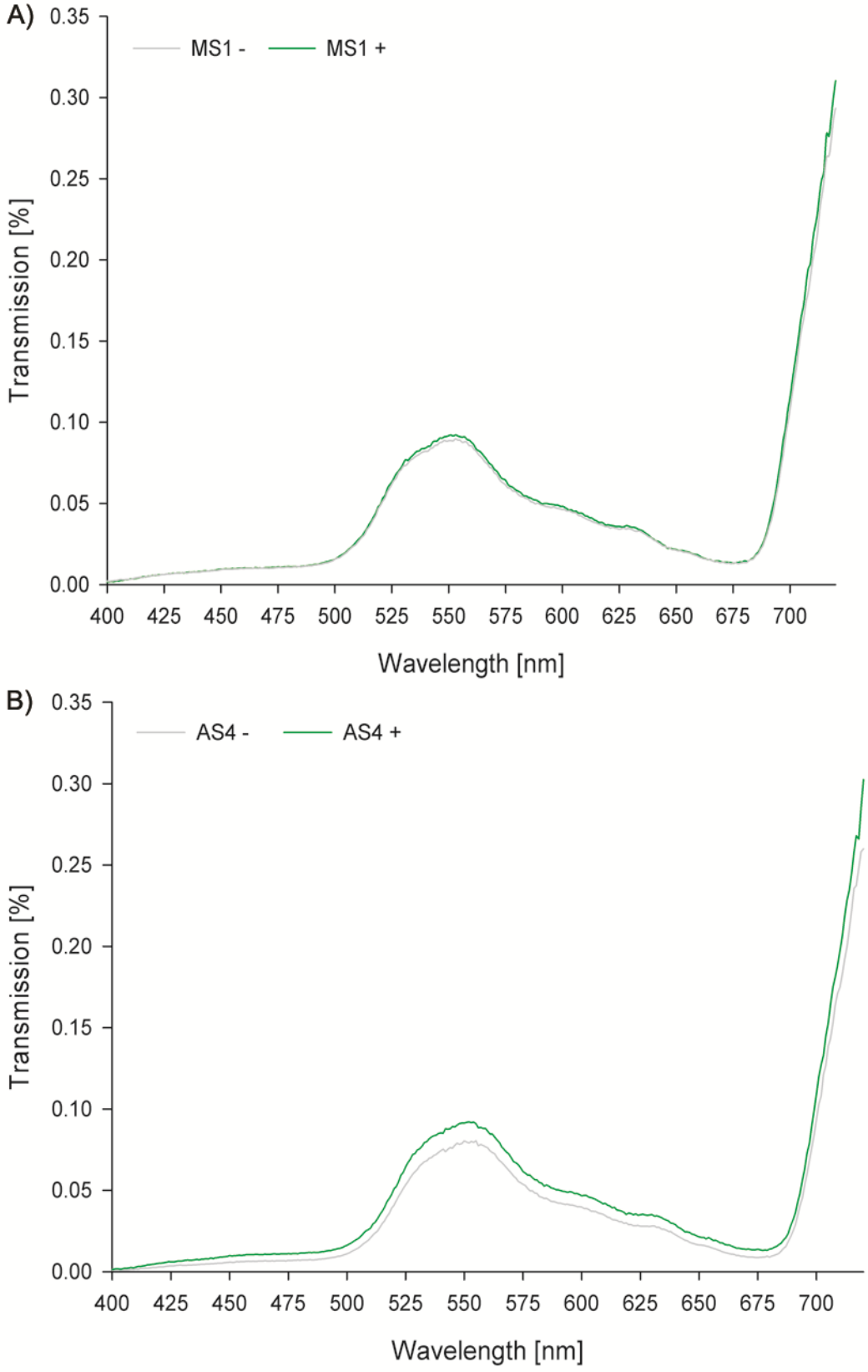


Figure 1.18: Transmission of PAR through the adaxial side of flag leaf blades

Transmission of PAR through the adaxial side of flag leaf blades of two BC_2F_3 NIL pairs is shown. In MS1 (A) transmission is similar in glaucous (grey line) and non-glaucous (green) NILs. In AS4 (B) the non-glaucous NIL appears to transmit more light than the glaucous NIL, although this is not significant. The difference is biggest in the green portion of the light (525 – 560 nm). The profile on the abaxial side is almost identical. Transmission is shown in decimals.

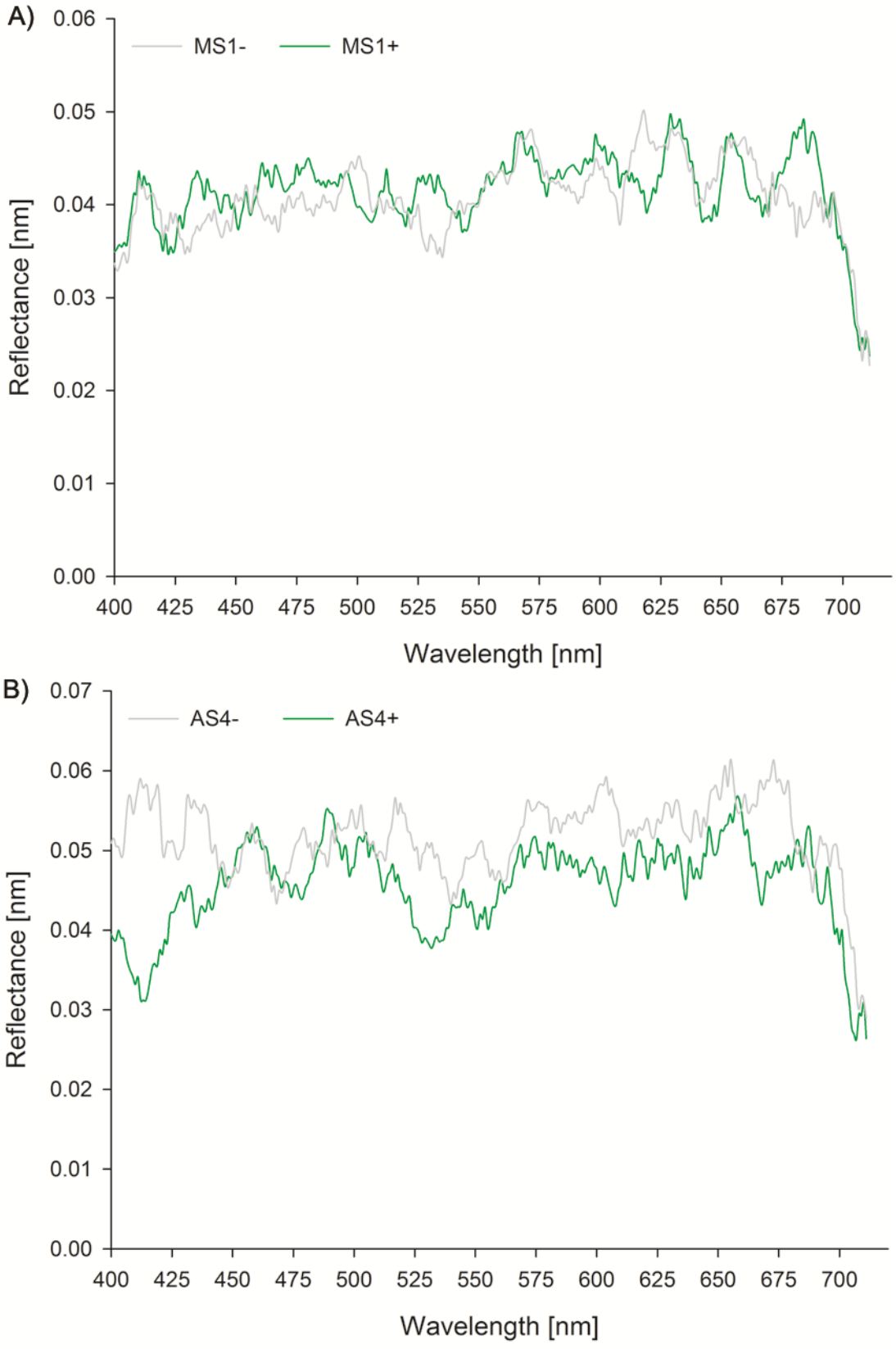


Figure 1.19: Reflectance of PAR through the adaxial side of flag leaf blades

Reflectance of PAR from the adaxial side of flag leaf blades of two BC_2F_3 NIL pairs is shown. In MS1 (A) reflectance is similar in both glaucous (grey line) and non-glaucous (green) NILs. In AS4 (B) the non-glaucous NIL reflects less light than the glaucous NIL. The difference is biggest in the blue portion of the light (400 – 450 nm). The profile on the abaxial side is almost identical. Reflectance is shown in decimals.

The results for AS4 agree with the PAR measurements of field-grown NIL pairs in terms of non-glaucous lines reflecting less of the incoming radiation than the glaucous lines. Furthermore, the lowest levels of reflection in the non-glaucous Alchemy NIL were observed in the blue and red portion of the light (Figure 1.19B). Unexpectedly there was a trend in the non-glaucous Alchemy NIL to transmit more PAR than its glaucous counterpart. This difference was not significant with the five replicates used here, but by using more biological replicates it might be possible to reduce the background noise and get a clearer result.

In contrast, the reflectance profile of MS1 is not different between glaucous and non-glaucous NILs, thus contradicting the reflectance results obtained from the same line in the field (see chapter 1.2.6.1). Comparing the reflectance values of the AS4 and MS1 NILs, it becomes obvious that the values for the non-glaucous lines are similar between the two genotypes (see Fig. 1.18A, B), while the reflectance values for the glaucous MS1 NIL are lower than those of the glaucous AS4 NIL. Likewise, the values for transmission of PAR are similar between the non-glaucous AS4 and MS1 lines, whereas transmission in the glaucous MS1 NIL is higher (almost as high as for non-glaucous AS4 and MS1 lines) than in the glaucous AS4 NIL. This indicates that the results for the MS1 NIL pair are not reliable, possibly because the (glaucous) MS1 plants did not cope as well with the stress of being dug up from the field as the other lines.

1.2.7 Effects of *Iw1* on green-canopy duration and yield

Previous data from a study of doubled-haploid lines showed an extended grain-filling period and an increase in yield in non-glaucous *Iw1* DH lines compared to the glaucous *iw1* DH lines (Simmonds et al. 2008). To elucidate whether these effects are caused by the *Iw1* introgression field trials using BC_2F_3 NILs were performed over several years in multiple locations in collaboration with the breeding companies Limagrain and RAGT.

Unfortunately the weather has been very challenging over the last few years. In the season 2010/2011 a serious drought in the spring caused the crop to form only one or two small tillers per plant: The UK Met Office reported that only 20% of the average rainfall from 1971-2000 was recorded in East Anglia, making this the driest spring in England and Wales since 1910 (Met_Office 2011). This drought had such a large impact that no reliable data could be obtained and thus no statistical analysis could be performed. In contrast, the season of 2011/2012 was dominated by heavy rain falls and low amounts of sunshine, especially during the months of June and July: The UK

Met Office reported that the average UK rainfall of 371 mm was the highest since the summer of 1912 (Met_Office 2012b).

1.2.7.1 Green-canopy duration

Green-canopy duration (GCD) was measured (see Materials and Methods 1.1.15) during three consecutive field seasons at Church Farm using BC_2F_3 NILs. GCD is correlated to the agronomic trait of grain-filling period (GFP), which is believed to have a positive impact on yield, but GFP itself was not measured directly. In the season of 2009/2010 the NILs were being bulked up in H90 plots and thus only one replicate per stream per NIL existed, i.e. four replicates (three for Einstein NILs) per genotype. All lines headed relatively late and senesced quickly, which is reflected in the relatively short GCD (see Table 1.4). On average GCD was extended in the glaucous lines by 0.64 days (p-value: 0.106). The differences between glaucous and non-glaucous lines were not significant in any of the pairs.

Table 1.4: Green-canopy duration of BC_2 NILs at Church Farm in 2009/2010
Green-canopy duration was measured in BC_2F_3 NILs at Church Farm in 2009/2010. Glaucous lines showed an extended GCD by 0.64 days on average compared to the non-glaucous NILs. However, none of these differences were significant. sem = standard error of the mean; cv = coefficient of variation

2009/2010	average [days]	Effect *+/w1*	sem	cv	p-value
AS-	43.3		0.48	2.21%	0.114
AS+	42.3	-1.0	0.25	1.18%	
ES-	46.0		0.58	2.17%	0.643
ES+	45.7	-0.3	0.33	1.26%	
HS-	44.3		0.25	1.13%	0.207
HS+	43.8	-0.5	0.25	1.14%	
MS-	44.0		0.41	1.86%	0.356
MS+	43.5	-0.5	0.29	1.33%	
RS-	46.5		0.65	2.78%	0.100
RS+	44.8	-1.7	0.63	2.81%	
XS-	44.3		0.48	2.16%	0.670
XS+	44.5	0.2	0.29	1.30%	

As was mentioned before the season of 2010/2011 did not yield any reliable data because of an historical drought during spring. The effects of the drought were so severe for all lines regardless of genotype that subtle changes in the microenvironment of a plot had dramatic effects on its performance in comparison with other plots.

The spring and summer of 2012 were very wet, greatly prolonging the GCD of all wheat lines. Five biological replicates of each stream of each NIL were analysed. All non-glaucous lines showed a slightly extended GCD when compared to their glaucous counterparts (Table 1.5). The GCD was not significantly different between individual BC_2 NIL pairs. However, when comparing all glaucous and non-glaucous lines a significant difference could be observed. On average the non-glaucous NILs had a significantly longer GCD (0.87 days, p-value: 0.048) than the glaucous NILs. In the BC_4 NILs however we could observe a significant extension of GCD in the non-glaucous NILs of Alchemy, Malacca, Robigus and Xi19 over their glaucous siblings. The overall effect was an increase in 1.73 days with p-value of 0.111.

Table 1.5: Green-canopy duration of BC_2 and BC_4 NILs at Church Farm in 2011/2012
Green-canopy duration was measured in BC_2F_3 NILs as well as BC_4F_3 NIL at Church Farm in 2011/2012. Overall the non-glaucous lines had a significantly extended GCD (0.87 days, p-value: 0.048) compared to the glaucous NILs. When comparing individual pairs of NILs only a non-significant trend towards an extended GCD in the non-glaucous NIL could be observed. sem = standard error of the mean; CV = coefficient of variation

	NIL	average [days]	Effect +/*w1*	sem	CV	p-value
BC2	AS-	73.15		0.6	3.66%	0.102
	AS+	74.35	1.2	0.39	2.36%	
	ES-	75.8		0.68	3.46%	0.464
	ES+	76.73	0.93	1.06	5.34%	
	HS-	75.7		0.4	2.35%	0.26
	HS+	76.35	0.65	0.41	2.38%	
	MS-	69.5		0.77	4.95%	0.385
	MS+	70.4	0.9	0.67	4.28%	
	RS-	70.46		0.46	3.36%	0.128
	RS+	71.41	0.95	0.4	2.91%	
	XS-	70.12		0.33	2.38%	0.24
	XS+	70.75	0.63	0.41	2.87%	
BC4	AS-	63.75		0.3	0.80%	0.019
	AS+	65.2	1.45	0.4	1.30%	
	ES-	72		0.7	2.00%	0.072
	ES+	74	2	0.6	2.20%	
	HS-	67		0.3	1.10%	0.407
	HS+	67.5	0.5	0.5	1.50%	
	MS-	63.8		0.2	0.70%	0.000
	MS+	65.6	1.8	0.2	0.80%	
	RS-	66.2		0.7	2.20%	0.036
	RS+	68.67	2.47	0.3	0.80%	
	XS-	62		1	2.30%	0.031
	XS+	64.14	2.14	0.3	1.40%	

1.2.7.2 Yield

Yield was measured in five locations in three consecutive years. The locations included Church Farm, two sites run by Limagrain (different locations each year) and two sites run by RAGT (different locations each year). Yield trials at Church Farm were performed using all streams of the BC_2 or BC_4 NILs (depending on the year) in five blocks, while only one stream of each NIL pair, but always the same one, was sent to the breeders for trialing in two blocks (see Materials and Methods 1.1.15).

As was mentioned before, an extreme drought period during the spring of the season 2010/2011 caused a poor development of the yield plots, especially with regards to the number of tillers (which was usually <2) and yield overall. The non-glaucous Alchemy and Hereward NILs yielded ~4.8% more than their glaucous counterparts, but the values were not significant (Table 1.6). Overall, the coefficient of variation for each NIL was very high, which indicates a high variability between the individual replicates. As the drought seemed to affect all NILs equally, the variations are most likely due to subtle changes in the microenvironment of each plot, e.g. soil structure, local nutrient content, etc. This makes it difficult to derive meaningful conclusions from this experiment.

Table 1.6: Comparison of yield between six BC_2 NIL pairs at Church Farm in 2010/2011
Yield trials for six BC_2F_3 NIL pairs were grown in one location in the season of 2010/2011. Yield is shown here both as the average yield per plot for each NIL as well as a percentage using the yield of the glaucous NIL as a baseline (100%) and comparing the yield of the non-glaucous NIL to it. The coefficient of variation for each NIL and the p-value for yield comparisons within NIL pairs is given.

NIL	Location	Average yield [kg]	CV	p-value	Yield as percentage
AS -	Church Farm	3.59	15.06%	0.128	104.87%
AS +	Church Farm	3.76	17.23%		
ES -	Church Farm	3.31	14.38%	0.749	98.61%
ES +	Church Farm	3.27	18.58%		
HS -	Church Farm	2.95	18.16%	0.09	104.88%
HS +	Church Farm	3.10	16.62%		
MS -	Church Farm	3.38	9.88%	0.973	99.91%
MS +	Church Farm	3.37	12.20%		
RS -	Church Farm	3.45	14.40%	0.418	97.13%
RS +	Church Farm	3.35	13.37%		
XS -	Church Farm	3.51	18.70%	0.553	97.95%
XS +	Church Farm	3.44	18.90%		

The results for the five locations for the season of 2011/2012 are shown below (Table 1.7). The Hereward derived non-glaucous NILs display a consistent trend of achieving a higher yield than their glaucous counterparts in all five locations. The size of this effect varies from 0.7% at Drinkstone to 9% at Ickleton, with a mean effect of 4.1% higher yield in non-glaucous than in glaucous NILs. These

differences in yield were only significant in two locations though (Wolverton and Church Farm, Table 1.7). In contrast, the Einstein derived non-glaucous NILs display a consistent trend of achieving a lower yield than their glaucous counterparts in all four out of five locations. Yield was reduced by 2.9% at Ickleton to 9.8% at Church Farm, with a mean effect of 3.8% lower yield in non-glaucous than in glaucous NILs. This reduction in yield was only significant in one location though (Dukes, Table 1.7). The remaining four NILs display varying effects of yield increase and decrease in different locations with a slightly positive but non-significant yield effect. Interestingly all non-glaucous NILs, except ES1, achieved high yields at Ickleton. This effect however was only significant in AS4.

Table 1.7: Comparison of yield between six BC_2 NIL pairs at five locations in 2011/2012

Yield trials for six BC_2F_3 NIL pairs were grown in five locations in the season of 2011/2012. Yield is shown here as a percentage using the yield of the glaucous lines as a baseline (100%) and comparing the yield of the non-glaucous NIL to it. Reduced yields are highlighted red, while increases in yield are highlighted green. Mean effect is the average yield effect of a non-glaucous NIL over all five locations. Asterisks indicate significance at probability <0.05 (*).

NIL pair	Drinkstone	Wolferton	Dukes	Ickleton	Church Farm	Mean effect
Alchemy	104.1%	97.9%	97.2%	108.9% *	95.5%	100.7%
Einstein	94.0%	105.7%	94.1% *	97.1%	90.2%	96.2%
Hereward	100.7%	107.2% *	102.2%	109.0%	101.4% *	104.1%
Malacca	99.5%	100.5%	102.1%	107.8%	99.0%	101.8%
Robigus	105.8%	94.4%	103.3%	105.4%	97.4%	101.3%
Xi19	103.8%	92.5%	103.0%	107.7%	104.0%	102.2%

During the season of 2011/2012 the BC_4 generation of NILs had been bulked up in the field and was ready to be sown for yield trials in 2012/2013. As in the year before the non-glaucous Hereward NILs displayed a positive yield trend in all locations, but the differences between glaucous and non-glaucous NILs were not significant in any location. Contrary to the previous year, the Einstein derived non-glaucous NILs also displayed a positive yield trend, albeit small and not significant in any location (Table 1.8). The other four NILs displayed again a mixture of positive and negative yield trends, none of which was significant, depending on location. The exception to that is AS4 at Church Farm, where the non-glaucous NIL showed a significant 5.7% increase in yield.

Table 1.8: Comparison of yield between six BC_4 NIL pairs at five locations in 2012/2013

Yield trials for six BC_4F_3 NIL pairs were grown in five locations in the season of 2012/2013. Yield is shown here as a percentage using the yield of the glaucous lines as a baseline (100%) and comparing the yield of the non-glaucous NIL to it. Reduced yields are highlighted red, while increases in yield are highlighted green. Mean effect is the average yield effect of a non-glaucous NIL over all five locations.

NIL pair	Woolpit	Wolferton	RAGT1	RAGT2	JIC	Mean effect
Alchemy	101.69%	98.84%	101.04%	98.65%	105.70%	101.18%
Einstein	101.28%	100.44%	112.66%	99.40%	101.76%	103.11%
Hereward	103.74%	105.28%	106.93%	105.64%	101.63%	104.64%
Malacca	99.31%	101.29%	100.57%	98.93%	101.05%	100.23%
Robigus	100.77%	97.57%	100.11%	102.86%	101.40%	100.54%
Xi19	98.82%	100.62%	92.03%	102.13%	101.35%	98.99%

In summary the field trials using six BC_2 and BC_4 NILs displayed mixed results for the two traits in question, namely GCD and yield. GCD was only assessed in two years with two different results. For the yield trials on the other hand a positive trend was observed in the non-glaucous Hereward NILs as well as an overall positive trend over all NILs, with the exception of the non-glaucous Einstein NILs. However, it is evident that more trials need to be performed before any clear conclusions can be drawn.

1.3 Discussion

The results of this chapter demonstrate the molecular effects that *Iw1* has on the cuticle and the composition of the epicuticular wax layer. The analysis of plant surfaces using cryo-SEM shows that *Iw1* specifically abolishes production or deposition of rod-like wax crystals onto the epicuticular wax layer, while platelet shaped wax crystals remain unaffected. The first type of compound has been associated with the presence of β-diketones, while the second one was postulated to consist mostly of primary alcohols (Netting and Wettstein-Knowles 1973). These initial results were substantiated by an extensive GC-MS analysis of cuticular waxes from different wheat lines. The analysis determined the composition of the cuticular waxes of glaucous and non-glaucous varieties and NILs, identifying 53 compounds and quantifying 26 of them. The obtained results match previous reports on the composition of wheat epicuticular waxes (Tulloch and Hoffman 1973, Tulloch and Hogge 1978). While numerous quantitative changes between cuticular waxes of glaucous and non-glaucous lines exist, the key difference between the two is the presence of large amounts of β-diketone aliphatics on the surface of glaucous plants, which are absent from cuticular waxes of non-glaucous plants. Instead cuticular waxes from flag leaf blades of non-glaucous plants consist mainly of C_{28}POH, a compound which is also highly abundant in cuticular waxes of glaucous flag leaf blades. These results are in accordance with previous studies, which reported loss of β-diketones in various non-glaucous plant species (Hall et al. 1965, Netting and Wettstein-Knowles 1973, Tulloch and Hoffman 1973). In addition, there is a significant increase in the amount of n-alkanes and C_{24}POH in cuticular waxes of non-glaucous plants.

The cuticular waxes of the exposed part of the peduncle from both glaucous and non-glaucous plants lacked C_{28}POH and instead showed a high presence of C_{24}POH and C_{29} primary alkanes. Again, cuticular waxes of non-glaucous plants showed a significant increase in these compounds over cuticular waxes from glaucous plants, while the total wax load was not changed significantly. This consistent increase in C_{24}POH and *n*-alkanes in *Iw1* non-glaucous lines raises the question whether it is directly connected with the absence of β-diketone aliphatics: This depends on the point in the biosynthetic pathway that is blocked in *Iw1* lines. The biosynthetic pathways of C_{24}POH, *n*-alkanes and β-diketone aliphatics diverges very early (Figure 1.2A, B). The precursors for C_{24}POH and *n*-alkanes are long carbon chain molecules that are further modified in the case of C_{24}POH (Figure 1.2A). In contrast, short carbon chain molecules serve as precursors for the β-diketone biosynthetic pathway, which are modified to form triketide-CoA intermediates that are subsequently elongated, thus yielding a molecule with mid-chain β-diketo groups (e.g. 15,17-β-diketo-C_{32}, Figure 1.2B). The presence of alkylresorcinols (ARs) in non-glaucous cuticular waxes suggests that the precursor

molecule common to ARs and β-diketone aliphatics is not affected by *Iw1*. This would argue against a direct rechanneling of carbon chains from the β-diketone pathway to the *n*-alkane/primary alcohol pathway. However, the pkKCS that forms the triketide-CoA molecule is different from the pkKCS that forms the tetraketide intermediate that is subsequently modified into ARs (Figure 1.2B), which suggests that *Iw1* affects the formation of triketide-CoA intermediates.

The deposition of wax compounds unto the plant surface appears to be a continuous process. Cuticular waxes of early leaf samples that were collected before the switch to the reproductive phase consisted solely of C_{28}POH. The composition of waxes became much more complex at anthesis with quantitative and qualitative differences in wax composition between flag leaf blades and peduncles as well as between glaucous and non-glaucous plants. Wax production and deposition continued until at least 42 days post anthesis (last data point collected) when plants were slowly starting to senesce. This demonstrates that wheat plants continue to invest into the production and maintenance of their cuticular waxes throughout their growth cycle.

The purpose of the cuticle and the cuticular waxes is to reduce the amount of water lost passively (outside of the photosynthetic cycle). This is accomplished by (i) forming a passive barrier (cuticle and embedded cuticular waxes) and (ii) by reflecting a portion of the incoming light and thus cooling the plant (epicuticular waxes). We have measured the thickness of the cuticle via transmission electron microscopy and have not found significant differences between our glaucous and non-glaucous germplasm. However, it has been argued that the thickness of the cuticle is not positively related to its ability to prevent passive water loss (Riederer and Schreiber 2001), which is why in addition we also analysed the cuticular conductance of our germplasm. While some of the analysed non-glaucous plants had a higher cuticular conductance than their glaucous counterparts in some locations, no significant differences could be detected overall. The observed effects were most likely due to specific genotype x environment interactions that were not related with the non-glaucous phenotype. The lack of a difference in cuticular conductance between glaucous and non-glaucous plants contrasts with previous scientific dogma that glaucousness reduces cuticular conductance (Blum 2005).

Our analysis of cuticular conductance was further substantiated by our comparison of carbon isotope discrimination in glaucous and non-glaucous NILs. The value for Δ can be used as a surrogate measurement for the ratio of the intercellular partial pressure of CO_2 to the ambient partial pressure of CO_2, which in turn allows us to predict differences in water-use efficiency (WUE) between samples (Farquhar et al. 1989). No significant differences could be observed between glaucous and non-

glaucous NILs in two consecutive years. However, the absolute values did differ between years for flag leaf and grain samples reflecting the difference in environmental conditions. A severe drought dominated most of the spring and summer of 2011 forcing the plants to restrict stomata aperture and preventing them from discriminating between C^{12} and C^{13}. In contrast, the summer of 2012 was the wettest summer on record, enabling plants to increase the aperture size of their stomata and discriminate against C^{13}. While the absolute values differed significantly between years ($p<0.001$) no difference was detected between glaucous and non-glaucous NILs. This contrasts with previously obtained results where glaucousness has been linked to an increase in WUE (Johnson et al. 1983, Richards et al. 1986, Premachandra et al. 1994). However, these studies were performed under Mediterranean conditions so it is not entirely surprising to discover that the results are conflictive.

We have also measured the reflectance and transmission properties of glaucous and non-glaucous germplasm and discovered a consistent reduction in reflectance of PAR in non-glaucous material. This is in agreement with previous studies (Johnson et al. 1983, Febrero et al. 1998). The extent of this reduction in reflectance of PAR varied between genotypes, field locations and field seasons, but the difference was always significant, with the exception of AS4 at Docking in 2011. Surprisingly, it appears that flag leaf blades of at least the non-glaucous Alchemy NILs transmit ~2% more light than the glaucous NILs. This experiment needs to be repeated before any definitive conclusions can be drawn, but it exemplifies the potential benefits of non-glaucousness for non-Mediterranean environments. By reflecting less of the (sparse) incoming radiation and instead transmitting excess light down into the canopy to other photosynthetically active tissue the amount of carbon fixed in a given period of time could be increased. However, whether this would have a direct impact on yield is unclear at this stage as yield is a very complex trait.

We have assessed yield in several locations over consecutive years using BC_2 (2010-2011 and 2011-2012) and BC_4 (2012-2013) NILs. In 2011 an extreme drought period during the spring resulted in very poor yields overall. The yield data for this year shows a high coefficient of variation, which indicates a high variability of the yield values, most likely due to differences in the microclimate of individual plots. This makes it difficult to conclude anything from the data. In 2012 and 2013, a positive yield effect for the non-glaucous Hereward NILs over all locations could be observed. The effect varied between the different years and locations from 0.7% to 9%, with an average yield increase of 4.37%, which is very close to the yield effect observed in the Shango x Shamrock DH lines (Simmonds et al. 2008). In contrast the non-glaucous Einstein NILs showed a variation in yield from -10% to +12.66% depending on the year and location, with a decrease in yield by -0.34% on average. Out of the six winter wheat varieties used to create the NILs Hereward was the lowest yielding one

(HGCA 2006). In contrast, out of the six varieties used Einstein had the second highest yield after Alchemy, yielding ~12% more than Hereward (HGCA 2006), and has recently achieved a world record in New Zealand with a yield of 15.637 t/ha (Limagrain 2011). This leads to the hypothesis that while Hereward might benefit from the introduction of the genetic segment containing *Iw1* in terms of yield, that same introgression might not affect, or even interact negatively with, the high yielding set of alleles in Einstein. However, Malacca has a similar yield to Hereward and has not shown such a consistent positive effect. This highlights the complexity of the yield trait, which is controlled by a multitude of genetic and environmental factors. Since the Hereward NILs responded positively to the *Iw1* introgression in terms of yield we will pursue the underlying genetic locus using recombinant lines in which the 20cM introgression has been broken up into smaller pieces. This will allow us to fine-map the yield locus and determine whether it coincides with the *Iw1* locus.

We have measured the green-canopy duration (GCD), which serves as an approximation for the grain-filling period (GFP); GFP is commonly associated with yield by farmers. We did not observe a significant difference in GCD over two different years in any of the six BC_2 NIL pairs. In these two years the weather had been very different, severely reducing the GCD across all NILs in 2009/2010, while increasing GCD in all NILs in 2011/2012. However, we could observe a significant increase in GCD in four BC_4 NIL pairs in 2011/2012 with an average increase of GCD of ~1.97 days in the non-glaucous NILs. This value is similar to the effect observed in the Shango x Shamrock DH lines (Simmonds et al. 2008). This also highlights the importance of highly isogenic germplasm for comparing physiological and agronomic traits. The BC_2 NILs used in the present study were ~87.5% isogenic, while the BC_4 NILs were 96.9% isogenic. This higher similarity reduces the variability of the phenotype, thus allowing detection of differences that would be otherwise masked by environmental factors. Likewise, the high isogenicity ensures that any effects observed are most likely caused by the genetic region under scrutiny. Thus, although the generation of NILs requires a lot of time and effort, the quality and reliability of any obtained results will compensate for that effort. Furthermore, it is imperative to create NILs as highly isogenic as possible to increase the chances of identifying significant effects.

In summary, the physiological effects of *Iw1* and thus non-glaucousness on wheat in a UK environment are not consistent with previous reports. Based on our results we can hypothesise that non-glaucousness is a beneficial trait in UK conditions. By reflecting less of the incoming light the plant surface heats up more. As the optimal temperature for photosynthesis in wheat is ~25°C (Kobza and Edwards 1987), this might have a beneficial impact on RUBISCO activity with regards to the relatively low average temperatures in the summer. With possibly more light being transmitted

through the canopy the potential for an enhanced photosynthetic rate is further increased. At the same time non-glaucous plants do not have a higher cuticular conductance or reduced WUE compared to glaucous plants. This was also recognized by the UK breeding community, where recently a non-glaucous winter wheat cultivar (Crusoe) was accepted into the HGCA recommended list for its high quality grain, good yield and good disease resistance. This cultivar is a descendant of the variety Shamrock and thus also carries *Iw1*.

Contrary to what the name "*Inhibitor of Wax 1*" suggests, it is evident that only a discrete type of wax is actually affected by this gene. The same observation has been made in a class of barley mutants, called *eceriferum*, which is Latin for "bearing no wax", where β-diketones are reduced or completely abolished, depending on the specific mutant, but other types of wax are still present (Wettstein-Knowles and Netting 1976). So while glaucousness can be attributed to the presence of β-diketones non-glaucousness does not equal non-waxiness. This highlights the need for a more formal phenotypic description of "glossy", "waxless" and "highly waxy" varieties, cultivars and land races in order to postulate accurate scientific theories about the interaction of a plant with its environment.

Chapter 2 Genetic and physical mapping of the *Iw1* locus

Introduction

In the following paragraphs an overview of genetic loci known to be involved in production or inhibition of glaucousness in wheat will be given. Numerous loci have been reported to control glaucousness to various extents and in various plant organs.

Non-domesticated grass species, like *Aegilops speltoides*, *Aegilops tauschii* and *Triticum turgidum* ssp. *dicoccoides* (TTD), show a huge range of variation in terms of glaucousness. This variation however was not transferred into domesticated species, like *Triticum aestivum* (common wheat; glaucous) or *Triticum monococcum* subsp. *monococcum* (Einkorn; non-glaucous), where the level of glaucousness is almost completely fixed and there is only little variation.

Glaucousness on a whole plant level in tetraploid wheat is controlled by two distinct loci situated at the distal end of the short arm of chromosome 2B (2BS). The first one, designated *W1*, is a dominant gene that promotes glaucousness on all aerial parts of a plant, e.g. the leaf sheath, peduncle, spike, etc. (Allan and Vogel 1960, Tsunewaki 1964). Plants lacking *W1* or having two copies of a non-functional version of *W1*, i.e. genotype *w1*/*w1*, are non-glaucous. In contrast, the second locus, called *Inhibitor of Wax 1* (*Iw1*), is a dominant inhibitor of glaucousness on all plant aerial surfaces (Jensen and Driscoll 1962, Driscoll and Jensen 1964). It is also dominant over *W1*, i.e. the presence of both genes will result in a non-glaucous plant due to the actions of *Iw1*. It was suggested that *W1* and *Iw1* are allelic, but it was shown that these two loci can recombine, which proves that although they are very close to each other they are also distinct from one another (Tsunewaki and Ebana 1999).

Homoeologues of *W1* and *Iw1*, called *W2* and *Iw2*, are present on the distal end of the short arm of chromosome 2D (2DS) in a homoeologous position (Tsunewaki 1966). These homoeologues loci work in the same way as their 2B counterparts: *W2* promotes glaucousness, while *Iw2* is dominant over *W2* (and *Iw1*) and causes a non-glaucous phenotype. In hexaploid wheat a single copy of either *W1* or *W2* is sufficient to elicit a whole-plant glaucous phenotype. Likewise, a single copy of either *Iw1* or *Iw2* is sufficient to repress the action of either or both *W1* and *W2* and thus cause a non-glaucous phenotype (Tsunewaki and Ebana 1999). Functional copies of the *W1* and *W2* genes are present in different combinations in modern wheat germplasm. In contrast the *Iw1* and *Iw2* loci are non-functional or perhaps even absent from modern wheat germplasm and only exist in non-domesticated wheat species (Tsunewaki and Ebana 1999). Exceptions are the synthetic hexaploid wheat lines and their derivatives generated by CIMMYT and at the National Institute of Agricultural

Botany (NIAB). Most of the almost 400 *Aegilops tauschii* accessions at NIAB used to make synthetic hexaploid wheat are non-glaucous. Only a single accession was found to be glaucous, which indicated that most of these accessions carry a functional copy of *Iw2* or lack *W2* and are thus non-glaucous.

In addition to these major loci, a number of genes controlling glaucousness on a plant organ level have been described.

Three loci were reported to affect spike glaucousness. The first one was found at the distal end of the short arm of chromosome 1B in a set of chromosome substitution lines. In these lines, chromosome 1B of durum wheat cultivar Langdon was replaced by chromosome 1B from a TTD accession. The non-waxy spike phenotype was noted to be dominant over the waxy spike phenotype, hence the locus was named *Iw3*, in accordance with the two other known inhibitors of wax, *Iw1* and *Iw2* (Dubcovsky et al. 1997). The second locus, named *WS*, was described a few years later by another research group who utilized it as a morphological marker in their mapping project (Peng et al. 2000). This marker was also mapped to the distal end of the short arm of chromosome 1B (1BS) in a cross between durum wheat cultivar Langdon and TTD accession Hermon H52, but no mention was made whether the underlying gene is dominant. It is not known whether *Iw3* and 1BS *WS* are allelic or not. The third locus, also named *WS* and also used as a morphological marker for spike glaucousness, mapped to the distal end of the short arm of chromosome 1A (1AS) in a population derived from the durum wheat cultivars Ciccio and Svevo (Gadaleta et al. 2009). No further mention was made on the properties and effects of the 1AS *WS* locus. Given the similar location it is likely that 1AS *WS* and 1BS *WS* are homoeologous loci.

Recently, a novel QTL responsible for quantitative variation in flag leaf glaucousness was detected on chromosome 3A in a bread wheat doubled-haploid population of Kukri and RAC875 (Bennett et al. 2012). The QTL was named QW.aww-3A and was shown to be causal for up to 52% of the variation in flag leaf glaucousness, i.e. plants carrying the QTL had visibly more wax on their flag leaves than plants without it. In addition, two studies using the 'International Triticeae Mapping Initiative' reference population identified QTL for waxiness on 1DL, 2DL and 4AL (Börner et al. 2002) and on 1A, 1D, 2B, 2D, 6A, 7A and 7D (Kulwal et al. 2003). No mention was made about their mode of action (dominant or recessive), which organ(s) were affected or how big these effect were.

Loci controlling glaucousness were also discovered in other grass and cereal species. A whole plethora of loci was detected in *Hordeum vulgare* (barley) by way of mutagenesis: 1,580 mutants affecting epicuticular waxes in barley have been described (Lundqvist and Lundqvist 1988). These so-

called *eceriferum* (*cer*) mutants have been localized into 79 different loci, which in turn were ordered into five categories: Leaf blade mutants (25 loci), spike mutants (23 loci), spike and leaf sheath mutants (8 loci), spike, leaf blade and leaf sheath mutants (4 loci) and partial mutants (19 loci) (Lundqvist and Lundqvist 1988). The loci underlying the *cer* mutations are also known as *glossy sheath* (*gsh*). The three mutants *cer-c*, *-q* and *–u* inhibit different steps leading to the production of β-diketones, hydroxy-β-diketones and esterified alkan-2-ols. Furthermore, these three mutants map to barley chromosome arm 2HS and are completely linked, even after screening 26,933 gametes (von Wettstein-Knowles and Søgaard 1980), which raises the question whether *cer-cqu* (*gsh6*) is a gene cluster or a multifunctional gene. Except for one mutation all *cer-cqu* mutations are recessive (King and von Wettstein-Knowles 2000), which together with its location suggests that the *cer-cqu* locus is the barley homolog of *W1*. Only one *cer* locus shows a consistent dominant inhibition of glaucousness as is seen in *Iw1*; all 18 *cer-yy* mutants exhibit a dominant inhibition of spike glaucousness, but these mutations map to barley chromosome 1H, making it unlikely that this locus is the barley homolog of *Iw1* (Lundqvist and Wettstein-Knowles 1982).

Only a single locus controlling glaucous was described in *Secale cereale* (rye). In a cross between the glaucous inbred line N6 and the non-glaucous line SI a mutation of the *waxless plant 1* (*wa1*) locus was mapped to the distal end of the long arm of rye chromosome 7 (7RL) (Korzun et al. 1997). The *wa1* mutation is recessive and it was proposed in the same study that the *wa1* is homologous to the *w1* and *w2*, as these are also recessive mutants leading to non-glaucousness.

At least 18 loci have been found in maize that alter the composition and quantity of cuticular waxes on leaves of young maize (*Zea mays*) seedlings (Neuffer et al. 1997). Note that leaves of young maize seedlings are glaucous, while leaves of adult plants are glossy (non-glaucous), which is exactly the opposite in wheat and barley where young plants are glossy and become glaucous after the switch to the reproductive phase. A number of the genes underlying the so-called *glossy* loci have been cloned in maize: The *Glossy 1* (*GL1*) gene encodes a transmembrane protein with desaturase/hydroxylase domains (Sturaro et al. 2005), *Glossy 2* (*GL2*) encodes a protein of the omega-hydroxypalmitate O-feruloyl transferase superfamily (Tacke et al. 1995), *Glossy 8* (*GL8*) probably encodes a β-ketoacyl reductase (Xu et al. 1997), while *Glossy 15* (*GL15*) encodes a member of the *Apetala2* transcription factor family (Moose and Sisco 1996).

The number and diversity of loci involved in controlling glaucousness highlights the complexity of the underlying genetic network. But one part of this genetic network seems to be unique to wheat, the presence of dominant inhibitors of glaucousness. None of the described loci in barley, rye or maize

fits all the criteria that a true homolog of *Iw1* or *Iw2* should fulfil. This raises the question whether *Iw1* and *Iw2* are truly unique to wheat, i.e. whether these genes were acquired after the divergence from the other grass species. An alternative hypothesis is that a gene common between grass species acquired a new function in wheat, a process known as neo-functionalization, which can occur frequently in wheat (Teshima and Innan 2008, Bartoš et al. 2012). The apparent absence of a homologous locus from other grass species might negatively affect the identification and validation process of *Iw1*, but at the same time it highlights the importance of crop research. Genes of interest cannot always be cloned in genetically simple model species, but instead need to be pursued in the organism of interest. Prominent examples of this include the domestication locus *Q* (Faris et al. 2003), the chromosome pairing locus *Ph1* (Griffiths et al. 2006), the grain quality controlling locus *GPC-B1* (Uauy et al. 2006) or disease resistance genes like *mlo* (Büschges et al. 1997) or *Lr21* (Huang et al. 2003) to name a few.

Here, we have performed a large recombinant screen followed by the initiation of a physical map to isolate the highly interesting gene *Iw1* via a positional cloning approach.

2.1 Materials and Methods

2.1.1 Plant material

Shango, Shamrock and the Shango x Shamrock DH population have been described in Chapter 1.1.1. Shamrock was used as pollen donor for a cross to Shango to generate a large F_2 mapping population. TTD140 is a non-glaucous *Triticum turgidum* ssp. *dicoccoides* accession from Israel that was shown to carry *Iw1* (Rong et al. 2000). Langdon (LDN; *Triticum turgidum* ssp. *durum*) is a glaucous durum wheat cultivar. TTD140 was used as pollen donor for a cross to Langdon to obtain several hundred F_2 plants.

2.1.2 96 well DNA extraction from cereal leaves

The procedure is adapted from Shiaoman Chao and Daryl Somers (Chao and Somers). Briefly, it consists of the following steps with catalogue numbers included:

1. Add a 3 mm Tungsten Carbide bead (Qiagen, #69997) to each well of a 96-well 1.2 mL storage plate (ABgene, #AB-0564) using a 96 well 3 mm bead dispenser (Qiagen, #69973).
2. Add 100 mg leaf tissue to each well.
3. Lyophilise samples in plate overnight.
4. Preheat extraction buffer (0.1M Tris-HCl pH 7.5, 0.05M EDTA pH 8.0, 1.25% SDS) to 65°C.
5. Seal plate with a 96 Cap Sealing Mat (Abgene, #AB0674). Transfer plate to Genogrinder 2000 and disrupt tissue at 800-1,200 rpm for 3-5 min.
6. Add 500 µL of extraction buffer to each well, seal the plates and shake thoroughly. Incubate the plate at 65°C for 30-60 min.
7. Place the plates in the fridge to cool them down to room temperature (about 15 min) before adding 250 µL 6M ammonium acetate, which is stored at 4°C. Shake vigorously to mix in the ammonium acetate and then leave to stand for 15 min in the fridge.
8. Centrifuge the plate for 15 min at 5,000 rpm in a Sigma 4-15 centrifuge to collect the precipitated proteins and plant tissue.
9. Recover 600 µL of the supernatant into new storage plate containing 360 µL of propan-2-ol in each well. Mix thoroughly and allow the DNA to precipitate for 5 min.
10. Centrifuge the samples for 15 min at 5,000 rpm in order to pellet the DNA and then tip off the supernatant. Allow the remaining fluid to drain off the DNA pellet by inverting the tubes onto a piece of paper towel.
11. Wash the pellet in 500 µL of 70% ethanol.
12. Centrifuge the plate for 15 min at 5,000 rpm and discard the supernatant.
13. Resuspend the pellet in 100 µL of ddH_2O and vortex plate.
14. Place the plate at 65°C for 20 min and vortex again afterwards.

15. Spin down the un-dissolved cellular debris by centrifuging the plate for 20 min at 5,000 rpm.
16. Store DNA at 4°C.

For smaller number of samples, the same protocol was applied but using 1.5 mL Eppendorf tubes instead of 96-well plates.

2.1.3 Polymerase chain reaction (PCR)

Unless otherwise stated, polymerase chain reaction (PCR) was performed using a standard recipe and protocol:

PCR recipe:

Template	50 ng
Forward Primer	0.5 µL of a 10 µM working stock
Reverse Primer	0.5 µL of a 10 µM working stock
5x Promega Reaction Buffer	3 µL
dNTPs	0.3 µL of 10 mM stock (0.2 mM final concentration)
$MgCl_2$	1.2 µL of 25 mM stock (2 mM final concentration)
Taq	0.07 µL
ddH_2O	fill up to total reaction volume of 15 µL

PCR protocol:

Initial denaturation	94°C for 2 min
Number of cycles	40 cycles
Denaturation	94°C for 15 sec
Annealing	60°C for 20 sec
Elongation	72°C for X sec (60 sec/1 kb product size)
Final Elongation	72°C for 2 min
Store	16°C forever

2.1.4 PCR on bacterial colonies

Bacterial colonies are picked from LB plates using standard pipette tips. The colonies are used to inoculate a so-called master (LB agaorse + 50 µg/mL Carbenicillin) plate and then dipped into a tube containing 50 µL of ddH_2O. The colonies on the master plate are allowed to grow at 37°C overnight. The water containing some of the bacterial colony is used as a template for standard PCR. Positive clones can be picked from the master plate on the next day.

2.1.5 DNA sequencing using BigDye

This protocol uses BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) to prepare DNA samples for Sanger sequencing. The protocol includes the following steps:

1. Measure DNA concentration of the samples to be sequenced.
2. Recommended amounts are:

    100-200 ng of standard cloning plasmids (pBluescript, pGemT, etc.)

    500 ng of binary vector

    50 ng of PCR product of 1 kb length; for products of different lengths calculate accordingly

3. BigDye recipe:

Template	X µL DNA (see above for recommended amounts)
5x BigDye sequencing buffer	1.5 µL
Primer	3.5 µL (1 µM concentration)
BigDye reagent	1 µL
ddH_2O	4-X µL

4. PCR protocol:

Initial denaturation	94°C for 2 min
Number of cycles	26 cycles
Denaturation	94°C for 10 sec
Annealing	55°C for 30 sec
Elongation	60°C for 4 min
Store	16°C forever

5. Submit samples to TGAC for sequencing.

2.1.6 PAGE for SSCP products

Products of single-strand conformation polymorphism (SSCP) markers were visualized using polyarcylamide gel electrophoresis (PAGE). The reagents, solutions and steps are outlined below:

Reagents:

Gel solution mix for one gel:

MDE gel solution (Cambrex Bio Science)	16.5 mL
TTE Buffer (National Diagnostics)	2 mL
50% glycerol solution	12 mL
ddH_2O	35 mL

To polymerise the gel add:

10% ammonium persulfate (Sigma)	400 µL
TEMED (Sigma)	36 µL

Denaturing loading dye:

Formamide	98 mL
EDTA (0.5M, pH 8.0)	0.4 mL
Bromophenol Blue	20 mg

Xylene Cyanol FF	20 mg

1x Running buffer:

20xTTE	50 mL
ddH_2O	950 mL

Staining solutions:

Fixer: Mix 200 mL of 10x acetic acid with 1,800 mL of distilled water.

Silver stain: Add 12 mL Silver nitrate solution (1N) and 3 mL formaldehyde (40%) to 2 L of distilled water (this solution can be used 10 times).

Developer: Dissolve 60 g Sodium carbonate (anhydrous) in 2 L ddH_2O, chill to 5°C in cold room. Just before use add 3 mL formaldehyde and 300 µL Sodium thiosulphate.

1. Clean large glass plate with Liquinox (Sigma) and water. Rinse and dry using 100% Ethanol. Apply approximately 100 µL Repelcote (Fisher Scientific) all over the plate with blue roll. Allow this to dry a few minutes, then wipe with more Ethanol and dry.
2. Clean small glass plate with Liquinox and water. Rinse and dry using 100% Ethanol. Apply 30 µL Bind Silane (GE Healthcare), spread all over plate with blue roll. Allow to dry, rinse with 100% Ethanol twice and dry.
3. Ensure spacers are clean and dry. Align them on the edges of large plate. Place the small plate on top and re-align. Ensure there are no spaces between the plate edge and the sponge on the spacers. Secure base of plate sandwich with bulldog clips.
4. Invert mixture gently; pour between plates using a syringe or small plastic squeezy bottle. Insert clean dry comb so that teeth lie flush to the edge of large glass plate. Secure edges and top of plate with bulldog clips. Allow at least 60 min for polymerisation.
5. Remove comb, clean to remove all bits of gel. Clean up gel plates with water. Place gel in gel rig and secure. Pour cold running buffer into top and bottom reservoirs. Check for leaks. There is no need to pre-run the gel.
6. Prepare PCR samples. Add an equal volume of formamide dye to samples (for 100 well combs use 4 µL each). Denature samples in PCR machine for 3 min, and then plunge into an ice bath (mixture of ice and water).
7. Flush out any bubbles or bits of gel from between the plates using a plastic pipette (any bits of gel will prevent you loading your samples into the wells). Replace comb, teeth down and very carefully so that the points of the teeth are just into the gel. Load samples. Run gel at a maximum of 6W overnight at 5°C until the light blue dye (Xylene Cyanol FF) has just run to the bottom of the gel.

8. Remove gel from rig. Separate the glass plates using a plastic wedge. Place small glass plate with bound gel in grey box. Add *Fixer* (in fume hood) and place on the shaker for 30 min.
9. Tip *Fixer* carefully back into container. Wash gel with distilled water for 15 min until "greasy" look disappears, drain and add *Silverstain* to gel. Shake for 30 min.
10. Prepare your chilled *Developer* by adding formaldehyde and sodium thiosulphate. Add distilled water to another grey box and set aside.
11. Lower the gel into the water and agitate for about 6 sec. Drain gel quickly and then lower into the *Developer*.
12. Agitate the gel in the *Developer* and stop the reaction by adding the *Fixer* when the bands are clear. Wash gel in water for at least 5 min to remove sodium carbonate. Leave gel to dry.
13. When dry expose to duplicating film or scan.

2.1.7 KASPar genotyping system

Unless otherwise stated, KASPar assays were performed using a standard recipe and protocol:

PCR recipe:

Template	10-20 ng
Primer mix	0.07 µL
2x V4 KASPar mix	2.43 µL
ddH_2O	fill up to total reaction volume of 5 µL

PCR protocol:

Hotstart	95°C for 15 min
Number of cycles	10 cycles
Denaturation	95°C for 20 sec
Touchdown (-1°C per cycle)	65°C for 25 sec
Number of cycles	30 cycles
Denaturation	95°C for 10 sec
Annealing	57°C for 60 sec
Store	16°C forever

KASPar amplicons are usually smaller than 120 bp and require no extension step in the PCR protocol. 384-well sample plates (Cat. No. 04729749001, Roche Diagnostics) were read on a Roche Lightcycler® II 480. Fluorescence was detected at ambient temperature (20-25°C; RAMP speed 0.05°C per s) with four detection steps per °C. Additional amplification cycles (usually five to ten) were applied if the signature genotyping groups had not formed after the initial amplification. Data analysis was performed manually using the inbuilt Roche Lightcycler® 480 software (Version 1.50.39).

2.1.8 Subcloning of PCR amplicons

PCR products were run on an agarose gel to determine size and intensity and thus correct amplification. Bands of desired size were excised under UV light with a razor blade, trying to minimize exposure time of bands to UV light. PCR products were recovered from the gel fragments by using QIAquick Gel Extraction Kit (Qiagen) and their concentration measured afterwards (Implen Nanophotometer). The PCR products were then ligated into pGEM-T vector (Promega) using T4 DNA ligase (NEB) at 4°C overnight.

Ligation recipe:

2X Rapid Ligation Buffer, T4 DNA Ligase	5 µL
pGEM®-T Vector (50 ng)	0.5 µL
PCR product (insert:vector molar ratio 3:1)	X µL*
T4 DNA Ligase (3 Weiss units/µL)	1 µL
Nuclease-free water to a final volume of	10 µL

$$^{*}\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert:vector molar ratio} = \text{ng of insert}$$

The ligation products were transformed into chemically competent *E.coli* DH5α cells (Invitrogen) according to the manufacturer's protocol. The transformed cells were streaked onto LB Agar plates containing 50 µg/mL Carbenicillin and incubated over night at 37°C.

2.1.9 Annotation of wheat genomic DNA

Genomic sequences of wheat were annotated in a two-step procedure: First, transposable elements (TEs) were annotated using TREP (Matthews et al. 2008) and NCBI databases using the default settings. Target site duplications were annotated where possible. The orientation and similarity scores for the TEs were noted down to allow the orientation of contigs based on TE homology. In the second step, the non-repetitive sequences were used to predict open reading frames (ORFs) using FGENESH (Salamov and Solovyev 2000) and BLASTX (States and Gish 1994). For FGENESH, "wheat" was selected as target organism. Otherwise, default settings were used for both programs.

2.1.10 Screening of TTD140/RSL65 BAC library for positive clones by PCR

The 2D-pools of a BAC library were screened using standard PCR (Materials and Methods 2.1.3) and the PCR products were visualized on a 1.5% agarose gel. Positive samples were noted down and the 1D-pools responsible for the positive signal were screened again by PCR. PCR products were

extracted from the agarose gel and purified using QIAquick Gel Extraction Kit (Qiagen, Cat no. 28704) and then sequenced (Materials and Methods 2.1.5). Positive clones were picked and streaked out unto LB agarose plates containing 17 µg/mL Chloramphenicol for selection. Individual clones were picked and screened by colony PCR (using 17 µg/mL Chloramphenicol for the master plate). A single positive clone was used to inoculate a 2 mL starter culture (LB + 17 µg/mL Chloramphenicol) and incubated over night at 37°C. The starter culture was then used to inoculate 100 mL of LB + 17 µg/mL Chloramphenicol and grown at 37°C until it reached an optical density (OD) of 600. The bacteria were harvested and plasmid DNA was extracted using QIAGEN Plasmid Midi Kit (Qiagen, Cat no. 12143).

2.1.11 BAC end sequencing (BES)

BAC end sequencing is a variation of the standard Sanger sequencing method (Materials and Methods 2.1.5). A BAC is small in size compared to an entire genome, thus the number of DNA molecules per µL is very high. This allows for direct sequencing of BAC DNA using a single primer, with no prior amplification needed. The TTD140 and RSL65 BAC libraries use pIndigo-BAC-5 as vector, which has a T7-promotor on the left side of the multiple cloning site (MCS) and a M13r-promotor on the right side of the MCS. The recipe and protocol are as follows.

BigDye recipe:

Template	2 µL BAC DNA
5x BigDye sequencing buffer	2 µL
Primer	1 µL (10µM concentration)
BigDye reagent	1 µL
ddH_2O	4 µL

PCR protocol:

Initial denaturation	94°C for 2 min
Number of cycles	80 cycles
Denaturation	94°C for 10 sec
Annealing	55°C for 30 sec
Elongation	60°C for 4 min
Store	16°C forever

2.1.12 Screening of TTD140/RSL65 BAC library for positive clones by radioactive hybridization

A non-repetitive PCR product of approximately 500 bp was amplified and purified via an agarose gel. The probe was then send on dry ice to CNRGV in Toulouse, where it was radioactively labeled and hybridized to a set of six filter membranes carrying ~55,000 clones each. A film was exposed to each

membrane, developed and digitalized. The position and identity of every BAC on each membrane is stored digitally, thus allowing for a quick identification of positive clones.

2.1.13 Sequencing of BACs using Roche 454 or Illumina MiSeq

BACs were sequenced at The Genome Analysis Centre (TGAC) using either Roche 454 or Illumina MiSeq technology. High quality BAC DNA (Materials and Methods 2.1.10) was supplied to TGAC, who prepared the sequencing library: 350 bp paired-end libraries for Roche 454 sequencing; ~175 bp paired end libraries resulting in ~350 bp long pseudo-reads for the MiSeq. The libraries were multiplexed and run on a single 454/MiSeq lane.

2.1.14 BAC assembly parameters

The obtained reads of a sequenced BAC were mapped to the pIndigo-BAC-5 vector sequence, as well as to the *E. coli* genome to remove unwanted contaminants: Usually 15% of reads (sometimes more) would map to these sequences. The unmapped reads were used for a *De-novo* assembly using CLC genomic software (www.clcbio.com). *De-novo* assembly criteria were kept strict with reads having to be 95% similar over 75% of their length in order to be assembled together. These settings usually produced good results in terms of stringency, contig size and number. Stricter settings needed to be applied to BAC 13N10 though, which contained a lot of small multicopy repeats (99% similarity over 75% of read length).

2.2 Results

2.2.1 Previous work

This section presents previous work by James Simmonds, a senior RA in the Uauy lab. The data was acquired prior to the start of my PhD and formed the foundation for all my subsequent work.

2.2.1.1 *Vir* and *Iw1* are identical loci

In a previous study on a wheat doubled-haploid population a locus causing non-glaucousness was mapped to the distal end of the short arm of chromosome 2B (Simmonds et al. 2008). This locus was named *Viridescence* (*Vir*) because of the bright green colour of *Vir* plants. The map location coincides with the already mentioned *Inhibitor of Wax 1* (*Iw1*) locus, a dominant gene causing non-glaucousness. To determine whether the two loci are identical, the non-glaucous parent of the doubled-haploid population, Shamrock, was crossed to Shango and six other elite UK winter wheat varieties (Alchemy, Einstein, Hereward, Malacca, Robigus and Xi19; Material and Methods 2.1.1). The F_1 plants from all crosses uniformly displayed a non-glaucous phenotype and the F_2 progeny of all seven lines segregated 3:1 for non-glaucous to glaucous phenotypes, thus strongly supporting the hypothesis that *Vir* and *Iw1* are identical (Table 2.1).

Table 2.1: **Phenotypic segregation in F_2 progeny of crosses with Shamrock**
Seven glaucous winter wheat cultivars have been crossed to the non-glaucous cultivar Shamrock. The F_2 progeny of all lines were phenotyped. Chi-square tests show that the obtained ratio of non-glaucous to glaucous progeny is not significantly different to the expected ratio. Thus, the null hypothesis that all F_2 progeny show a 3:1 segregation ratio for non-glaucous to glaucous phenotype could not be disproven. Note that the table continues on the next page.

F_2 population	Shango x Shamrock	
	Observed	Expected
Non-glaucous	69	66.75
Glaucous	20	22.25
Total	89	
Chi-square value / p-value	0.3	0.58

F_2 population	Alchemy x Shamrock		Einstein x Shamrock		Hereward x Shamrock	
	Observed	Expected	Observed	Expected	Observed	Expected
Non-glaucous	40	39.75	42	40.5	51	53.25
Glaucous	13	13.25	12	13.5	20	17.75
Total	53		54		71	
Chi-square value / p-value	0.01	0.94	0.22	0.64	0.38	0.54

F_2 population	Malacca x Shamrock		Robigus x Shamrock		Xi19 x Shamrock	
	Observed	Expected	Observed	Expected	Observed	Expected
Non-glaucous	36	37.5	56	57	59	54.75
Glaucous	14	12.5	20	19	14	18.25
Total	50		76		73	
Chi-square value / p-value	0.24	0.62	0.07	0.79	1.32	0.25

2.2.1.2 Anchoring *Iw1* to *Brachypodium* and rice

The gene order in grass genomes is highly conserved despite up to 65 million years between the last common ancestor, differences in haploid chromosome number, high content of repetitive DNA as well as genetic inversions and translocations (Moore et al. 1995, Gale and Devos 1998, Sorrells et al. 2003). This conservation represents a unique tool for marker design in large, unsequenced grass genomes like that of common wheat. By establishing the syntenic relationship between a map location in wheat and a small sequenced grass genome, like *Brachypodium distachyon* (International Brachypodium Initiative 2010) or rice (*Oryza sativa*) (Kawahara et al. 2013), it is possible to extract relevant gene sequences from these small grass genomes and search for homologous sequences in various wheat databases or alternatively to design primers from these sequences to amplify the homologous sequences in wheat. Recently, the amount of publicly available DNA sequences of common wheat increased vastly, first by the release of a low-coverage draft genome of the spring wheat cultivar Chinese Spring (Wilkinson et al. 2012) and later by pre-publication access to survey sequences of all chromosomes of Chinese Spring (International Wheat Genome Sequencing Consortium 2011). However, a fully sequenced, ordered and annotated genome sequence is still not available for common wheat. Thus, anchoring a map location to *Brachypodium* or rice and exploiting synteny between grass species is still a necessity for any map-based cloning strategy.

The map location of the *Vir* locus (and hence *Iw1*) was established previously using simple sequence repeat (SSR) markers (Simmonds et al. 2008), with *Vir* mapping closely to marker *Xgwm614*. Synteny is the conservation of gene order between related species, and since SSRs are not based on genes their position cannot be used to establish synteny. Instead, single-strand conformation polymorphism (SSCP) markers were developed from ten expressed sequence tags (ESTs) (Table 2.2), which have previously been mapped to the distal deletion bin of 2BS (2BS3-0.84–1.00) (Conley et al. 2004). The SSCP markers were mapped unto the existing Shango x Shamrock DH population (87 lines) and the results compared to the gene order in *Brachypodium* and rice. A clear syntenic relationship could be observed between the distal end of wheat 2BS and chromosomes 4 and 5 of rice and *Brachypodium* respectively (Table 2.2).

Table 2.2: Markers used for fine mapping of *Iw1* in Shango x Shamrock and Langdon x TTD140
This is a list of markers used in the fine-mapping of *Iw1*. The syntenic relationship between gene-based markers is shown where possible. ([a]) indicates a marker for which both SSCP and KASPar assays were designed, ([b]) indicates the genetic map established in (Simmonds et al. 2008). Primer sequences can be found in the Appendix (A3).

Marker name	Wheat Sequence	Initial Evidence	*Brachypodium* homologue	*Oryza* homologue	Marker type
JIC001	CJ522609	bin mapped EST	*Bradi5g01880*	*LOC_Os04g02500*	SSCP
JIC002	tplb0034e07	bin mapped EST	*Bradi5g01730*	*LOC_Os04g02730*	SSCP
JIC003	CK167138	bin mapped EST	*Bradi5g01420*	*LOC_Os04g02900*	SSCP
JIC004[a]	TA69057_4565	bin mapped EST	*Bradi5g01410*	*LOC_Os04g02910*	SSCP
			Bradi5g01410	*LOC_Os04g02910*	KASPar
JIC005	TA70616_4565	bin mapped EST	*Bradi5g01280*	*LOC_Os04g04020*	SSCP
JIC006	TA76401_4565	bin mapped EST	*Bradi5g01230*	*LOC_Os04g04254*	SSCP
JIC007[a]	TA95426_4565	bin mapped EST	*Bradi5g01220*	*LOC_Os04g04320*	SSCP
			Bradi5g01220	*LOC_Os04g04320*	KASPar
JIC008	454 5x CS	colinearity	*Bradi5g01210*	*LOC_Os04g04330*	SSCP
JIC009	TA80162_4565	bin mapped EST	*Bradi5g01180*	*LOC_Os04g05010*	SSCP
JIC010[a]	TA105895_4565	colinearity	-	*LOC_Os04g05030*	SSCP
			-	*LOC_Os04g05030*	KASPar
JIC011[a]	tplb0015p16	colinearity	*Bradi5g01160*	-	SSCP
			Bradi5g01160	-	KASPar
JIC012[a]	RFL_Contig2535	bin mapped EST	*Bradi5g01130*	*LOC_Os04g05050*	SSCP
			Bradi5g01130	*LOC_Os04g05050*	KASPar
JIC013	wPt-4453	genetic map[b]	-	-	SSCP
JIC014	TA49863_4565	bin mapped EST	*Bradi5g01020*	*LOC_Os04g10680*	SSCP
JIC015[a]	RFL_Contig1863	Colinearity	*Bradi5g01190*	*LOC_Os04g01240*	SSCP
			Bradi5g01190	*LOC_Os04g01240*	KASPar
			Bradi5g01190	*LOC_Os04g01240*	KASPar
JIC016	2BS contig 5198726	Barley contig 46434	*Bradi3g18920*	*LOC_Os08g14620*	presence / absence
Xgwm614		genetic map[b]	-	-	SSR
Xwmc25		genetic map[b]	-	-	SSR

Five additional SSCP markers could be developed from collinear genes in *Brachypodium* and rice (Table 2.2). Using these markers *Iw1* was mapped to a ~2.3-cM interval between markers JIC007 and JIC012, with markers JIC009, JIC010 and JIC011 completely linked to the *Iw1* locus (Figure 2.1).

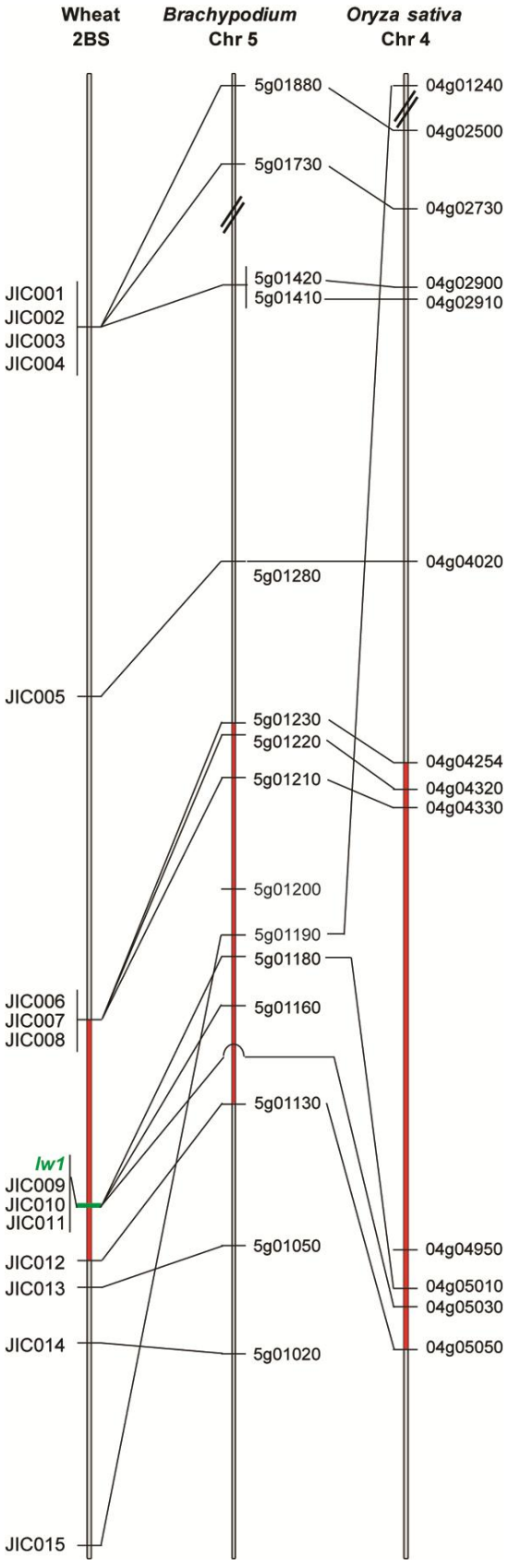


Figure 2.1: 2BS map of wheat anchored on *Brachypodium* and rice

SSCP markers were used to establish synteny between wheat 2BS, *Brachypodium* and rice. Gene order is generally well conserved between the three species in this region. The 2.3-cM *Iw1* interval and the corresponding *Brachypodium* and rice regions are represented by red lines. The putative position of *Iw1* is indicated.

The mapping data also revealed a recent gene duplication and translocation in *Brachypodium*; genes *Bradi5g02380* and *Bradi5g02390* have been duplicated in *Brachypodium* and these paralogous copies (*Bradi5g01200* and *Bradi5g01190*; >95% similarity to *Bradi5g02380* and *Bradi5g02390*) have been translocated to a position between genes *Bradi5g01210* and *Bradi5g01180*. This duplication seems to be limited to these two genes only as the flanking genes *Bradi5g02370* and *Bradi5g02400* both exist as single-copies in *Brachypodium*. In the DH population JIC015 (*Bradi5g01190*) mapped proximal to JIC009 (*Bradi5g01180*) (Figure 2.1).

The results from this section suggest that *Iw1* is very likely identical to the previously mapped *Vir* locus. The locus segregates in a clear Mendelian fashion in seven different test crosses. Development of markers from genes (SSCP markers) allowed anchoring of the map location to *Brachypodium* and rice, two grass species with small, sequenced genomes. Furthermore, additional markers could be developed by exploiting the existing collinearity between these three grass species. This data represents a robust framework to start a map-based cloning approach.

2.2.2 Conversion of SSCP markers into KASPar markers for high-throughput recombinant screen

Recently a new low-cost high-throughput genotyping system was developed by KBiosciences Ltd (now LGC Genomics). This genotyping system distinguishes single nucleotide polymorphisms (SNPs), but can also be used with small insertions/deletions (InDels). The principle behind the method is a variation of polymerase chain reaction (Materials and Methods 2.1.7).

If the KASPar genotyping platform could be adapted to work on wheat it would speed up the genetic mapping of *Iw1* tremendously. In the Shango x Shamrock DH population *Iw1* was mapped between the two SSCP markers JIC007 and JIC0012 (2.3-cM), while the increased yield and delayed senescence traits were mapped between SSCP markers JIC004 and JIC015 (~11-cM). Converting these four SSCP markers into KASPar markers would allow screening for recombinants for all three traits.

To simplify this task group 2 chromosomes of winter wheat variety Chinese Spring were sorted (>95% purity) by Jaroslav Doležels group using a flow cytometry approach (Vrána et al. 2000). SSCP primers for the four genes *Ta_1410*, *Ta_1220*, *Ta_1130* and *Ta_1190* were tested for genome specificity on these chromosome suspensions (Table 2.2). Amplicons of non-genome-specific primers were Sanger sequenced and genome-specific primers were designed from these. No genome-specific primers could be developed for *Ta_1190*, so instead non-genome-specific PCR products of Shango and Shamrock were subcloned into competent *E. coli* DH5α cells (Invitrogen Catalogue Number 18265-

017) (Materials and Methods 2.1.8). Sixteen clones per genotype were Sanger sequenced to obtain amplicons from all three genomes. Sequences from each marker were aligned using BioEdit software (Wessler et al. 1995) and analysed for the presence of SNPs between the three genomes as well as between the B-genome copies of Shango and Shamrock.

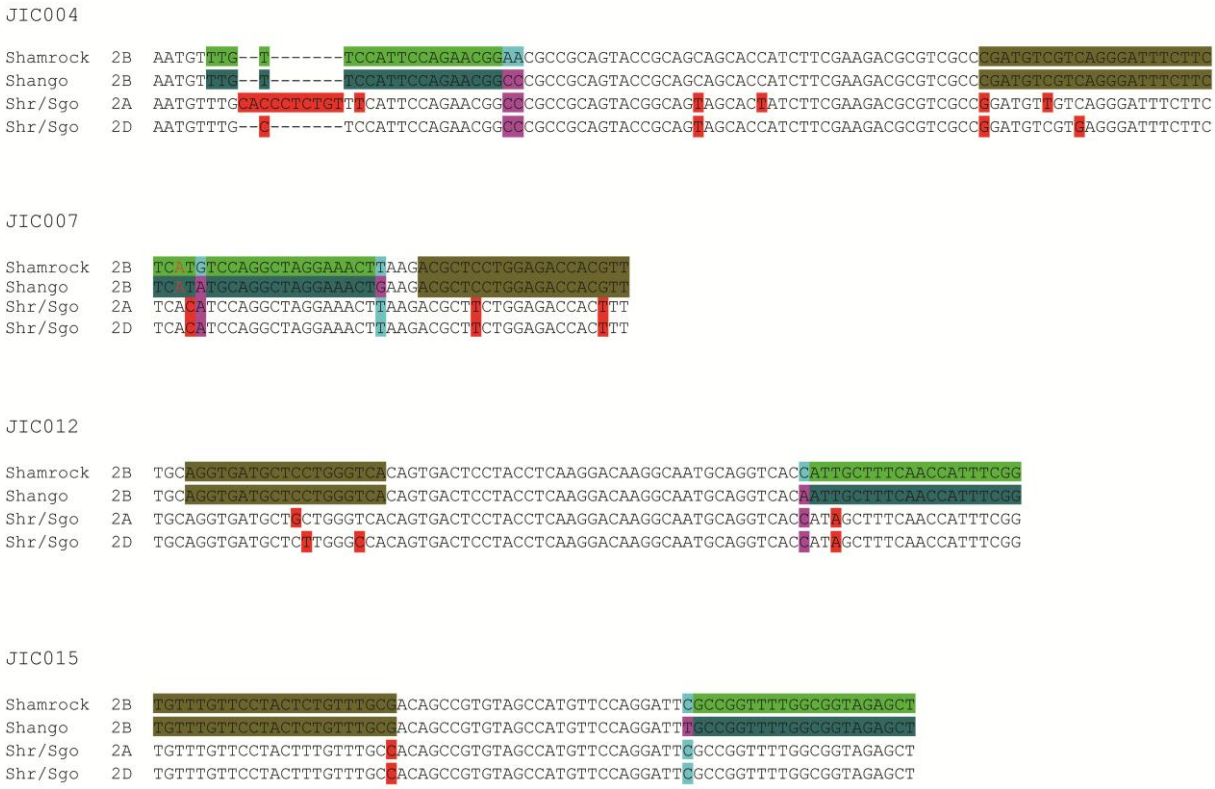


Figure 2.2 KASPar markers used for *Iw1* recombinant screen

Chromosome 2A, 2B and 2D sequence of Shango and Shamrock used to design KASPar assays JIC004, JIC007, JIC012, and JIC015. The allele-specific primers are shown in light and dark green highlight, whereas the common primer is in brown. Allelic SNPs are in blue and purple highlight, whereas homoeologous SNPs between the B and A/D genomes are highlighted red. The VIC and FAM tails of the allele specific primers are not shown.

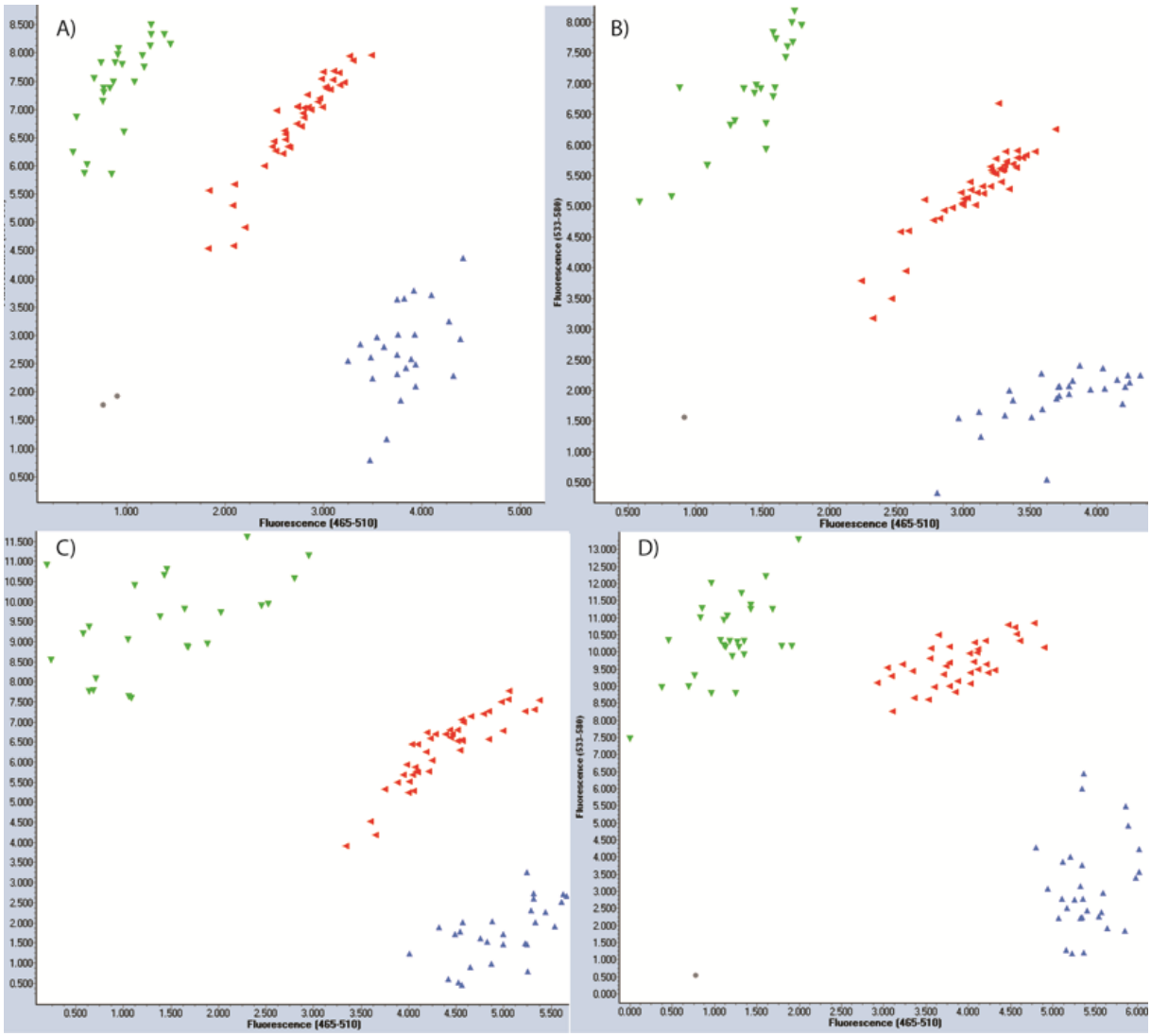


Figure 2.3 KASPar markers scores

Representative scores of a 96 samples from the *Iw1* recombinant screen using KASPar assays JIC004 (A), JIC007 (B), JIC012 (C), and JIC015 (D). The X and Y axis represent the relative fluorescence of the FAM and VIC fluorophores, respectively.

After obtaining sequences for genomic copies of the four genes from both genotypes, KASPar markers were designed based on allelic SNPs between the parental genotypes and homoeologous polymorphisms where possible (Figure 2.2). These markers were then tested on twenty lines of the Shango x Shamrock DH population. KASPar genotyping scores were compared with previously obtained SSCP genotyping scores and B-genome-specific KASPar markers were obtained for all four genes (Table 2.2).

2.2.3 Genetic mapping of *Iw1* in hexaploid wheat

As was mentioned before, the *Inhibitor of wax 1* (*Iw1*) had been mapped to a ~2.3-cM region on the short arm of chromosome 2B. To fine-map the *Iw1* locus 2,350 F_2 plants of a Shango x Shamrock cross were screened for recombinants with four KASPar markers (previous section, Figure 2.2 and 2.3). For 2,073 plants good genotyping calls were obtained and 43 recombinants were discovered between markers JIC007 and JIC012. The recombinants were transferred to new pots and genotyped again using new DNA extractions. Out of the 43 recombinants only 36 showed the same genotyping scores as before and were kept, while the seven false positive lines were discarded.

A further 2,350 Shango x Shamrock F_2 plants screened for recombinants, but initially only with two KASPar markers (JIC007 and JIC012). Good genotyping calls were obtained for 2,304 plants, resulting in the identification of 60 recombinants. The recombinants were transferred to new pots and genotyped with all four KASPar markers using new DNA extractions. Out of the 60 initial recombinants only 34 proved to be true recombinants between JIC007 and JIC012.

The 70 recombinants were genotyped with four additional SSCP markers (JIC005, JIC009, JIC010 and JIC011) showing that the majority of lines (56) have a recombination event between JIC007 and JIC009 (0.64-cM distance), while the remaining 14 lines have a recombination event between the three linked markers JIC009/010/011 and JIC012 (0.16-cM distance). The phenotype of the 70 recombinants, which is a dominant morphological marker in itself, agreed in all cases with the genotyping scores of JIC009/010/011. The three markers JIC009, JIC010 and JIC011 remained completely linked between each other and with the phenotype, even after analysing 4,377 plants (8,754 gametes).

DNA was extracted from twelve F_3 seeds of each recombinant line and genotyped with five KASPar markers (JIC004, JIC007, JIC010, JIC011 and JIC012) and one SSCP marker (JIC009). Two homozygous recombinants were selected from each set of twelve, one as the main recombinant line to be used for all future experiments and the other one as a backup. The phenotypes of all homozygous recombinants matched the genotyping scores for JIC009/010/011.

Table 2.3: Recombinants obtained from the screen of Shango x Shamrock F_2 plants

This table summarizes the results of the large recombinant screen in a Shango x Shamrock F_2 mapping population, from which 70 recombinants were obtained that can be distinguished into six types. Note that markers obtained after the initiation of the physical maps are also included here ([a]).

Marker	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
JIC004	Shr	Shr	Shr	Sgo	Sgo	Sgo
JIC005	Shr	Shr	Shr	Sgo	Sgo	Sgo
JIC006	Shr	Shr	Shr	Sgo	Sgo	Sgo
JIC007	Shr	Shr	Shr	Sgo	Sgo	Sgo
JIC008	Shr	Shr	Shr	Sgo	Sgo	Sgo
JIC016	Shr	Shr	Sgo	Shr	Sgo	Sgo
JIC033[a]	Shr	Shr	Sgo	Shr	Sgo	Sgo
JIC035[a]	Shr	Shr	Sgo	Shr	Sgo	Sgo
JIC009	Shr	Sgo	Sgo	Shr	Shr	Sgo
JIC023[a]	Shr	Sgo	Sgo	Shr	Shr	Sgo
Iw1	non-glaucous	glaucous	glaucous	non-glaucous	non-glaucous	glaucous
JIC010	Shr	Sgo	Sgo	Shr	Shr	Sgo
JIC011	Shr	Sgo	Sgo	Shr	Shr	Sgo
JIC012	Sgo	Sgo	Sgo	Shr	Shr	Shr
JIC015	Sgo	Sgo	Sgo	Shr	Shr	Shr
No. of recombinants	7	24	4	12	16	7

2.2.4 Genetic mapping of *Iw1* in tetraploid wheat

At the same time that the hexaploid Shango x Shamrock F_2 population was screened for recombinants, James Simmonds started to map five 2BS-specific SSCP markers (JIC004, JIC008, JIC010, JIC012 and JIC015) on 94 F_2 lines of a cross between Langdon and TTD140 (Materials and Methods 2.1.1). One of the 94 lines (LDNxTTD-69) had a recombination event between the *Iw1* phenotype (morphological marker) and the two linked markers JIC010/011. This recombination event was confirmed in the F_3 progeny of LDNxTTD-69, thus showing that the *Iw1* locus is more closely linked to JIC009 but not to JIC010/011 (Figure 2.4). This has implications for the 70 recombinants from the hexaploid mapping population. The 14 lines with recombination events between JIC010/011 and JIC012 (Type 1 and Type 6; Table 2.3) have essentially become parental lines, as their recombination events are outside of the *Iw1* interval.

To obtain more recombinants inside the *Iw1* interval all remaining F_2 lines of the Langdon x TTD140 population were screened with two KASPar markers (JIC007 and JIC011). Out of the initial 94 F_2 lines screened above, 39 lines showed heterozygosity between JIC004 and JIC015. Their F_3 progeny can thus still recombine in the *Iw1* interval, essentially behaving like a regular F_2 population in this

genomic region. In total, 850 plants (1,700 gametes) were screened and 29 recombinants identified, indicating a genetic distance between JIC007 and JIC011 of 1.7-cM. The majority of plants (25) had a recombination event between JIC007 and JIC009, while four plants had a recombination event between JIC010/011 and JIC009. One of these four plants died prematurely before setting seed, reducing the number of recombinants to 28. Homozygous recombinants were obtained for all 28 lines as described for the Shango x Shamrock recombinants (Section 2.2.3). This work was performed in coordination with Dr. Adrian Turner.

Table 2.4: Recombinants obtained from the screen of Langdon x TTD140 F_2 plants

This table summarizes the results of the large recombinant screen in a Langdon x TTD140 F_2 mapping population, from which 28 recombinants were obtained that can be distinguished into six types. Note that markers obtained after the initiation of the physical maps are also included here ([a]).

Marker	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
JIC007	TTD140	TTD140	Lgd	Lgd	Lgd	Lgd
JIC016	Lgd	TTD140	Lgd	Lgd	Lgd	TTD140
JIC034 [a]	Lgd	TTD140	Lgd	Lgd	Lgd	TTD140
JIC035 [a]	Lgd	TTD140	Lgd	Lgd	Lgd	TTD140
JIC036 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC024 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC023 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC033 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC009	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
lw1	glaucous	non-glaucous	glaucous	glaucous	non-glaucous	non-glaucous
JIC022 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC020 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC019 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC018 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC017 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC032 [a]	Lgd	TTD140	Lgd	Lgd	TTD140	TTD140
JIC031 [a]	Lgd	TTD140	Lgd	TTD140	TTD140	TTD140
JIC011	Lgd	Lgd	TTD140	TTD140	TTD140	TTD140
JIC010	Lgd	Lgd	TTD140	TTD140	TTD140	TTD140
No. of recombinants	11	1	1	1	1	13

2.2.5 A contig from *Hordeum vulgare* cultivar Morex

Barley (*Hordeum vulgare*) diverged from wheat approximately twelve million years ago (Chalupska et al. 2008) and thus is very similar to wheat in genetic terms, making it a good model species for the Triticeae tribe. In an international effort, a gene-rich scaffold of the barley genome (Mayer et al. 2011) was created and Dr. Nils Stein gave us pre-publication access to barley contig 46434, which has an approximate size of 685 kb. The Bacterial Artificial Chromosomes (BACs) making up the contig were arranged using fingerprinted contig (FPC) data (Soderlund et al. 1997) and contained the *Ta_1180* gene. All BACs were annotated by hand (Material and Methods 2.1.9) to identify putative barley genes within contig 46434 (Table 2.5). This analysis was followed by a BLASTN search against the wheat survey sequence (International Wheat Genome Sequencing Consortium 2011) database using all putative barley genes as query.

Primer pairs were designed for the MU-homology domain gene, the Cox2 genes and the Transferase gene. No sequences with clear homology to Cytochrome P450 and Haemolysin-III related genes were detected in the wheat survey sequence. Primer pairs in the Cox2 and Transferase genes only amplified products in Shango and Shamrock, but not in Langdon or TTD140, which suggests that they are present in the D-genome only. Primer pairs for the Mu-Homology domain gene (JIC016) amplified only in Shango and Langdon, but when tested on the Shango x Shamrock and Langdon x TTD140 F_3 homozygous recombinants, JIC016 mapped inside the *Iw1* interval between JIC007 and JIC009 (Table 2.3 and Table 2.4). Out of the 84 recombinants 40 lines had a recombination event between JIC007 and JIC016, essentially reducing the number of recombinant lines by half and thus reducing the size of the *Iw1* interval to 0.42-cM. However, this new marker is a dominant presence/absence marker, i.e. one parent lacks the sequence that the primer amplifies. Both Shamrock and TTD140 lack this sequence, so the new marker cannot be used in the construction of the physical map of the *Iw1* locus.

Table 2.5: List of BACs allocated to contig 46434 and the genes detected within

This table lists the BACs allocated to barley contig 46434 by FPC. BACs were annotated by hand and genes were detected in almost every BAC. Primers were designed for several of the genes and the MU-homology domain gene appears to have a homologue on the short arm of chromosome 2B. Note that the 2BS copy of *Ta_1180* in wheat is split into two contigs.

BAC name	genes	Wheat homologue / marker
HVVMRXALLmA0018H15	none	-
HVVMRXALLhA0692E09	MU-homology domain gene	2BS contig 5198726 / (JIC016)
HVVMRXALLhA0692E09	Cytochrome P450	-
HVVMRXALLhA0692E09	Cytochrome P450	-
HVVMRXALLhA0036B08	Multicopper oxidase Cox2 genes	2DS contigs 5340809 and 5344521
HVVMRXALLeA0092B21	Transferase gene	2DS contig 5294824
HVVMRXALLhA0085J09	Multicopper oxidase Cox2 genes	2DS contigs 5340809 and 5344521
HVVMRXALLhA0612N23	CBS domain gene (*Hv_1180*)	2BS 5007649 and 5182000 (*Ta_1180*) / JIC009
HVVMRXALLrA0131J23	Cytochrome P450	-
HVVMRXALLrA0131J23	Cytochrome P450	-
HVVMRXALLrA0242M21	none	-
HVVMRXALLeA0241J02	Haemolysin III-related	-

Concluding remarks on the genetic mapping process

In total, by screening 5,227 plants (10,454 gametes) from two mapping populations 44 recombinants inside the *Iw1* interval between the markers JIC016 and JIC010 were obtained, representing a genetic distance of 0.42-cM (Figure 2.4). Physical maps are usually initiated after narrowing down the genetic distance to ~0.3-cM or below by screening large numbers of plants from a mapping population in order to reduce the physical size of the map (Konishi et al. 2006, Li et al. 2006). However, there appears to be a lack of recombination inside the *Iw1* interval as even after screening more than 10,000 gametes the genetic distance could not be narrowed down further than 0.42-cM. Out of the 44 recombinants obtained in total, 41 lines had a recombination event between JIC016 and JIC009, which means that 93.2% of all recombination events occurred between JIC016 and JIC009. This indicates that JIC016 is still quite distal from the actual *Iw1* gene, with at least some genes in between allowing for recombination to happen. However, neither *Brachypodium*, rice nor barley have syntenic genes in this region that have not been already exploited for marker development, thus making it impossible to develop markers closer to the *Iw1* interval. In contrast, only three plants had a recombination event between *Iw1*/JIC009 and JIC010/011 and all of these came from the Langdon x TTD140 population. This imbalance between the two flanks of the genetic map could be due to differences in recombination frequencies (so-called hot- and cold spots of recombination)

(Lichten and Goldman 1995, Sidhu and Gill 2005) or due to a deletion within the modern germplasm, thus making recombination in this region impossible. Regardless of which of these is true, it seems that screening more plants for recombination events inside the *Iw1* interval will yield little new information, which is why at this point the physical map of the *Iw1* interval was initiated.

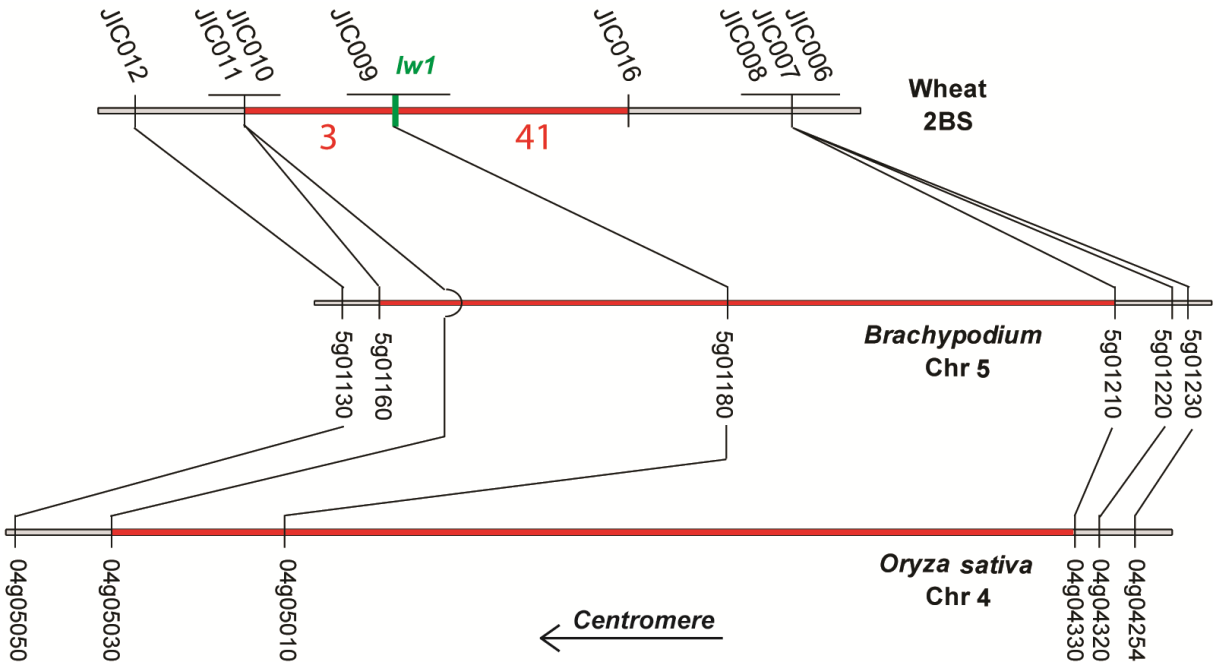


Figure 2.4: Updated genetic map of the *Iw1* interval

After screening 10,454 gametes from two mapping populations, 44 recombinant lines were obtained. The numbers in red represent the number of recombinants between two adjacent markers. The red line represents the *Iw1* interval, which is 0.42-cM in size in wheat. The location of the centromere is indicated.

2.3 Construction of a physical map of the *Iw1* locus

This section will summarize the ongoing and not yet completed effort to create a physical map of the *Iw1* interval using two BAC libraries.

2.3.1 TTD140 BAC library construction

By screening more than 10,000 gametes the *Iw1* interval has been narrowed down to a sub-cM distance (0.42-cM), thus allowing construction of a physical map of this genomic region to start. Since all available bacterial artificial chromosome (BAC) libraries were made from glaucous plants a new BAC library needed to be constructed. The decision was made to create a BAC library from TTD140 rather than Shamrock as it is a tetraploid, thus simplifying the physical map construction and in addition it would represent the first BAC library of a TTD accession, thus making it useful for other

projects in addition to this one. Rather than constructing the library at JIC, dark-grown leaf material of TTD140 was sent to the group of Dr. Hélène Bergès at CNRGV (Toulouse) as they have the necessary skills and equipment for constructing BAC libraries as well as a proven track record of doing so. The library was shipped to JIC at the beginning of January 2012. The library consists of 331,776 gridded clones organized in 864 384-well plates. The average size of each clone is ~120 kb giving the library a size of 39,813,120 kb, which covers every part of the TTD genome (11,000,000 kb) three and a half times on average. DNA from each clone was extracted at CNRGV and the DNA from an entire 384-well plate was pooled into a single well of a 96-well plate, thus the entire library is represented by nine 96-well plates. These so-called one-dimensional pools (1D-pools) were further pooled by row and column into three 96-well plates, thus creating so-called two-dimensional pools (2D-pools) which allow screening of the entire BAC library on a single 384-well plate (Figure 2.5).

1D-pool

	IC1	IC2	IC3	IC4	IC5	IC6	IC7	IC8	IC9	IC10	IC11	IC12
IR1	1	9	17	25	33	41	49	57	65	73	81	89
IR2	2	10	18	26	34	42	50	58	66	74	82	90
IR3	3	11	19	27	35	43	51	59	67	75	83	91
IR4	4	12	20	28	36	44	52	60	68	76	84	92
IR5	5	13	21	29	37	45	53	61	69	77	85	93
IR6	6	14	22	30	38	46	54	62	70	78	86	94
IR7	7	15	23	31	39	47	55	63	71	79	87	95
IR8	8	16	24	32	40	48	56	64	72	80	88	96

IC1: Plate **I**, Column **1**
IR1: Plate **I**, Row **1**

2D-pools

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A		IC7	IIR7		IIIC7	IVR7		VC7	VIR7		VIIC7	VIIIR7		IXC7
B		IC8	IIR8		IIIC8	IVR8		VC8	VIR8		VIIC8	VIIIR8		IXC8
C	IR1	IC9	IIC1	IIIR1	IIIC9	IVC1	VR1	VC9	VIC1	VIIR1	VIIC9	VIIIC1	IXR1	IXC9
D	IR2	IC10	IIC2	IIIR2	IIIC10	IVC2	VR2	VC10	VIC2	VIIR2	VIIC10	VIIIC2	IXR2	IXC10
E	IR3	IC11	IIC3	IIIR3	IIIC11	IVC3	VR3	VC11	VIC3	VIIR3	VIIC11	VIIIC3	IXR3	IXC11
F	IR4	IC12	IIC4	IIIR4	IIIC12	IVC4	VR4	VC12	VIC4	VIIR4	VIIC12	VIIIC4	IXR4	IXC12
G	IR5		IIC5	IIIR5		IVC5	VR5		VIC5	VIIR5		VIIIC5	IXR5	
H	IR6		IIC6	IIIR6		IVC6	VR6		VIC6	VIIR6		VIIIC6	IXR6	
I	IR7		IIC7	IIIR7		IVC7	VR7		VIC7	VIIR7		VIIIC7	IXR7	
J	IR8		IIC8	IIIR8		IVC8	VR8		VIC8	VIIR8		VIIIC8	IXR8	
K	IC1	IIR1	IIC9	IIIC1	IVR1	IVC9	VC1	VIR1	VIC9	VIIC1	VIIIR1	VIIIC9	IXC1	
L	IC2	IIR2	IIC10	IIIC2	IVR2	IVC10	VC2	VIR2	VIC10	VIIC2	VIIIR2	VIIIC10	IXC2	
M	IC3	IIR3	IIC11	IIIC3	IVR3	IVC11	VC3	VIR3	VIC11	VIIC3	VIIIR3	VIIIC11	IXC3	
N	IC4	IIR4	IIC12	IIIC4	IVR4	IVC12	VC4	VIR4	VIC12	VIIC4	VIIIR4	VIIIC12	IXC4	
O	IC5	IIR5		IIIC5	IVR5		VC5	VIR5		VIIC5	VIIIR5		IXC5	
P	IC6	IIR6		IIIC6	IVR6		VC6	VIR6		VIIC6	VIIIR6		IXC6	

Figure 2.5: Pooling strategy for the TTD140 BAC Library
Layout of the pooling strategy for the DNA pools of the TTD140 BAC library. DNA is extracted from all clones of a single 384-well plate and pooled into one well of a 96-well plate (1D-pool). Individual rows of the 1D-pool plate are pooled into a single well of a new 96-well plate and the same is done for individual columns (2D-pools). Thus the entire library is encompassed within 180 wells.

2.3.2 Construction of a physical map of the *Iw1* interval in TTD140

In section 2.2.3 it was mentioned that the recombination frequency at the two ends of the *Iw1* interval is drastically different, with 93.2% of recombination events lying between JIC016 and JIC009. This could indicate a big physical distance between JIC016 and JIC009 with many genes present between them that are non-syntenic with *Brachypodium* or rice. In contrast only three recombination events were discovered between JIC009 and JIC011, indicating a small distance between these genes (Figure 2.4). Thus it was decided to start the construction of the physical map using these two markers.

Table 2.6: List of markers used in physical map construction

This is a list of markers used in the physical map construction in TTD140 and RSL65. ([a]) indicates markers based on non-repetitive sequence, ([b]) indicates markers based on junctions between two repetitive elements, ([c]) indicates markers based on single repetitive elements, ([d]) indicates BACs identified in non-gridded pools at CNRGV Toulouse. Primer sequences can be found in the Appendix (A4). Note that the table continues on the next page.

Marker name	Wheat Sequence	Marker type	TTD140 BACs detected	RSL65 BACs detected
JIC009[a]	TA80162_4565	InDel	321F24, 529K13, 551I08, 784K20	264M08
JIC011[a]	tplb0015p16	InDel	305N15, 370A02, 427A02	91G23
JIC016[a]	2BS contig 5198726	presence / absence	97O06, 305G20	-
JIC017[b]	TTD140 BAC 305N15	presence / absence	190O12, 305N15, 427A02, 774P06	-
JIC018[b]	TTD140 BAC 774P06	presence / absence	13N10, 412O05, 774P06	-
JIC019[b]	TTD140 BAC 13N10	presence / absence	13N10, 23N19, 412O05	-
JIC020[b]	TTD140 BAC 23N19	presence / absence	23N19, 83F21, 158I09, 363M21, 571B20, 802A16, 807A06	-
JIC021[c]	TTD140 BAC 83F21	non-specific	83F21, 170H13	not tested
JIC022[b]	TTD140 BAC 170H13	presence / absence	170H13, 68G17[d]	-
JIC023[b]	TTD140 BAC 321F24	presence / absence	321F24, 577P12, 784K20	-
JIC024[b]	TTD140 BAC 529K13	presence / absence	274M11, 329M10, 478J01, 529K13, 551I08, 636L08	-
JIC025[b]	TTD140 BAC 329M10	presence / absence	329M10	-
JIC026[b]	TTD140 BAC 636L08	presence / absence	274M11, 329M10, 478J01, 636L08	-
JIC027[a]	TTD140 BAC 478J01	presence / absence	274M11, 329M10, 478J01, 529K13, 551I08, 636L08	-

Table 2.6 continued

Marker name	Wheat Sequence	Marker type	TTD140 BACs detected	RSL65 BACs detected
JIC028[b]	TTD140 BAC 551I08	presence / absence	551I08, 32C13[d]	-
JIC029[a]	TTD140 BAC 577P12	presence / absence	321F24, 551I08, 577P12, 784K20	-
JIC030[a]	TTD140 BAC 577P12	presence / absence	577P12, 784K20, 85B11[d]	-
JIC031[a]	RSL65 BAC 91G23	presence / absence	-	91G23
JIC032[b]	2BS contig 5157821	presence / absence	-	1041P07
JIC033[a]	RSL65 BAC 264M08	presence / absence	-	119H03, 264M08, 294I15, 304G05, 618L11, 1326P06
JIC034[a]	RSL65 BAC 97O06	presence / absence	-	97O06, 374D10, 634B01
JIC035[a]	RSL65 BAC 305G20	presence / absence	-	305G20, 784L08, 1031F12, 1150M10
JIC036[a]	2BS contig 5204867	InDel	not tested yet	not tested yet

2.3.2.1 TTD140 physical map starting at JIC011

2.3.2.1.1 JIC011

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC011 (Materials and Methods 2.1.10). Three positive clones (305N15, 370A02, 427A02) were identified by this screen (Table 2.6). BAC end sequences (BES) were obtained from all three clones (Materials and Methods 2.1.11). The T7-end of 305N15 consisted of two exons and an intron from *Ta_5030*, while the M13r-end consisted of a long terminal repeat (LTR) retrotransposon (TREP3530: RLC_Angela_B_consensus-1) that had inserted itself into a different LTR retrotransposon (TREP3457: RLG_Danae_consensus-1).

The junctions between transposable elements (TEs) are virtually unique because of the high turnover rate of the repetitive portion of the wheat genome (Wicker et al. 2003, Dubcovsky and Dvorak 2007), thus making them ideal sites for marker development (Paux et al. 2006). The drawback of these markers is their dominant nature; due to their high specificity, they often only amplify the correct product in the genotype they were designed on (here TTD140), but usually not in other genotypes (e.g. Langdon, Shango or Shamrock). Primer pairs were designed with one primer sitting inside one TE and the other primer inside the adjacent TE (insertion site based polymorphism (ISBP) (Paux et al. 2006)). The primer pairs were tested on the Langdon x TTD140 F_3 homozygous recombinants to verify their specificity. The mapping data revealed three recombination events between JIC011 and

the new marker JIC017 (Table 2.4). This result orientated BAC clone 305N15 relative to the centromere. However, there are no more recombination events left on the proximal side of the physical map, which means that new BAC clones can only be orientated by physical overlap with existing clones (Figure 2.6).

2.3.2.1.2 JIC017

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC017. Four positive BAC clones were identified, two of which were already known (305N15 and 427A02) and two new BAC clones (190O12 and 774P06) (Table 2.6). Primer pairs were designed from the BES of the new clones. Markers from the M13r-end of clone 774P06 yielded PCR products of correct size and sequence from BACs 305N15, 427A02 and 190O12; while a marker from the T7-end of 774P06 (JIC018) did not amplify DNA from any of these three BACs. Marker JIC018 was tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

2.3.2.1.3 JIC018

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC018. Three positive BAC clones were identified, one of which was already known (774P06) and two new BAC clones (13N10 and 412O05) (Table 2.6). Primer pairs were designed from the BES of the new clones. Markers from the T7-end of clone 13N10 yielded PCR products of correct size and sequence from BACs 774P06 and 412O05, while a marker from the M13r-end (JIC019) did not amplify DNA from BAC 774P06. Marker JIC019 was tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

2.3.2.1.4 JIC019

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC019. Three positive BAC clones were identified, two of which were already known (13N10 and 412O05) and the other one was a new BAC clone (23N19) (Table 2.6). Primer pairs were designed from the BES of the new clone. Markers from the T7-end of clone 23N19 yielded PCR products of correct size and sequence from BAC 13N10, while a marker from the T7-end (JIC020) of clone 23N19 did not amplify DNA from BAC 13N10. Marker JIC020 was tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

2.3.2.1.5 JIC020

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC020. Seven positive BAC clones were identified, one of which was already known (23N19) and six new BAC clones (83F21,

158I09, 363M21, 571B20, 802A16 and 807A06) (Table 2.6). Primer pairs were designed from the BES of the new clones. Markers from the M13r-end of clone 83F21 yielded PCR products of correct size and sequence from all seven BACs (23N19, 83F21, 158I09, 363M21, 571B20, 802A16 and 807A06), while a marker from the T7-end (JIC021) did not amplify DNA from any of these seven BACs. However, marker JIC021, and all other primer pairs from the T7-end of BAC 83F21, could not be tested on the Langdon x TTD140 F_3 homozygous recombinants, as the primer pairs lacked specificity: The T7-end of BAC 83F21 consisted of a single LTR retrotransposon (TREP3254: RLG_Sakura_10k23-4), which did not allow for the development of specific primers.

2.3.2.1.6 JIC021

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC021, despite its lack of specificity. Two positive clones were identified, one of which was already known (83F21) and the other was a new BAC clone (170H13) (Table 2.6). Another plate was identified several times in the 2D-pools (plate 144) but even after screening the plate numerous times no positive clone could be detected. Primer pairs were designed from the BES of the new clone 170H13. Markers from the M13r-end of clone 170H13 yielded PCR products of correct size and sequence from BAC 83F21, while primer pairs from the T7-end of clone 170H13 did not amplify DNA from BAC 83F21. The T7-end of BAC 170H13 consisted of a single LTR retrotransposon (TREP3457: RLG_Danae_consensus-1), which did not allow for the development of specific primers. Furthermore, the various T7 primer pairs only detected clone 170H13 and plate 144 when used to screen the TTD140 2D-pools, but no other BACs. After sequencing BAC 170H13 (see below) specific primer pairs were designed as close as possible to the T7-end of BAC 170H13. Marker JIC022 was tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

2.3.2.1.7 JIC022

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC022, but no new clones were detected (Table 2.6). A probe made from JIC022 was sent to the lab of Dr. Hélène Bergès at CNRGV in Toulouse where the TTD140 BAC library was created. This probe was then used to detect clones on filter membranes of the BAC library by radioactive hybridization (Materials and Methods 2.1.12). No new clones were detected even after several different screens. As a last option non-gridded pools were prepared from the original BAC library plugs and screened by PCR. After screening several plates of pools a positive clone (68G17) was detected using primer JIC022 (Table 2.6). The BES of 68G17 consisted of single retroelements on both ends and could not be used to design new specific markers. Sequencing of this BAC clone is underway.

2.3.2.1.8 Sequencing of the minimum tiling path (MTP)

In the seven chromosome walking steps starting from JIC011 sixteen BACs were identified (Figure 2.6). A minimum tiling path (MTP) of seven clones was selected and sequenced using Roche 454 (305N15, 774P06 and 13N10) and Illumina MiSeq (23N19, 68G17, 83F21 and 170H13) (Materials and Methods 2.1.13). The sequencing data for each BAC was filtered from contaminants and then assembled using CLC Genomics (Materials and Methods 2.1.14). Sequence reads from adjacent BACs were mapped against the assemblies to verify their position in the MTP and to elucidate by how much they overlap with each other. This process verified that the supposed MTP obtained by chromosome walking was correct, with every BAC in the MTP overlapping several thousand base pairs with its neighbouring BACs with more than 99.9% similarity. The size of the MTP is approximately 560.9 kb (excluding 68G17).

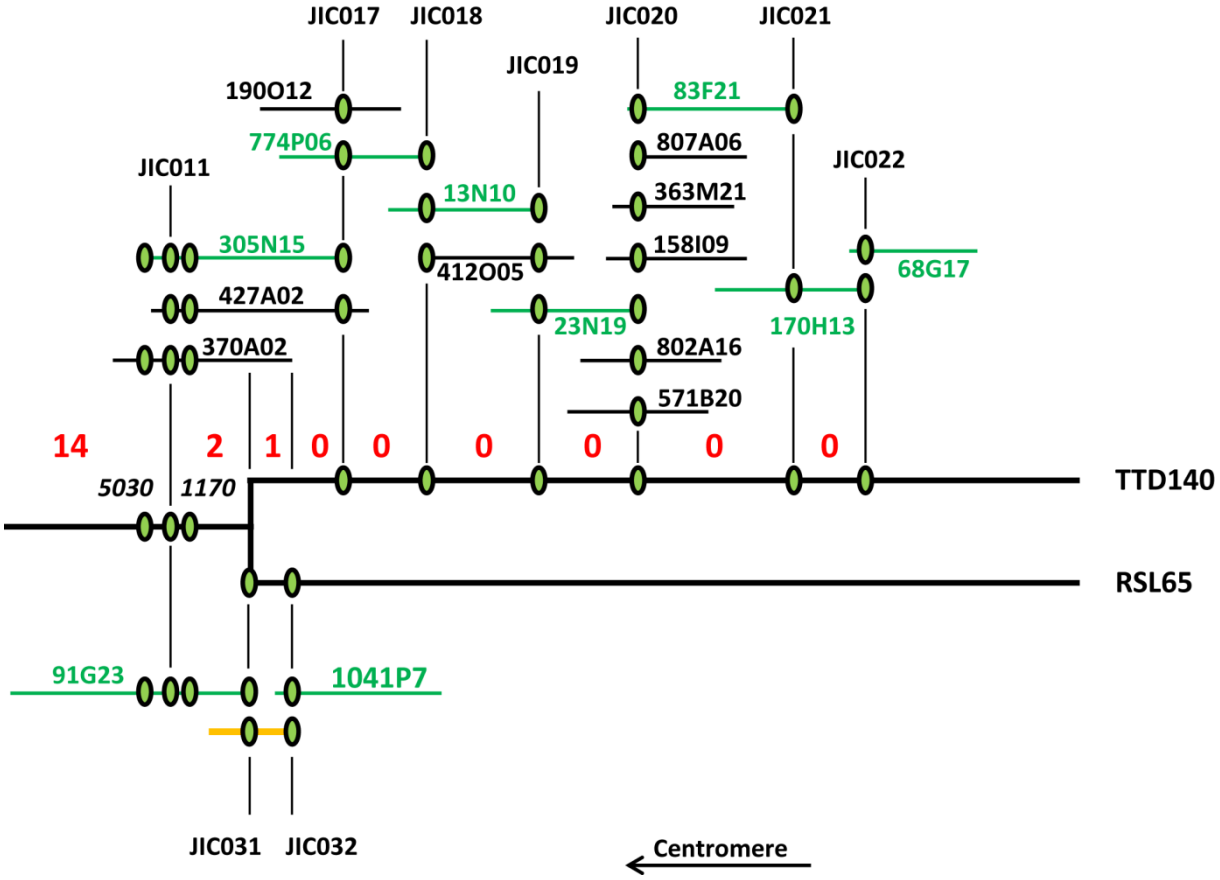


Figure 2.6: Physical map starting at JIC011

This diagram depicts the progress on the physical maps of TTD140 and RSL65 on the proximal side of the *Iw1* interval. BACs labelled green have been sequenced. The orange line represents the 2BS contig 5157821. The numbers in red represent the number of recombinants from both mapping populations between two adjacent markers. The bifurcation of the map represents an hypothesized set of two different haplotypes in TTD140 and RSL65. The location of the centromere is indicated. Note that BAC overlap and scale is only approximate for the purpose of clarity.

2.3.2.2 TTD140 physical map starting at JIC009

2.3.2.2.1 JIC009

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC009. Four positive clones (321F24, 529K13, 551I08 and 784K20) were identified by this screen (Table 2.6; Figure 2.7). BAC end sequences (BES) were obtained from all four clones and from these primer pairs were designed to orientate the BACs relative to each other (Table 2.6). A specific marker from the M13r-end of BAC 321F24 (JIC023) overlapped only with BAC 784K20 and was tested on the Langdon x TTD140 F_3 homozygous recombinants as well as the Shango x Shamrock F_3 homozygous recombinants and proved to correctly map to 2BS. Similarly, a marker from the T7-end of BAC 529K13 (JIC024) overlapped only with BAC 551I08 and was mapped on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

2.3.2.2.2 JIC024

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC024. Six positive BAC clones were identified, two of which were already known (529K13 and 551I08) and four new BAC clones (274M11, 329M10, 478J01, 636L08) (Table 2.6). Primer pairs were designed from the BES of the new clones. A marker from the T7-end of BAC 329M10 (JIC025) did not amplify DNA from any of the other BACs. Apart from that the new markers (JIC026, JIC027) did not enable us to orientate the new BACs relative to each other, as they amplified PCR products of correct size in all four BACs. Sequencing of the PCR products showed that they consisted of mixed traces, indicating that more than one product was amplified. Subcloning of PCR products showed that most markers amplified at least two different sequences, if not more.

After sequencing BACs 529K13, 551I08, 274M11 and 329M10 it became clear that BACs 529K13 and 551I08 overlap extensively with greater than 99.9% similarity and likewise BACs 274M11 and 329M10 overlapped extensively with greater than 99.9% similarity. However, the similarity between seemingly overlapping contigs of BACs 529K13/551I08 and 274M11/329M10 was only 98%, which indicated a recent duplication event. Specific PCR markers were developed using the fully sequenced BACs, but no new clones adjacent to 274M11 or 329M10 could be detected. Probes made from these specific primers (Appendix, A5) were sent to the lab of Dr. Hélène Bergès at CNRGV in Toulouse where the TTD140 BAC library was created. These probes were then used to detect clones on filter membranes of the BAC library by radioactive hybridization (Materials and Methods 2.1.12). No new clones were detected even after several different screens. As a last option non-gridded pools were prepared from the original BAC library plugs and screened by PCR, but again no new clones could be detected.

2.3.2.2.3 JIC028

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC028, but no new clones were detected (Table 2.6). A probe made from JIC028 was sent to the lab of Dr. Hélène Bergès at CNRGV in Toulouse where the TTD140 BAC library was created. This probe was then used to detect clones on filter membranes of the BAC library by radioactive hybridization (Materials and Methods 2.1.12). No new clones were detected even after several different screens. As a last option non-gridded pools were prepared from the original BAC library plugs and screened by PCR. After screening several plates of pools a positive clone (32C13) was detected using primer JIC028 (Table 2.6). The BES of 32C13 consisted of single retroelements on both ends and could not be used to design new specific markers. Sequencing is currently underway for this BAC which will hopefully help resolve the putative duplication in this region.

2.3.2.2.4 JIC023

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC023. Three positive BAC clones were identified, two of which were already known (321F24 and 784K20) and the other one was a new BAC clone (577P12) (Table 2.6). Primer pairs were designed from the BES of the new clone. A specific marker from the M13r-end of clone 577P12 (JIC029) yielded PCR products of correct size and sequence from BACs 321F24 and 784K20, while no specific markers could be developed from the T7-end because it consisted of a single LTR retrotransposon (TREP3529: RLC_Angela_A_consensus-1). After sequencing BAC 577P12 specific primer pairs were designed close to T7-end of the BAC. Marker JIC030 was tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

2.3.2.2.5 JIC030

The TTD140 BAC library 2D-pools were screened by PCR with marker JIC030, but no new clones were detected (Table 2.6). A probe made from JIC030 was sent to the lab of Dr. Hélène Bergès at CNRGV in Toulouse where the TTD140 BAC library was created. This probe was then used to detect clones on filter membranes of the BAC library by radioactive hybridization (Materials and Methods 2.1.12). No new clones were detected even after several different screens. As a last option non-gridded pools were prepared from the original BAC library plugs and screened by PCR. After screening several plates of pools a positive clone (85B11) was detected using primer JIC030 (Table 2.6). The BES of 85B11 consisted of single TEs on both ends and could not be used to design new specific markers. However, the BES of the T7-end is identical (identities: 907/907) to BACs 784K20 and 577P12, suggesting that this is a genuine connection.

2.3.2.2.6 Sequencing of the minimum tiling path (MTP)

In the five chromosome walking steps starting from JIC009 eleven BACs were identified (Figure 2.7). A MTP of seven clones was selected and sequenced using Roche 454 (274M11, 329M20, 551I08 and 784K20) and Illumina MiSeq (32C13 and 577P12). The sequencing data for each BAC was filtered from contaminants and then assembled using CLC Genomics (Materials and Methods 2.1.14). Sequence reads from adjacent BACs were mapped against the assemblies to verify their position in the MTP and to elucidate by how much they overlap with each other. This process verified that the supposed MTP obtained by chromosome walking was correct, with every BAC in the MTP overlapping several kb with its neighbouring BACs on both sides with more than 99.9% similarity. The size of the MTP is approximately 381.7 kb (excluding 32C13).

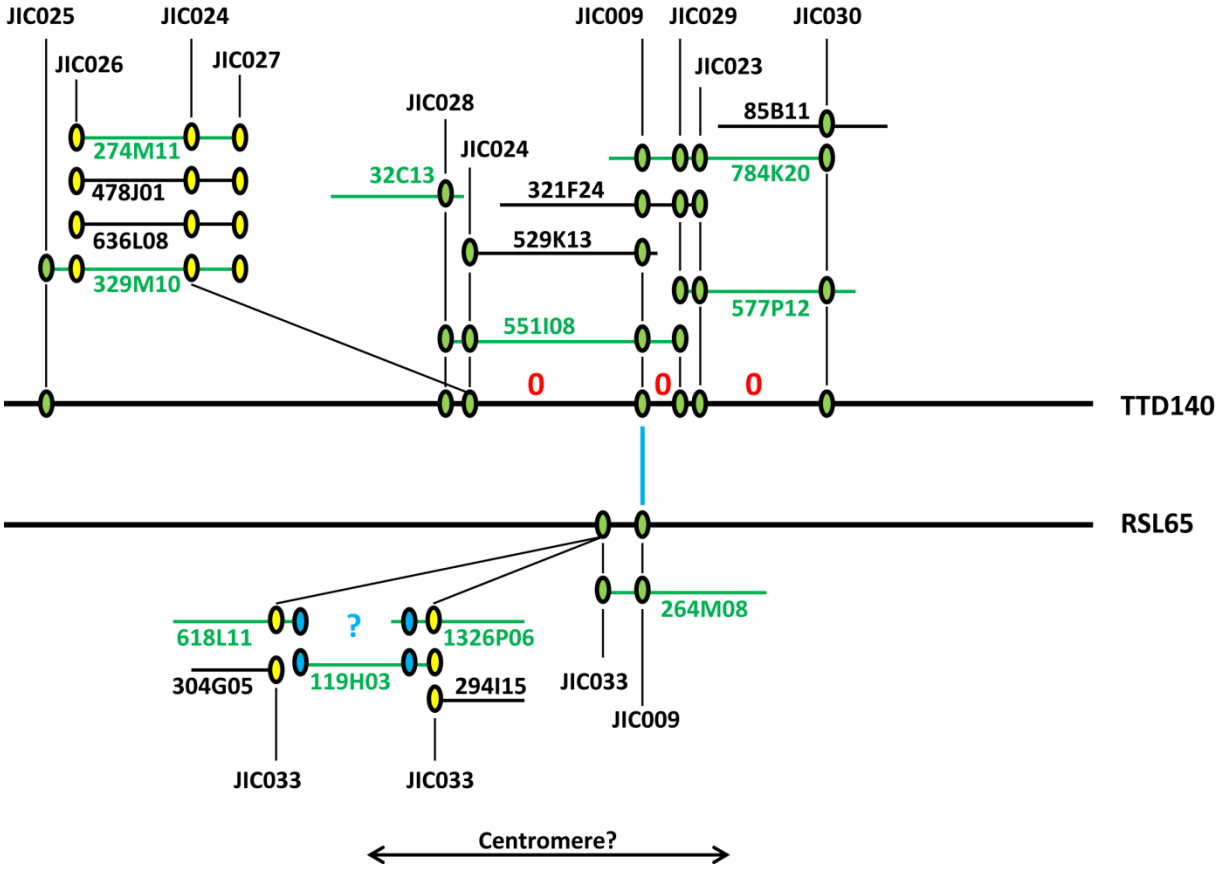


Figure 2.7: Physical map starting at JIC009

This cartoon depicts the progress on the physical maps of TTD140 and RSL65 on the central part of the *Iw1* interval. BACs labelled green have been sequenced. The dotted blue line represents the paralogous *Ta_1180* locus. The numbers in red represent the number of recombinants from both mapping populations between two adjacent markers. Note that the location of the centromere is unknown due to lack of recombination inside this region. The blue filled dots represent a putative overlap based on BES. The yellow filled dots represent a possible duplication. The parallel lines represent a hypothesized set of two different haplotypes in TTD140 and RSL65. Note that BAC overlap is only approximate for the purpose of clarity.

2.3.2.3 Concluding remarks on the TTD140 physical map

Before initiating the construction of the TTD140 physical map an extensive recombinant screen was performed to reduce the size of the *Iw1* interval to sub-cM size. The physical map was initiated at a marker flanking the target region proximally (JIC011) and at a marker that is completely linked to *Iw1* (JIC009). Construction of the distal end of the physical map was not started, as 81 recombinants were detected between the *Iw1* interval and the next closest marker (JIC007).

After numerous chromosome walking steps, two MTPs with a combined size of >1 Mb were obtained. Unfortunately, due to technical errors from the sequencing service provider a number of BACs could not be sequenced in time to include them in this thesis (68G17, 32C13, 85B11). These two MTPs do not overlap and the 0.42-cM interval is also not completely covered yet (Figures 2.6 and 2.7). Genomic sequence contigs from both MTPs were used for BLASTN searches in various wheat databases, including the Chinese Spring wheat survey sequence database. Apart from contigs containing known genes like *Ta_1160* and *Ta_1180*, these database queries did not yield positive hits. This suggests a possible deletion of the *Iw1* interval in Langdon and other modern day wheat varieties, which could not be uncovered by the genetic mapping but becomes apparent at the physical map stage. We further hypothesise that because of this deletion the *Iw1* interval in RSL65 will be small compared to TTD140, thus facilitating the creation of the RSL65 physical map. This new physical map could yield new marker information that would allow us to start construction of the distal end of the TTD140 physical map.

2.3.3 Construction of a physical map of the *Iw1* interval in RSL65

A copy of the 5x RSL65 BAC library (Cenci et al. 2003) is present at JIC. This library is a based on a recombinant substitution line of durum wheat cultivar Langdon, which carries a 30-cM introgression from a wild emmer wheat accession on the short arm of chromosome 6B. It has been used successfully to clone the *GPC-B1* gene (Uauy and University of California 2007). The library has been pooled in the same way as the TTD140 BAC library (Section 2.3.1). Here we want to use this library to create a physical map of the *Iw1* interval from a glaucous wheat cultivar. We hypothesise that *Iw1* was deleted in RSL65 and other glaucous germplasm. This is based on the results of the TTD140 physical map, where we discovered a large segment of genomic DNA (>1 Mb), most of which was not represented in public databases including the Chinese Spring survey sequence. If our hypothesis is correct, the *Iw1* interval in RSL65 will be small compared to TTD140. We hope that this map will help us in our efforts of constructing a complete physical map of the *Iw1* interval in TTD140.

2.3.3.1 RSL65 physical map starting at JIC011

2.3.3.1.1 JIC011

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC011 (Materials and Methods 2.1.10). One positive clone (91G23) was identified (Table 2.6) and its BES was obtained. The BES consisted of unique sequence and was thus used for a BLASTN query of the wheat survey sequence database. The T7-end was 99.8% (identities: 841/843) identical to 2BS contig 5226543, while the M13r-end was 99.9% (identities: 785/786) identical to 2BS contig 5157821. The latter contig contained the two genes *Ta_1160* and *Ta_1170* and based on the orientation of these genes relative to the BES of 91G23 it was concluded that the M13r-end points towards gene *Ta_1180* (compare with section 2.3.2.1 JIC011). Several specific primer pairs were designed from the M13r-end of 91G23 and used to screen the RSL65 2D-pools, but no positive clone could be identified.

The 2BS contig 5157821 from the wheat survey sequence extends 6,447 bp from the M13r-end of BAC 91G23. Several markers were designed along this part of contig 5157821, tested on the Langdon x TTD140 F_3 homozygous recombinants and JIC032 proved to correctly map to 2BS within the *Iw1* interval (Figure 2.6; Table 2.4).

2.3.3.1.2 JIC032

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC032. One positive BAC clone (1041P07) was identified (Table 2.6) and its BES obtained. The M13r-end of BAC 1041P07 was 99.4% (identities: 960/966) identical to 2BS contig 5157821, while the T7-end did not match with the contig's sequence. The T7-end consists of a single LTR retrotransposon (TREP: 1435; RLC_Inga_AY268139-1) and could not be used to design specific primer pairs.

2.3.3.1.3 Sequencing of the minimum tiling path (MTP)

In the two chromosome walking steps starting from JIC011 two BACs and one wheat survey sequence contig were identified (Figure 2.6). The BACs were sequenced using Illumina MiSeq. The sequencing data for each BAC was filtered from contaminants and then assembled using CLC Genomics (Materials and Methods 2.1.14). Sequence reads from the BACs were mapped against the assemblies to verify their position in the MTP and to elucidate by how much they overlap with each other. The two BACs 91G23 and 1041P07 do not overlap with each other: There is a gap of 2,397 bp between them, which is spanned by the WSS 2BS contig 5157821, resulting in a continuous MTP (Figure 2.6). The size of the MTP is approximately 205 kb (excluding 1041P07).

2.3.3.2 RSL65 physical map starting at JIC009

2.3.3.2.1 JIC009

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC009. One positive clone (264M08) was identified (Table 2.6) and its BES was obtained. The entire BAC was sequenced shortly after its identification using Illumina MiSeq (Materials and Methods 2.1.13). The T7-end of the BAC was assembled into a ~7 kb contig, from which a specific primer pair was designed (JIC033), tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4). The M13r-end of the BAC was only assembled into a ~1 kb contig, which consisted of a single DNA transposon (TREP: 771; DTC_Clifford_TREP771-1) and was not suitable to design specific primer pairs from. The assemblies from both ends of BAC 264M08 were queried against the wheat survey sequence database, but either did not yield significant hits (M13r-end) or only matched contigs that did not extend beyond the assembled contig (T7-end).

2.3.3.2.2 JIC033

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC033. Six positive BAC clones were identified, one of which was known already (264M08) and five of which were new clones (119H03, 294I15, 304G05, 618L11, 1326P06) (Table 2.6). BES was obtained for all new clones and primer pairs were designed from them. The BES of the five clones (119H03, 294I15, 304G05, 618L11 and 1326P06) did either only overlap in parts with 264M08 (119H03, 304G05 and 1326P06) or showed a consistent pattern of SNPs with 264M08 but not between themselves (294I15 and 618L11). Thus it seems that actually none of these BACs truly overlaps with 264M08. Furthermore, it seems that 618L11 and 1326P06, although identified using the same marker (JIC033) are not overlapping either. Only two small contigs (5 kb and 7 kb) overlap with <99% similarity, which suggests recently duplicated sequences. BAC 119H03 could be a link between 618L11 and 1326P06 as its M13r BES is almost identical with 618L11 (identities: 1012/1013), while its T7 BES is identical with 1326P06 (identities: 930/930) (Figure 2.7). New markers need to be developed to screen the RSL65 library again and recover overlapping BACs on both sides of BAC 264M08.

2.3.3.2.3 Sequencing of the minimum tiling path (MTP)

In the two chromosome walking steps starting from JIC009 six BACs were identified (Figure 2.7). A minimum tiling path (MTP) of four clones was selected and sequenced using Illumina MiSeq (119H03, 264M08, 618L11 and 1326P06). The sequencing data for each BAC was filtered from contaminants and then assembled using CLC Genomics (Materials and Methods 2.1.14). Sequence reads from adjacent BACs (apart from 119H03) were mapped against the assemblies to verify their position in the MTP and to elucidate by how much they overlap with each other. This process revealed that two

separate contigs were created during the chromosome walking steps. The three BACs 119H03, 618L11 and 1326P06 form one MTP, which is not yet physically connected to BAC 264M08, but markers from both MTPs map to the *Iw1* interval on 2BS (Figure 2.7).

2.3.3.3 RSL65 physical map starting at JIC016

2.3.3.3.1 JIC016

The marker JIC016 is based on a gene that is predicted to encode a MU-homology domain gene (MU). This gene was absent from the syntenic intervals in *Brachypodium* and rice and was discovered on a barley contig. This MU-gene appears to be missing from TTD140 and Shamrock, but is present in Langdon and Shango.

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC016. Two positive BAC clones (97O06 and 305G20) were identified (Table 2.6) and their BES was obtained, which consisted for both BACs on both ends of single LTR retrotransposons. All BES were queried against the wheat survey sequence database using BLASTN. The T7-end of BAC 97O06 matched a 2BS contig from the wheat survey sequence (contig 5247170) with 100% (identities: 937/937), while the next best hit was only 88.1% similar to it. Contig 5247170 extended 4,212 bp further away from BAC 97O06. A marker was designed (JIC034), tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4).

Likewise, the M13r-end of BAC 305G20 hit 2BS contig 5194426 with 99.8% similarity (identities: 999/1001), while the next best hit was only 88.8% similar to it. Contig 5194426 extended 700 bp further into BAC 305G20. A marker was designed (JIC035), tested on the Langdon x TTD140 and Shango x Shamrock F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.3 and Table 2.4).

The mapping data did not uncover any recombination events, which prevented the orientation of the two BACs relative to the physical map. As such, chromosome walking steps needed to be performed for both ends (Figure 2.8).

2.3.3.3.2 JIC034

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC034. Four positive BAC clones were identified, two of which were already known (97O06 and 305G20) and two new clones (374D10 and 634B1) (Table 2.6). BES was obtained from the new clones and used to design primer pairs. The

primers for both BACs were tested on the Langdon x TTD140 F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.4). However, no recombination event was detected with any of the markers.

2.3.3.3.3 JIC035

The RSL65 BAC library 2D-pools were screened by PCR with marker JIC035. Four positive BAC clones were identified, one of which was already known (305G20) and three new clones (784L08, 1031F12 and 1150M10) (Table 2.6). BES was obtained from the new clones and used to design primer pairs. The T7-end of BAC 1031F12 was 99.9% similar (identities: 885/886) to 2BS contig 5204867, while the next best hit was only 81.2% similar to it. Contig 5204867 extended 21,494 bp further away from BAC 1031F12. A marker was designed (JIC036), tested on the Langdon x TTD140 and Shango x Shamrock F_3 homozygous recombinants and proved to correctly map to 2BS within the *Iw1* interval (Table 2.3 and Table 2.4). A recombination event was detected in the Langdon x TTD140 population that orientated this end of the physical map towards *Ta_1180* (Figure 2.8). Furthermore, JIC036 is an InDel marker and can thus be used to screen the TTD140 BAC library for BAC clones on the distal end of the TTD140 physical map.

2.3.3.3.4 Sequencing of the minimum tiling path (MTP)

In the three chromosome walking steps starting from JIC016 seven BACs and two wheat survey sequence contigs were identified (Figure 2.8). A minimum tiling path (MTP) of three clones was selected and sent for sequencing using Illumina MiSeq (97O06, 305G20, 1031F12). Unfortunately, due to technical difficulties none of these BACs could be sequenced in time to include them in this thesis.

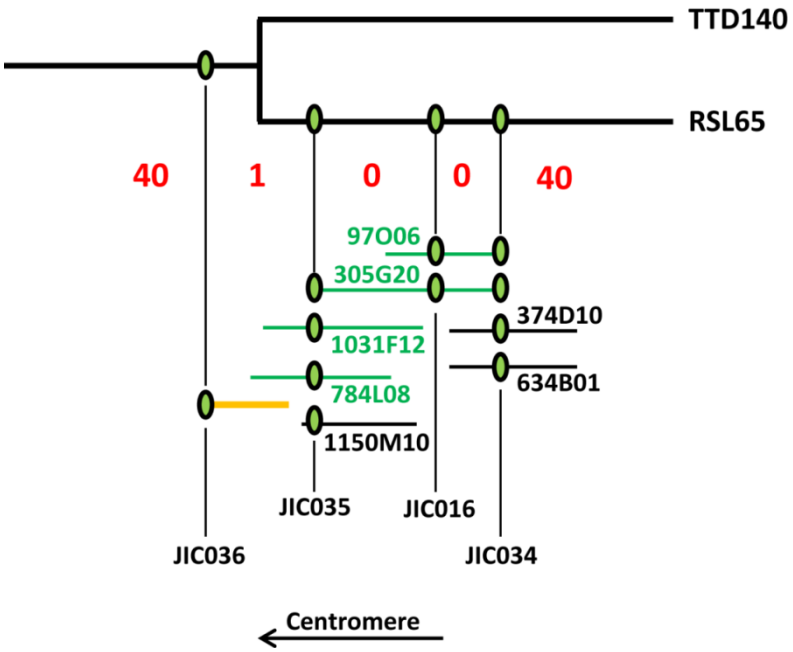


Figure 2.8: Physical map starting at JIC016
This cartoon depicts the progress on the physical maps of RSL65 on the distal side of the *Iw1* interval. BACs labelled green have been sequenced. The orange line represents the 2BS contig 5204867. The numbers in red represent the number of recombinants from both mapping populations between two adjacent markers. The bifurcation of the map represents a hypothesized set of two different haplotypes in TTD140 and RSL65. The location of the centromere is indicated. Note that BAC overlap is only approximate for the purpose of clarity.

2.3.3.4 Concluding remarks on the RSL65 physical map

The construction of the RSL65 physical map was started to elucidate the extent of a possible deletion event inside the *Iw1* interval in glaucous cultivar Langdon. This work is still ongoing, but is facilitated by the readily available wheat survey sequence, which at several points helped to extend the physical map and thus identify new BAC clones or detect recombination events in our mapping populations. This suggests that the *Iw1* interval is well conserved between RSL65 (cultivar Langdon) and cultivar Chinese Spring. Although not yet complete the construction of the RSL65 map already helped to discover a new marker, which is also present in TTD140, on the distal side of the *Iw1* interval. So far the closest marker on the distal side of the *Iw1* interval was JIC016, but the gene that it was designed on is not present in TTD140. As a consequence JIC016 could not be used to screen the TTD140 BAC library. The new marker JIC036 is co-dominant between Langdon and TTD140 and allows us now to start the construction of the distal end of the TTD140 physical map.

2.3.4 Preliminary comparison of the TTD140 and RSL65 physical maps

A comparison between the TTD140 and RSL65 physical maps, although both are still unfinished, will help our understanding of the *Iw1* interval. Here, we take a preliminary look (maps unfinished yet) at the sequence similarity between TTD140 and RSL65 MTPs to try and understand the apparent presence of two different haplotypes at the *Iw1* interval.

2.3.4.1 JIC011 MTP

The MTP from RSL65 is short compared to its TTD140 equivalent, but already a stark difference in similarity can be observed. Three closely linked genes, *Ta_5030*, *Ta_1160* and *Ta_1170*, are present in both MTPs. The coding sequence (CDS) of gene *Ta_5030*, which encodes a protein of unknown function, is highly conserved (99.8%; identities: 1,782/1,786) between TTD140 and the wheat survey sequence contig 2BS 5186802 (i.e. Chinese Spring), but less conserved between TTD140/Chinese Spring and RSL65 (96.9%; identities: 1,731/1,786). The gene *Ta_1160*, which also encodes a protein of unknown function, was used as proximal flank in both physical maps. The CDS of *Ta_1160* was 99.4% (identities: 1,215/1,222) similar between TTD140 and RSL65 and 100% similar (identities: 1,222/1,222) between TTD140 and the wheat survey sequence contig 5157821.

In contrast the CDS of *Ta_1170*, which is predicted to encode a NB-ARC domain protein, is only 98.5% similar between TTD140 and RSL65/Chinese Spring. This difference in sequence mostly stems from a 25-bp deletion in *TTD140_1170*, which leads to a premature STOP codon in the predicted protein after 192 amino acids (aa). In contrast, *RSL65_1170* is 99.7% similar to the corresponding open reading frame (ORF) in Chinese Spring.

The similarity of the non-repetitive region following *Ta_1170* drops to 96% (identities: 2,523/2,628) between TTD140 and RSL65 and 94.8% (identities: 2,737/2,887) between TTD140 and Chinese Spring and decreases to ~85% at the end of Chinese Spring contig 2BS 5157821. At this position a gene is predicted in TTD140, RSL65 and Chinese Spring by both FGENESH and BLASTX algorithms, which here will be called *Ta_Lectin-Pkc* after the domains it is predicted to encode. The similarity at this gene is 99.7% between RSL65 and Chinese Spring, but only ~88% between TTD140 and RSL65/Chinese Spring. This suggests that a split into two different haplotypes has occurred at this point between TTD140 and RSL65/Chinese Spring (Figure 2.9).

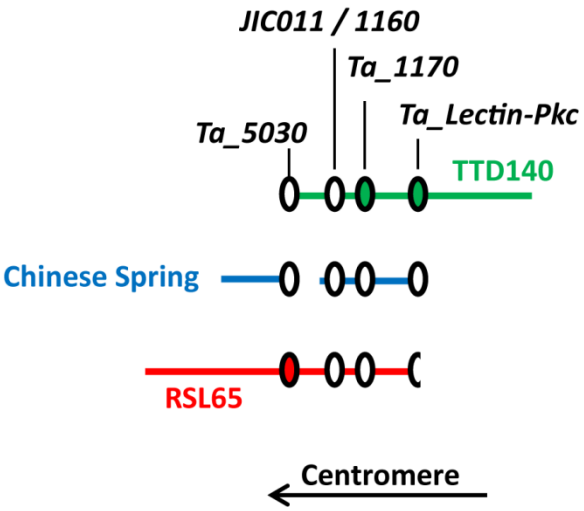


Figure 2.9: Comparison between the physical maps of TTD140 and RSL65 starting at JIC011
The physical maps of TTD140 (green line) and RSL65 (red line) are compared with contigs from the wheat survey sequence (blues lines labeled Chinese Spring). The ovals represent the four predicted genes. Gene *Ta_5030* from RSL65 is different from TTD140/Chinese Spring (red oval), while gene *Ta_1160* is equally conserved among the three genotypes (white oval). Genes *Ta_1170* and *Ta_Lectin-Pkc* are different between TTD140 and RSL65/Chinese Spring (green ovals). *Ta_Lectin-Pkc* is only partly present on the RSL65 map, because BAC 91G23 ends at that point (half oval).

2.3.4.2 JIC009 MTP

The MTPs from TTD140 and RSL65 surrounding JIC009 are very different from each other in terms of gene content and sequence similarity. Apart from *Ta_1180*, the MTPs from TTD140 and RSL65 have no genes in common: Predicted genes encoding Cytochrome P450 enzymes differ in exon-intron structure and sequence and cannot be homologous to each other. Overlaps in sequence outside of repetitive elements are small in size (less than 1 kb) and low in similarity (92% and below). The only real commonality between the MTPs was the gene *Ta_1180*, which encodes a protein with a CBS_pair domain. The CDS of *Ta_1180* is only 97.1% similar (identities: 647/666) between TTD140 and RSL65, while it is 99.4% (identities: 671/675) similar between RSL65 and Chinese Spring. The genomic sequence of *Ta_1180* is even less conserved between TTD140 and RSL65/Chinese Spring: While the first two exons, including the intron between them, is 98% similar (identities: 579/591) between TTD140 and RSL65/Chinese Spring, the last three exons, including the introns, are only 82.2% similar (identities: 722/878) between them. Therefore it is difficult to establish with absolute certainty if the TTD140 and RSL65/Chinese Spring genes are indeed homoeologues.

In addition, a DNA transposon (DHH_Helios_42j2-1) with >7 kb in size has inserted 2,324 bp downstream of the second exon of *Ta_1180* in RSL65 and 2BS contig 5007649, but not in TTD140. The intron between exons 2 and 3 is only 992 bp in size in TTD140, suggesting that the DNA transposon inserted some extra sequence into the intron.

While it is possible that the MTPs were unknowingly constructed in opposite directions, as there are no recombinants at this position to orientate the two maps relative to the centromere, the differences are too numerous to assume that the MTPs are of the same haplotype.

Interestingly, a number of genes predicted in the RSL65 MTPs are very similar to the genes found in barley 2HS contig 46434. A Haemolysin-III encoding gene is 91.5% similar (identities: 910/994) between RSL65 and barley, while one of the four Cytochrome P450 encoding genes in barley is 89.3% similar (identities: 1,391/1,557) to the ones identified in the RSL65 MTPs. A difference between the RSL65 MTPs and barley is the absence of DNA transposon DHH_Helios_42j2-1 inside *Hv_1180*, suggesting that the insertion occurred after the divergence from barley.

These results suggest that two different haplotypes exist in TTD140 and RSL65/Chinese Spring and highlights the divergence between modern cultivars (RSL65 and Chinese Spring) and wild emmer accession TTD140.

2.3.4.3 JIC016 MTP

The gene that marker JIC016 is based on is absent from TTD140/Shamrock, which has prevented the use of JIC016 to screen the TTD140 BAC library. As a consequence this part of the physical map is only represented by a MTP from RSL65. However, in the process of constructing the RSL65 physical map at the distal end of the *Iw1* interval a co-dominant marker (JIC036) was discovered. This enables us to start construction of the TTD140 physical map on the distal side of the *Iw1* interval. Forty recombinants are left between JIC036 and JIC009, suggesting that there are a number of genes shared between Langdon and TTD140 in this in region. However, based on the results from the proximal and central parts of the physical map we would expect Langdon and TTD140 to diverge at some point.

2.4 Discussion

This chapter described the progress that was made in trying to clone the dominant inhibitor of glaucous *Iw1*. A lot of work has been done prior to the start of my PhD to demonstrate that a single locus from a wild emmer wheat accession is responsible for inhibiting glaucousness in seven different UK winter wheat varieties. Furthermore, this locus had been mapped to a discrete location on the short arm of chromosome 2B between markers JIC007 and JIC012. In addition, the distal end of wheat chromosome 2B had been anchored to the *Brachypodium* and rice genomes, which allowed the design of new markers based on syntenic genes in this interval. Lastly, a big F_2 population of the cross Shango x Shamrock had been prepared as well as a smaller F_2 population of the cross Langdon x TTD140.

It is at this point that I started my PhD project by converting existing SSCP markers into KASPar markers. This new type of marker promised low-cost and high-throughput screening of wheat populations and over the course of my PhD established itself as the new standard genotyping method across several wheat laboratories worldwide. Screening 4,377 F_2 plants (8,754 gametes) of the hexaploid Shango x Shamrock population was accomplished in approximately two weeks (including the preparation of DNA), whereas this would have taken up at least half a year using SSCP and SSR markers. This exemplifies the power that KASPar markers offer to a geneticist. From this screen 70 recombinants were identified. Shortly afterwards 850 plants (1,700 gametes) of the tetraploid Langdon x TTD140 population were screened with KASPar markers as well. From these, 28 recombinants were discovered, including three lines that showed recombination events between two previously linked markers, thus rendering 14 recombinants from the hexaploid Shango x Shamrock population redundant.

These recombinants were developed to the F_3 generation and homozygous lines were obtained. The homozygous recombinants were screened with all available markers, completely exhausting the syntenic genes from *Brachypodium* and rice. At this time an international consortium had been working on a draft sequence of barley and we were granted pre-publication access to a barley contig with syntenic relationship to the *Iw1* interval. This contig was annotated by hand and the predicted genes from this contig were queried against various wheat databases. Primer pairs were designed from homologous wheat contigs, one of which mapped to the short arm of wheat chromosome 2B. This marker was positioned much closer to the *Iw1* interval than the previously used marker JIC008 and reduced the number of recombinants on the distal side of the interval by half. In total, 5,227 plants (10,454 gametes) from two mapping populations were screened and 44 recombinants inside the *Iw1* interval were obtained, which equals a genetic distance of 0.42-cM.

The distribution of recombinants is heavily skewed though, with only three recombination events on the proximal side of the *Iw1* interval and the remaining 41 on the distal side. This could be due to a recombination cold spot on the proximal side or a recombination hotspot on the distal side of the *Iw1* interval. Alternatively, it could be that this genomic region contains many genes that are unique to wheat and absent from the syntenic intervals in barley, *Brachypodium* and rice.

There is also an imbalance of recombination events between the two mapping populations. Out of the 41 recombination events on the distal side of the *Iw1* interval, 40 are present in the lines obtained from the Shango x Shamrock population. Likewise, the three recombination events on the proximal side of the *Iw1* interval are exclusive to progeny from the Langdon x TTD140 population. The imbalance on the distal side could be due to the vastly different amount of plants screened between the two populations (4,377 vs. 850), but the number of screened plants does not explain the imbalance of recombination events on the proximal side.

A possible explanation is gene content; genetic recombination can only take place between genes and if the parents of the two mapping populations differ in gene content this could result in the observed imbalance. Shango and Langdon are both modern domesticated cultivars, which means that they are very similar to each other. In contrast the other two parents are likely to be very different. The wild emmer wheat accession used in the cross yielding cultivar Shamrock is different from TTD140. The non-domesticated wild emmer wheat accessions in the Fertile Crescent display a huge array of genetic diversity (Luo et al. 2007), so it is possible to assume a different genetic diversity and thus gene content for the wild emmer wheat introgression in Shamrock and the accession TTD140. Another explanation could be again hot- and cold spots of recombination present in one but not the other population (Faris et al. 2000, Sidhu and Gill 2005).

Screening large numbers of plants is vital for any positional cloning effort (Konishi et al. 2006, Li et al. 2006), especially for species with large genomes. A high number of plants screened will increase the chance to find recombination events on either side of the locus of interest and also in close proximity to it. Formulas have been derived that estimate the number of plants or gametes needed to achieve this (Durrett et al. 2002). These formulas depend on estimates of the physical to genetic distance (Mb/cM) in order to be accurate. Cytogenetic studies have shown an exponential decrease in recombination with distance from the telomere (Dvořák et al. 1984, Lukaszewski and Curtis 1993) and a rapid increase in the physical to genetic distance ratios closer to the centromere. The genome-wide estimate is 3 Mb/cM (Bennett and Smith 1991), but smaller ratios than that are common for distal chromosomal regions (Stein et al. 2000). The genetic map of *Iw1* has a size of 0.42-cM, which would translate to a physical size of approximately 480 kb to 1,200 kb, if the physical to genetic distance is constant along the interval.

In previous map-based cloning projects a wide range of physical to genetic distance ratios has been discovered. In some cases this ratio could be as little as 0.09 Mb/cM (*mlo* (Büschges et al. 1997)), whereas in other instances the ratio was much bigger than the genome-wide estimate by Bennett and Smith, ranging from 7.4 Mb/cM (*Vrs1* (Komatsuda et al. 2007)) to 13.5 Mb/cM (*VRN1* (Yan et al. 2003)). However, in all these cases the size of the genetic interval of the gene of interest had been small, ranging from 0.04 cM (*VRN1*) to 0.36 cM (*mlo*). This was achieved by screening large numbers of gametes (4,044 for *mlo;* 6,190 for *VRN1*; 9,831 for *Vrs1*) for recombinants. The final size of the physical amps from which these genes were identified ranged from only 30 kb for *mlo* to 550 kb for *VRN1*.

We have screened 10,454 gametes from two mapping populations with the size of the genetic interval of *Iw1* being 0.42-cM. We have also created several BAC contigs with a combined size of 1,200 kb and are still adding more clones to these contigs. This highlights the complexity of the *Iw1* interval, which we hypothesise is caused by the existence of different haplotypes of the *Iw1* interval. The comparison of physical maps between TTD140 and RSL65, although yet limited, suggests a high level of divergence in terms of gene content and general sequence similarity. Interestingly, we found that the RSL65 map showed great similarity in terms of gene content with a contig from the glaucous barley cultivar Morex. Because of the lack of dominant inhibitors of glaucousness in barley, we hypothesise that an ancestral "glaucous" haplotype existed, which diverged at least in some wild emmer species and creating *Iw1* in the process. The existence of at least two haplotypes would explain the lack of recombination in the *Iw1* interval and it would also mask the interval's genetic size. This is reminiscent of another map-based cloning project which suffered from lack of recombination and allelic diversity. The *Ph1* locus was mapped to a physical interval of 2.5 Mb and was eventually validated using a set of deletion mutants (Griffiths et al. 2006). This highlights the fact that although positional cloning in polyploid species with large genomes is a challenge, it is feasible. In addition, new genetic resources like the wheat survey sequence (International Wheat Genome Sequencing Consortium 2011) or the barley genome (Mayer et al. 2011) have already proven to be valuable for map-based cloning and will facilitate future cloning efforts.

In summary, *Iw1* was fine-mapped to a sub-cM position on the short arm of wheat chromosome 2B. Physical maps for TTD140 and RSL65 were started, but not completed yet. Based on a limited comparison of the two physical maps it appears as though the *Iw1* interval is quite different between TTD140 and RSL65. We hypothesise the existence of two haplotypes of the *Iw1* interval, which prevent recombination inside the interval. We further hypothesise that the "glaucous" haplotype is the ancestral form as there is no evidence for the existence of *Iw1* so far in rye, barley or other grass

species. Instead *Iw1* has possibly emerged within a new “non-glaucous” haplotype in wild emmer wheat. It is possible to assume the existence of a glaucous accession within wild emmer wheat, which carries the “non-glaucous” haplotype or at least a haplotype that is more closely related to it. We have acquired a set of >100 wild emmer wheat accessions and are currently screening the glaucous ones for similarity to the “non-glaucous” haplotype. This would create an alternative for cloning *Iw1*.

Chapter 3 Validation of *Iw1* candidate genes

Introduction

The genomes of members of the Triticeae tribe range in size from >3 Mb to 8 Mb (Bennett and Smith 1976, Arumuganathan and Earle 1991, Doležel et al. 1998, Jakob et al. 2004) in diploid species and to multiples of that in polyploid species. These immense genomes exist because of the exceedingly high presence of repetitive DNA, accounting for 76, 80 and 92% of the genome size in *Hordeum vulgare*, *Triticum monococcum* and *Secale cereale* respectively (Flavell et al. 1974, Bennetzen et al. 1998, Shirasu et al. 2000, Lagudah et al. 2001). This high amount of repetitive elements is associated with a very dynamic genome, with transposable elements (TEs) causing an altered expression of genes (Kloeckener-Gruissem and Freeling 1995, Fu et al. 2005, Zhang and Saier 2009), gene deletions (Harberd et al. 1987, Chopra et al. 1999), gene duplications (Akhunov et al. 2007) or increases in genome size (SanMiguel et al. 1998, Kalendar et al. 2000), thus being a major driver for evolution. One would assume that because of this plasticity the genomes of the Triticeae would be highly different, but it was discovered that genes and gene order are well conserved among grass species (Casacuberta and Santiago 2003). Genes in grasses are organized into so-called gene islands (Barakat et al. 1997) that are highly conserved, while the surrounding intergenic regions turn over at a fast pace (in evolutionary terms) (Wicker et al. 2003, Dubcovsky and Dvorak 2007). However, this concept only refers to the structure of gene islands in general, but not to their content.

The genome structure of grass species makes the identification of genes quite challenging, but because of the high academic and agronomic interest in these species new ways to cope with these difficulties are being created. Transposable elements are discovered and stored in large databases like the Triticeae Repeat Sequence Database (TREP, (Wessler 1998) or the MIPS Repeat Element Database (Nussbaumer et al. 2013) to allow plant scientists to correctly annotate and interpret genomic sequences. After annotating the repetitive portion of a given sequence the search for open reading frames (ORFs) can begin. Tools like FGENESH (Salamov and Solovyev 2000) and BLASTX (States and Gish 1994) have been used widely for their ease of use (can be run in a browser window) and good “accuracy” in predicting ORFs. These ORFs or candidate genes can be validated *in silico* by exploring expressed sequence tag (EST) databases (Kawaura et al. 2005), but ultimately they need to be validated by wet lab experiments.

Numerous methods for the validation of candidate genes have been developed over the years. One of the most widespread methods is expression profiling (Uzarowska et al. 2009), in which the expression of candidate genes is compared between a panel of lines differing in the phenotype of

interest. The rationale behind this method is that expression of the causal candidate gene should be correlated with the presence of the phenotype of interest.

Similar to that, candidate genes can be tested in a high-throughput manner by inhibiting their expression via post-transcriptional gene silencing (PTGS) (Waterhouse and Helliwell 2003). The PTGS assays can be designed to knock down the expression of a range of related genes, which is very useful in polyploid species like wheat, which can have three or more homoeologous copies of a single locus (Uauy et al. 2006). The knockdown of a causal candidate gene by PTGS should mimic a loss-of-function phenotype.

In recent years Targeting Induced Local Lesions IN Genomes (TILLING) has established itself as a viable method for high-throughput validation of candidate genes in both model organisms and polyploid species like wheat (McCallum et al. 2000, Slade et al. 2005, Uauy et al. 2009). In fact, a polyploid species has a much higher mutational density than a diploid species because of its genetic redundancy in form of homoeologous gene copies. It is estimated that a TILLING population of hexaploid wheat has on average one mutation per 32 kb of sequence, compared to one mutation per 380 kb of sequence for a diploid (Wang et al. 2012). This high mutation density significantly reduces the size of the population that needs to be screened in order to obtain a knock-out mutant for the gene-of-interest in a polyploid species compared to a diploid species. However, the genes of interest obviously need to be present within the TILLING population, which could be problematic for rare alleles.

Another method to validate candidate genes is the analysis of allelic diversity for the trait of interest within a natural population (Peleg et al. 2008). Analysing the nucleotide sequence of candidate genes within a number of accessions from a broad geographic range should yield a good correlation between candidate ORFs and the phenotype of interest. It is also possible that candidate genes are not present at all in some members of a diversity panel. Allelic diversity studies can easily be combined with expression analysis to increase the robustness of the correlations.

Genetic complementation is regarded as one of the best methods for validating candidate genes. By transforming a single gene into another organism the effect of this gene can be readily studied in stable transgenic individuals. This method is often used to complement mutants with a functional gene copy or to transfer a single gene with a clear phenotype as resistance against a pathogen (Huang et al. 2003) or grain hardness (Beecher et al. 2002). However it is still difficult and expensive to stably transform many crop species like wheat as it requires regeneration of transformed callus tissue into mature plants and at the same time to ensure the correct expression pattern of the transgene (Harwood 2012).

The method of choice for validation depends on the function of the gene of interest, the ease of distinguishing the phenotype caused by the gene of interest and on the budget. Usually several

methods are used in a complementary way in candidate gene validation to achieve the strongest correlation possible.

We hypothesized that *Iw1* is either non-functional, not expressed or even deleted in modern germplasm due to its clearly observable phenotype, its dominant nature and the lack of it across modern varieties. This hypothesis was underlined by the results of our physical maps (Chapter 2.3), which suggested the existence of two different haplotypes at the *Iw1* locus. We have acquired over 100 accessions of *Triticum turgidum* ssp. *dicoccoides* (TTD) differing in their glaucous phenotype. These lines were used to test the expression profile of candidate genes and to understand their allelic diversity. Candidate genes that passed these tests were transformed into the glaucous spring wheat cultivar Fielder at the National Institute of Agricultural Botany (NIAB). We were not aware of the existence of a TILLING population containing *Iw1*, so instead we created EMS populations of Shamrock and TTD140 to knock out *Iw1*. In addition, we have used our BC_4 NILs in an RNA-seq experiment to identify *Iw1*.

3.1 Materials and Methods

3.1.1 Plant material

Shamrock, Shango and the Shango x Shamrock DH population have been described in Chapter 1.1.1. Shamrock was used as pollen donor for a cross to Shango to generate a large F_2 mapping population. TTD140 is a non-glaucous wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*, TTD) accession from Israel that was shown to carry *Iw1*. Langdon (*Triticum turgidum* ssp. *durum*) is a glaucous durum wheat cultivar. TTD140 was used as pollen donor for a cross to Langdon to obtain several hundred F_2 plants.

We have acquired 124 TTD accessions from the National Institute of Agricultural Botany (NIAB). We have passport information for 96 of these accessions.

3.1.2 Total RNA extraction

Tissue was collected and immediately frozen in liquid nitrogen and stored at -80°C until needed. Approximately 1 mg of tissue was homogenized in 0.5 mL of TriReagent (15596-018, Ambion) and incubated for 5 min to separate the nucleoprotein complex. Then 0.1 mL of Chloroform was added, the sample was shaken vigorously and incubated it at room temperature (RT) for 3 min. Samples were centrifuged for 10 min at 10,000 x g. The upper aqueous layer was transferred to a new tube and 0.25 mL of propan-2-ol were added to precipitate the RNA. The sample was incubated for 10 min at RT and then centrifuged for 10 min at 10,000 x g. The supernatant was removed carefully, followed by the addition of 0.5 mL of 70% Ethanol. The samples were centrifuged for 10 min at 10,000 x g and the supernatant was removed. Excess liquid was removed with a pipette. The pellet was then immediately resuspended in 20 µL of nuclease-free H_2O. RNA was stored at -80°C.

3.1.3 DNase treatement of RNA

Prior to cDNA synthesis the RNA was treated with RQ1 DNase (M6101, Promega).

1. DNase recipe:

RNA	5 µL
RQ1 RNase-Free DNase 10X Reaction Buffer	1 µL
RQ1 RNase-Free DNase	1 µL
Nuclease-free water to a final volume of	3 µL

2. Incubate samples for 30 min at 37°C.
3. Add 1 µL of RQ1 DNase Stop Solution to each sample to terminate the reaction.
4. Incubate samples for 10 min at 65°C to inactivate DNase.
5. Store at -80°C.

3.1.4 cDNA synthesis from total RNA

Reagents:

- M-MLV Reverse Transcriptase (28025-013, Invitrogen)
- Oligo(dT) Primers (AM5730G , Invitrogen)
- RNaseOUT™ Recombinant Ribonuclease Inhibitor (10777-019, Invitrogen)
- dNTP mix (U1511, Promega)

1. Mix the following:

DNase treated RNA	6 µL
dNTPs (10 mM)	1 µL
Oligo(dT) Primers	1 µL

2. Incubate at 65°C for 5 min, then quick chill on ice.

3. Add the following to the mix:

5x First-Strand Buffer	4 µL
0.1 M DTT	2 µL
RNaseOUT	1 µL
M-MLV RT	1 µL

4. Incubate for 50 min at 37°C.

5. Incubate for 15 min at 70°C to inactivate reverse transcriptase.

6. Store cDNA at -20°C for short term and at -80°C for long term.

3.1.5 Ethyl methanesulfonate (EMS)-mediated mutagenesis

The following protocol was used to mutagenise wheat seeds with EMS.

1) Mix 91 mL of Tween 20 with 819 mL H_2O = 910 mL to obtain a 10% solution in a 1 L bottle.
2) Weigh 150 mL of seeds (~3,000 seeds) and add to each 1 L bottle.
3) Agitate on roller bar shaker at setting 7-8 for 15 min.
4) Take out 10% Tween 20 solution and rinse with water. Then add 910 mL H_2O and place on roller bar shaker at setting 7-8 for 5 min.
5) Repeat three more times.
6) Add 910 mL of H_2O and proceed to add EMS. For 0.85% EMS concentration 6.4269 mL EMS was added (MW 124.16 g/mol).
7) Agitate on roller bar shaker at setting 3-4 for 18 hours.
8) Eliminate EMS solution in NaOH (40 g/L) + Thioglycolic acid (12.5 mL/L) and add fresh H_2O. Rinse 5 times changing the water each time. Then agitate on roller bar shaker for 15 min at setting 4.

9) Combine all seeds into one bucket and place under running water for 3 hours. Place cheesecloth on top to be sure that seeds do not float out.

3.1.6 Sequencing Total RNA using Illumina HiSeq

Total RNA was extracted from 100 mg of tissue using twice the amount of volumes described in Materials and Methods 3.1.2. Total RNA was quantified using (Implen Nanophotometer) and 50 ng of total RNA were handed to TGAC for preparation of a paired-end 100 bp library. Libraries were multiplexed in sets of two to be run on a single lane; a leaf sheath sample was always paired with a peduncle sample from the same plant. The subsequent sequencing was performed using an Illumina HiSeq 2500 in 'Output Run' mode to generate >100,000,000 reads per lane.

3.2 Results

3.2.1 Phenotypic analysis of the TTD diversity panel

The 124 TTD accessions were sown out in pairs, grown for a complete growth cycle and the seeds were recovered. Only 43 accessions have a spring growth habit, while the remaining 81 accessions have a requirement for vernalization. The accessions showed a range of glaucous phenotypes summarized in Table 3.1. Apart from their glaucous phenotype the TTD accessions displayed a wide range of phenotypes for many agriculturally important traits like flowering time, height, tiller number, seed size and glume size.

Table 3.1: Glaucous phenotypes of the TTD diversity panel

This table summarizes the glaucous phenotypes observed in the 124 accessions of TTD. Glaucousness was recorded as visible wax on four plant organs: Leaf sheath, flag leaf, peduncle and spike. Presence of a waxy bloom is represented by (+), while (-) denotes absence of glaucousness.

Leaf Sheath	Flag leaf	Peduncle	Ear	No. of accessions	Origin
+	+	+	+	26	South/North
+	+	+	+	22	No passport information
+	+	+	-	2	South/North
+	+	-	-	2	South/North
+	-	-	+	2	North
+	-	-	-	10	South/North
-	-	-	-	55	South
-	-	-	-	5	No passport information

Wild emmer wheat grows in a discontinuous arc from southern Levant to northwestern (NW) Syria, southeastern (SE) Turkey, northern Iraq, and NW Iran (Figure 3.1). It was shown by AFLP analysis that wild emmer wheat formed two populations (Ozkan et al. 2002); a northern population (Turkey, Iraq, Iran, Armenia and Azerbaijan) and a southern population (Jordan, Israel, Palestine, Lebanon, and southwestern Syria). Emmer, like einkorn, was domesticated west of the Diyarbakir region in southeastern (SE) Turkey (Figure 3.2), and the domesticated emmer expanded to the southern Levant, where it inter-crossed with wild emmer, forming a second centre of diversity (Luo et al. 2007).

Figure 3.1: Map of the Fertile Crescent

The green shaded area is known as the Fertile Crescent, where cereals like einkorn, emmer wheat and barley have been grown by early farmers. Wild relatives of these species still grow in this area and represent a huge source for genetic diversity. Figure taken from Feuillet et al. (2008).

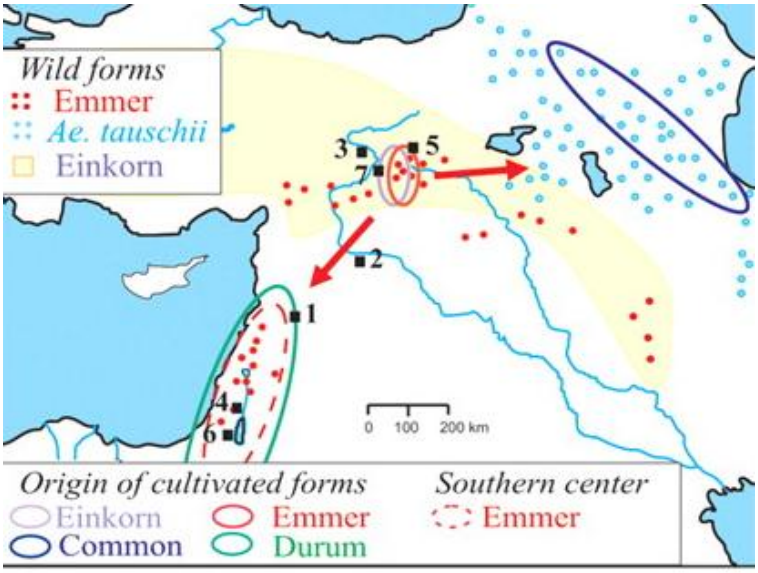


Figure 3.2: Distribution of wheat species in the Fertile Crescent

Emmer wheat was domesticated west of the Diyarbakir region (full red circle) and expanded southwest, where it mixed with wild emmer wheat (dotted red circle) forming a southern centre of diversity. Northeastern expansion of domesticated emmer led to hybridization with *Aegilops tauschii* (purple ring) and the emergence of common wheat (*T. aestivum*). Figure taken from Dubcovsky and Dvorak (2007).

The 96 accessions, for which we have information about their collection site for, seem to fall into two groups when analysing their phenotype for glaucousness. We considered a glaucous phenotype when at least one aerial plant organ of an accession showed a waxy bloom. Accessions originating from the northern population (24 in total) exclusively display a glaucous phenotype, while accessions from the southern population of wild emmer (72 in total) display both glaucous and non-glaucous phenotypes (Table 3.2). From these 72 southern accessions, 55 display a non-glaucous phenotype, while 17 display a glaucous phenotype. Test-crosses between 22 non-glaucous TTD accessions and the glaucous durum wheat cultivar Langdon have been made already, with the F_1 generation currently being grown in the glasshouse. The remaining non-glaucous TTD accessions will also be crossed with Langdon. This will elucidate whether the inhibitor *Iw1* or the absence of wax-producing gene *W1* is responsible for the non-glaucousness in these accession.

These preliminary results suggest that glaucousness is the ancestral phenotype of wild emmer wheat and that non-glaucousness controlled by *Iw1* only developed in the southern population, but not in the northern one. Examining more accessions from the northern wild emmer population would be necessary to confirm these results.

Table 3.2: Geographic distribution of glaucous and non-glaucous phenotypes among 96 TTD accessions for which passport information is available

Country	glaucous	non-glaucous	Population
Turkey	20	-	Northern
Iraq	2	-	Northern
Iran	1	-	Northern
Armenia	1	-	Northern
Jordan	-	10	Southern
Israel	6	30	Southern
Lebanon	5	3	Southern
SW Syria	6	12	Southern

3.2.2 Summary of genes predicted in the physical maps of TTD140 and RSL65

We have started construction of two physical maps for the *Iw1* interval, one in the non-glaucous TTD140 and another in the glaucous RSL65. The TTD140 map is comprised of two contigs made from eleven BAC clones with a total size of 942 kb, excluding overlap between BACs. Out of these 942 kb approximately 684 kb (~73%) correspond to known TEs while the remaining 258 kb (27%) correspond to low copy number sequence. Nine genes were predicted inside these 258 kb using FGENESH and BLASTX, which equals 1 gene every 29 kb. The predicted genes are listed in Table 3.3 based on their position in the physical map from proximal to distal of the *Iw1* interval.

Table 3.3: List of predicted genes in TTD140

The genes predicted in the TTD140 physical map are listed here in the approximate order from proximal to distal of the *Iw1* interval. (+) denotes presence of a predicted gene in a genotype, (-) denotes absence of a predicted gene in a genotype, (?) denotes an unclear status. (+) and (-) with a (?) next to them denotes hypothesized presence or absence. Chinese Spring information is based on the 2BS flow sorted survey sequence assemblies.

predicted gene	Langdon *Glaucous*	Shango *Glaucous*	TTD140 *Non-glaucous*	Shamrock *Non-glaucous*	Chinese Spring *Glaucous*
305N15_CYP450	-	-	+	+	-
305N15_CHS	?	-	+	?	-
774P06_WPK	-	-	+	+	-
23N19_FAE1	-	-	+	+	-
23N19_CYP450	-	-	+	+ ?	-
23N19_LRR	-	-	+	-	-
551I08_PGG	-	-	+	-	-
784K20_1180	- ?	- ?	+	- ?	-
784K20_PFF	- ?	- ?	+	- ?	-

The physical map of RSL65 is comprised of three contigs made from eight BAC clones. However, only four of these BAC clones have been sequenced so far, resulting in a total map size of just 388 kb. Out of these 388 kb approximately 209 kb (54%) correspond to known TEs while the remaining 179 kb (46%) correspond to low copy number sequence. This is a much lower ratio of repetitive elements than would be expected. However, the map is small and might depict a particularly gene-rich region. Seven genes were predicted inside the 179 kb of low copy number sequence using FGENESH and BLASTX, which equals 1 gene every 26 kb. The predicted genes are listed in Table 3.4 based on their position in the physical map from proximal to distal of the *Iw1* interval.

Table 3.4: List of predicted genes in RSL65

The genes predicted in the RSL65 physical map are listed here in the approximate order from proximal to distal of the *Iw1* interval. (+) denotes presence of a predicted gene in a genotype, (-) denotes absence of a predicted gene in a genotype, (?) denotes an unclear status. (+) and (-) with a (?) next to them denotes hypothesized presence or absence. Chinese Spring information is based on the 2BS flow sorted survey sequence assemblies.

predicted gene	Langdon *Glaucous*	Shango *Glaucous*	TTD140 *Non-glaucous*	Shamrock *Non-glaucous*	Chinese Spring *Glaucous*
264M08_PKc	+	+	-	+	+
264M08_CYP450	+	+	-	-	+
264M08_1180	+	+	-	-	+
618L11_CYP450	+	+ ?	- ?	- ?	+
618L11_HlyIII	+	?	?	?	+
1326P06_CYP450	+	+ ?	- ?	- ?	+
1326P06_HlyIII	+	?	?	?	+

3.2.3 Analysis of predicted genes in TTD140

The following section covers an initial analysis of the genes predicted within the TTD140 physical map. The aim is to characterize the putative candidate genes among the four parental lines of the genetic maps, namely TTD140, Langdon, Shamrock and Shango. This should yield and initial estimate whether the predicted gene is a likely candidate gene of Iw1, in which case further validation efforts will be undertaken. The sequences of the primers mentioned here can be found in the Appendix (A6). For the BLASTN queries against the wheat survey sequence database described here, hits with a similarity below 90% were ignored.

3.2.3.1 *305N15_CYP450*

FGENESH and BLASTX predicted a gene with two exons (CDS: 1,542 bp; genomic: 1,981 bp), which encodes a member of the Cytochrome P450 (PLN02687 (Marchler-Bauer et al. 2011), P450) family. A BLASTN query against the wheat survey sequence database yielded one hit with 93.7% similarity (identities: 2,321/2,477) from 2AS (contig 387072), while hits to 2BS and 2DS contigs had <85% similarity. BLASTN queries against several EST databases resulted in a single hit (BJ256128) with 94% similarity (identities: 561/591), but which was 100% identical with 2AS contig 387072. This *in silico* analysis suggested that Chinese Spring has a homoeologue of *305N15_CYP450* in the A genome, but no homologue of the TTD140 2BS gene could be identified.

Two primer pairs were designed to amplify the first (F1+R1) and the second (F2+R2) exon (Figure 3.3). PCR products of expected size were amplified and sequenced for both primer pairs in Langdon, Shango, TTD140 and Shamrock. The sequence of the Langdon and Shango F1+R1 amplicons was >99% similar to wheat 2AS contig 387072, but 97% (identities: 855/881) and 98% (identities: 803/823) similar to *305N15_CYP450* respectively. The sequence of the F2+R2 amplicons however was only 92% similar to 2AS contig 387072, but 95% to *305N15_CYP450*. A comparison between *305N15_CYP450* and contig 2AS 387072 revealed that SNPs are present at the 3' end of both F2 and R2, which prevented amplification of the A-genome homoeologue of *305N15_CYP450*. Surprisingly though, 2AS contig 387072 remained the top hit in the wheat survey sequence database using the Langdon and Shango F2+R2 amplimers as queries.

The sequence of both amplicons from TTD140 and Shamrock was >99.9% similar to *305N15_CYP450*. This suggested that this gene is present in both non-glaucous varieties while absent from the glaucous varieties, which supports the hypothesis that *305N15_CYP450* is a candidate gene of *Iw1*. Further experiments including expression analysis are necessary to test this hypothesis.

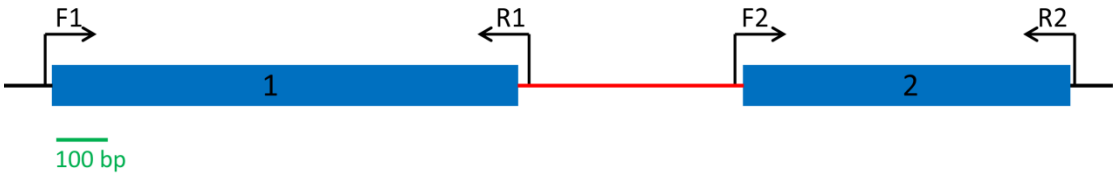


Figure 3.3 Genomic interval of *305N15_CYP450*

The figure represents the genomic interval of the predicted gene *305N15_CYP450*. Labelled blue boxes represent the exons, the red line represents the intron and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.2 *305N15_CHS*

FGENESH and BLASTX predicted a gene with two exons (CDS: 1,203 bp; genomic: 1,289 bp) and encodes a CHS-like domain (cd00831) protein. A BLASTN query against the wheat survey sequence database yielded one hit (contig 5273836) from 2AS with 98.1% similarity (identities: 1,264/1,289), while hits to 2BS and 2DS contigs had <85% similarity.

Two primer pairs were designed to amplify the first (F3+R3) and second (F4+R4) exon (Figure 3.4). PCR products of expected size were amplified and sequenced for Langdon, Shango, TTD140 and Shamrock. The sequence of the F4+R4 amplicon from Shango was 99.5% similar (identities: 874/878) to wheat 2AS contig 5273836. The sequencing of all the other amplicons failed though. The experiment needs to be repeated, possibly with new primer pairs. Until then *305N15_CHS* remains a candidate gene of *Iw1*.

Figure 3.4 Genomic interval of *305N15_CHS*
The figure represents the genomic interval of the predicted gene *305N15_CHS*. Labelled blue boxes represent the exons, the red line represents the intron and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.3 *774P06_WPK*

FGENESH and BLASTX predicted a gene with five exons (CDS: 654 bp; genomic: 1,530 bp) that encodes a protein with a partial SRPBCC_11 domain (cd08866) and a partial FR_SDR_e domain (cd08958). A BLASTN query against the wheat survey sequence database yielded one hit (contig 4362786) from 6BL with 96.6% similarity (identities: 1,096/1,134), but no hits to any group 2 contigs. The 6BL contig is truncated though and starts only in the middle of the second exon. Apart from this 6BL contig, no nucleic acid database yielded a hit to the entire *774P06_WPK* gene, but only either to the first two or the last two exons.

Two primer pairs were designed to amplify the entire gene (Figure 3.5). PCR products of expected size were sequenced for Langdon, Shango, TTD140 and Shamrock. The sequence of F5+R6 amplicons from Langdon and Shango was 99.9% similar (identities: 1,133/1,134) to wheat 6BL contig 4362786, while the sequence of F5+R6 amplicons from TTD140 and Shamrock was 100% (identities: 1,530/1,530) and 98% (identities: 1,499/1,528) similar to *774P06_WPK*, respectively. These results suggest that *774P06_WPK* is present in both non-glaucous varieties, whereas a paralogous copy is present in the glaucous varieties on 6BL. The expression of *774P06_WPK* and its allelic diversity in the TTD panel need to be tested to validate or rule out *774P06_WPK* as a candidate gene for *Iw1*.

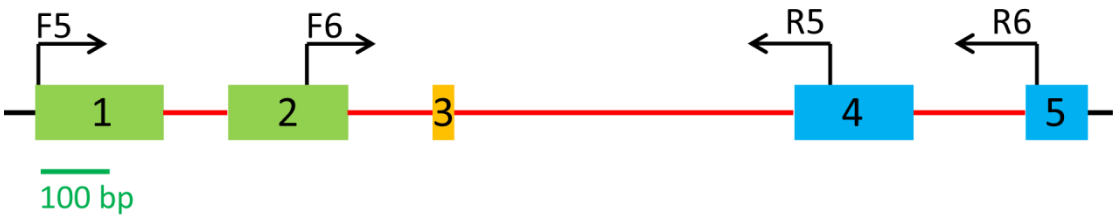


Figure 3.5 Genomic interval of *774P06_WPK*
The figure represents the genomic interval of the predicted gene *774P06_WPK*. The green boxes represent the SRPBCC_11 domain encoding exons, while the blue boxes represent the FR_SDR_e domain encoding exons. The orange box represents an exon that does not seem to belong to either of the two domains. The red lines represent introns and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.4 *23N19_FAE1*

FGENESH and BLASTX predicted a gene with one exon and a LTR retrotransposon (Ale_391M13_1, *Copia* family) of 6,884 bp in size inserted at position 1,049 counting from the ATG (target site duplication: GCCCG). The uninterrupted coding sequence is 1,512 bp in length and encodes a CHS-like domain (cd00831) protein, which consists of two domains itself: A FAE1_CUT1_RppA domain (pfam08392) and an ACP_syn_III_C domain (pfam08541). The inserted retrotransposon leads to a premature STOP codon at position 1,059 counting from the ATG. This is predicted to abolish the ACP_syn_III_C domain, but leaves the FAE1_CUT1_RppA domain intact.

A BLASTN query against the wheat survey sequence database using the uninterrupted CDS yielded one hit (contig 5275619) from 2AS with 97.8% similarity (identities: 1,449/1,481), but no hits from 2BS or 2DS. Using the sequence with the inserted retrotransposon for the BLASTN yielded the same 2AS contig in addition to multiple hits to the Ale_391M13_1 sequence. Two primer pairs were designed to amplify the putative gene, with primers inside the top part of the CDS, inside the repetitive sequence and inside the lower CDS (Figure 3.6). Using primer pair F7+R8, PCR products of 1500 bp size (size of the uninterrupted CDS) were amplified and sequenced from Langdon, Shango, TTD140 and Shamrock. These sequences were >99.5% similar to the wheat 2AS contig 5275619.

Primer R7 was designed with its 3' end inside the *23N19_FAE1* CDS but with half of the primer inside the retrotransposon. The sequence of F7+R7 PCR products from Langdon and Shango was 99.9% similar to the wheat 2AS contig 5275619. In contrast, the sequence trace files of the TTD140 F7+R7 amplicon showed the presence of overlapping single peaks, suggesting that both homoeologues of the gene had been amplified, as these overlapping peaks coincided with SNPs between *23N19_FAE1* and 2AS contig 5275619.

The sequence trace file of the Shamrock F7+R7 amplicon also shows overlapping single peaks, but is much more complicated than in TTD140. The Shamrock sequence trace file clearly shows that two products have been amplified with 17 SNPs between them and *23N19_FAE1*. Nine of these SNPs can be explained by homoeologous SNPs between *23N19_FAE1* and 2AS contig 5275619. Three further SNPs are present in one or the other amplicon, but they seem to be specific to the Shamrock genotype as *23N19_FAE1* and 2AS contig 5275619 are identical at these positions. The remaining five SNPs are polymorphic with *23N19_FAE1*, but identical with 2AS contig 5275619. This suggested that Shamrock carries the 2AS homoeologue of *23N19_FAE1*, but the identity of the other amplicon is unclear.

These results suggest that although half of primer R7 lies inside the retrotransposon the primer can still bind to the 2AS homoeologue. New primers need to be designed or the PCR products using the existing primer pairs need to be subcloned to disentangle the different amplicons. Until then *23N19_FAE1* remains a valid candidate for *Iw1*.

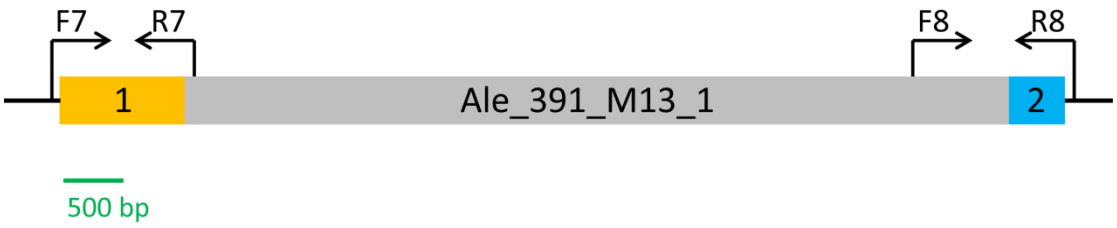


Figure 3.6 Genomic interval of *23N19_FAE1*

The figure represents the genomic interval of the predicted gene *23N19_FAE1*. The yellow box indicated the first half of the CDS, which encodes the FAE1_CUT1_RppA domain, while the blue box represents the ACP_syn_III_C domain encoding exon. The grey box represents the inserted transposable element and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.5 *23N19_CYP450*

FGENESH and BLASTX predicted a gene with six exons (CDS: 1,695 bp; genomic: 2,306 bp), which encodes a member of the Cytochrome P450 (pfam00067 (Finn et al. 2010)) family. A BLASTN query against the wheat survey sequence database yielded one hit with 97.7% similarity (identities: 2,254/2,306) from 2AS (contig 5275619), five hits with <85% similarity from 2BS and no hits from 2DS.

Two primer pairs were designed; one to amplify the first three exons (F9+R9) and the other to amplify the remaining three exons (F10+R10) of *23N19_CYP450* (Figure 3.7). PCR products of expected size from both primer pairs were obtained and sequenced for Langdon, Shango, TTD140 and Shamrock. The sequence of the Langdon and Shango amplicons was 99.7% similar to wheat 2AS contig 5275619 for both primer pairs.

The TTD140 sequence trace files for both primer pairs showed the presence of overlapping single peaks, suggesting two very similar but nonetheless different products having been amplified. These overlapping peaks coincide with SNPs between *23N19_CYP450* and the 2AS contig 5275619, suggesting that both homoeologues of the gene have been amplified in TTD140.

The sequence of the F10+R10 amplicon from Shamrock matches the 2AS contig 5275619 100% (identities: 1393/1393). In contrast the trace files of the F9+R9 amplicon showed the presence of overlapping single peaks, suggesting two very similar but nonetheless different products having been amplified, suggesting that both *23N19_CYP450* and its 2A homoeologue have been amplified. The lack of 2BS sequence in the F10+R10 amplicon of Shamrock might be caused by SNPs overlapping with the primer pair. This gene seems to be present in TTD140 and Shamrock and absent from Langdon and Shango, making it a viable candidate for *Iw1*.

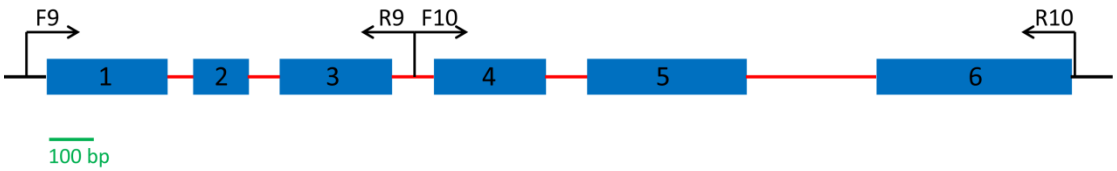


Figure 3.7 Genomic interval of *23N19_CYP450*

The figure represents the genomic interval of the predicted gene *23N19_CYP450*. Labelled blue boxes represent the exons, red lines represent introns and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.6 *23N19_LRR*

FGENESH and BLASTX predicted a gene with three exons (CDS: 666 bp; genomic: 1,073 bp), which encodes a partial LRR-domain (COG4886 (Tatusov et al. 2000)) protein. A BLASTN query against the wheat survey sequence database yielded no hits from 2AS, 2BS or 2DS. BLASTN queries of several EST databases also did not yield any hits. Two primer pairs were designed to amplify the putative gene (Figure 3.8). A PCR product could only be obtained from TTD140 using primer pair F11+R12, the sequence of which was 100% similar (identities: 1,218/1,218) to *23N19_LRR*. The absence of a PCR product in Shamrock suggested that this gene is not *Iw1*. However, this hypothesis needs to be verified by several more independent primer pairs or preferably by Southern Blot. Until then *23N19_LRR* remains a candidate gene of *Iw1*.

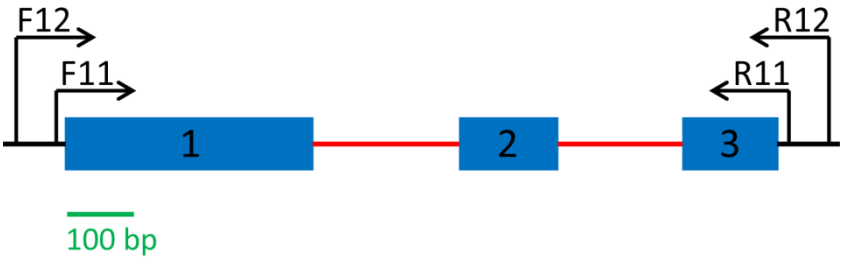


Figure 3.8 Genomic interval of *23N19_LRR*

The figure represents the genomic interval of the predicted gene *23N19_LRR*. Labelled blue boxes represent the exons, red lines represent introns and arrows indicate the position of primer pairs used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.7 *551I08_PGG*

FGENESH and BLASTX predicted a gene with two exons (CDS: 2,514 bp; genomic: 2,780 bp) that encodes a protein of unknown function with five PGG domains (pfam13962). A BLASTN query against the wheat survey sequence database yielded one hit (contig 5167874) from 2BS with 89.3% similarity (identities: 2,149/2,407), several partial hits with >92% similarity to 2DS and no hits to 2AS.

Two primer pairs (F13+R13 and F14+R14) were designed to amplify both exons (Figure 3.9). Both primer pairs amplified PCR products of expected size in TTD140, but not in the other three genotypes. The sequence from both TTD140 amplicons was 99.9% similar to *551I08_PGG*. This gene appears to be unique to TTD140. Since this gene could not be amplified from Shamrock repeatedly it was thus concluded that *551I08_PGG* is not *Iw1*.

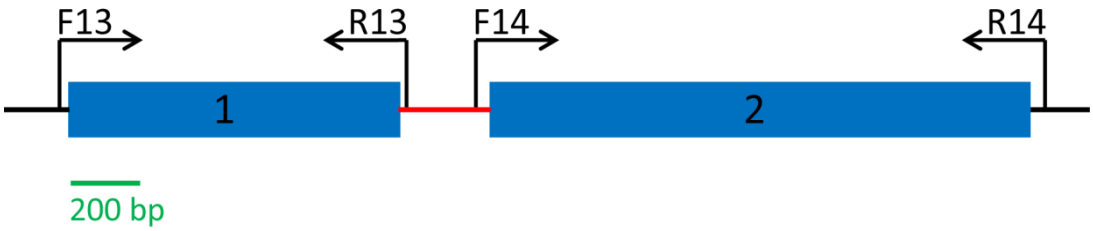


Figure 3.9 Genomic interval of *551I08_PGG*

The figure represents the genomic interval of the predicted gene *551I08_PGG*. The blue boxes represent exons, the red line represents the intron and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.3.8 *784K20_PFF*

FEGENESH and BLASTX predicted a gene with six exons (CDS: 1,146 bp; genomic: 4,681 bp) that encodes a protein of unknown function with no conserved domains. A BLASTN query against the wheat survey sequence database yielded one hit (contig 5198780) to 2AS with 92.9% similarity (identities: 4,277/4,604), but no hits to 2BS or 2DS.

Two primer pairs were designed to amplify the first three (F15+R15) and the last three (F16+R16) exons (Figure 3.10). The sequence trace files of the F15+R15 amplicon from TTD140 showed the presence of overlapping single peaks, suggesting two very similar but nonetheless different products having been amplified. These overlapping peaks coincide with SNPs between *784K20_PFF* and the 2AS contig 5198780, suggesting that both sequences have been amplified in TTD140. The sequence of F15+R15 amplicons from Langdon, Shango and Shamrock was >99% similar to the 2AS contig 5198780.

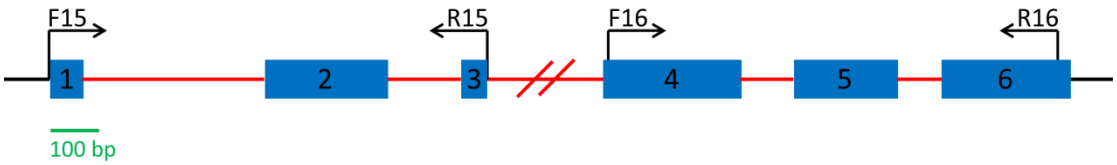


Figure 3.10 Genomic interval of *784K20_PFF*

The figure represents the genomic interval of the predicted gene *784K20_PFF*. The blue boxes represent the exons, red lines represent introns, while arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The intron between exon 3 and 4 is 2.8 kb in size and was shortened here for the purpose of clarity. The green bar indicated the scale.

The F16+R16 primer pair however amplified two different products in all four genotypes (overlapping peaks in the trace files), but none of these was identical with *784K20_PFF*. A simple sequence repeat (SSR) within the *784K20_PFF* sequence was different to the amplified products: *784K20_PFF* has ten repeats of the base pair sequence TA, while all amplified products only had five repeats of TA (Figure 3.11).

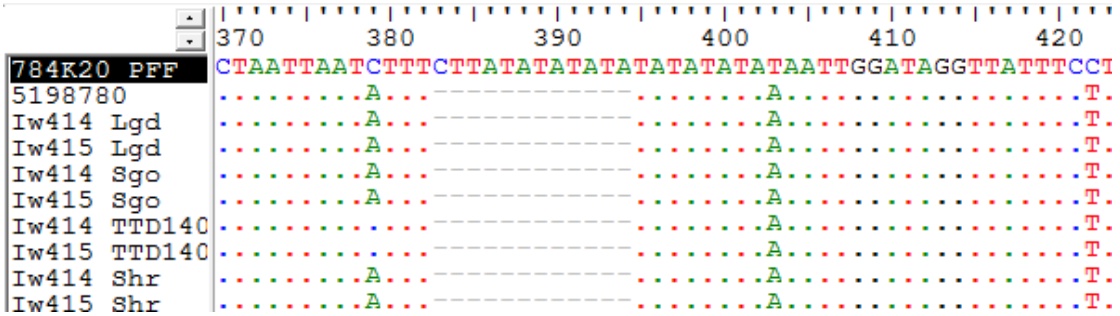


Figure 3.11: Alignment of PCR products of *784K20_PFF*

The F16+R16 primers pair amplified products from Langdon, Shango, TTD140 and Shamrock. None of these products was identical to *784K20_PFF*, as can be seen from the reduced copy number of the simple sequence repeat TA in the trace files of all PCR products. Lgd = Langdon, Sgo = Shango, Shr = Shamrock

The F16+R16 primer pair failed to amplify the correct sequence even from TTD140 even though there are no SNPs underlying the primers. From the F15+R15 amplification we would hypothesise that *784K20_PFF* is unique to TTD140 and thus cannot be *Iw1*, as appears to be missing from Shamrock. However, this hypothesis needs to be verified by several more independent primer pairs, as we cannot exclude the possibility that the lack of amplification using F15+R15 did not result from SNPs coinciding with the primer pairs. Until then *784K20_PFF* remains a candidate gene of *Iw1*.

3.2.4 Analysis of predicted genes in RSL65

3.2.4.1 *264M08_PKc*

FGENESH and BLASTX predicted a gene with seven exons (CDS: 1,044 bp; genomic: 5,348 bp), which encodes a PKc (cd00180) protein kinase. A BLASTN query against the wheat survey sequence database yielded one hit with 99.9% similarity (identities: 5,344/5,348) from 2BS (contig 5208300), but no contig was hit from 2AS or 2DS. A large intron (3,279 bp) separates the first four exons from the remaining three exons.

Two primer pairs were designed to amplify the first three exons (F17+R17) and the last three exons (F18+R18)(Figure 3.12). PCR products of expected size were obtained from both primer pairs for Langdon, Shango and Shamrock. The sequence of the F17+R17 amplicon from Langdon, Shamrock and Shango was 100% (identities: 1,034/1,034), 99.9% (identities: 975/976) and 100% (identities: 1,026/1,026) similar to *264M08_PKc*, respectively. The sequence of the F18+R18 amplicon from Langdon, Shamrock and Shango was 100% (identities: 907/907), 99.8% (identities: 912/914) and 100% (identities: 921/921) similar to *264M08_PKc*, respectively. The SNPs between *264M08_PKc* and Shamrock are predicted to cause a non-synonymous change in two amino acids: Glutamatic acid is replaced by Lysine at position 46 and Serine is replaced by Isoleucine at position 324. Several active sites are present in the predicted protein around position 50, but as Glutamatic acid and Lysine both belong to the same amino acid group (amino acids with electrically charged side chains) it seems unlikely that this exchange would affect protein function in Shamrock. No active sites are predicted at position 324 and in fact the catalytic domain of PKc ends at position 300, which makes it unlikely that the change from Serine to Isoleucine would affect protein function in Shamrock.

No PCR products could be amplified from TTD140 repeatedly. *264M08_PKc* was also not identified in any of the sequenced TTD140 BACs, although this could be due to the fact that the TTD140 physical map is not yet complete. Several more primer pairs will be designed to try and amplify the gene from TTD140, but it appears that *264M08_PKc* is not *Iw1*.

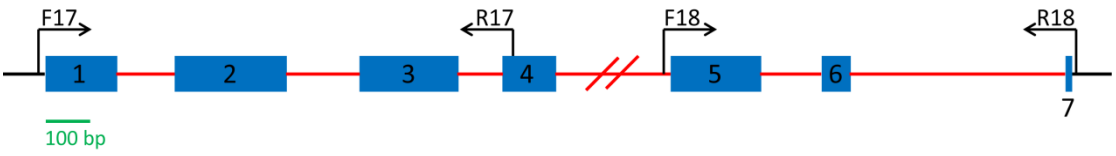


Figure 3.12 Genomic interval of *264M08_PKc*

The figure represents the genomic interval of the predicted gene *264M08_PKc*. The blue boxes represent exons, red lines represent introns, while arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The intron between exon 4 and 5 is >3 kb in size and was shortened here for the purpose of clarity. The green bar indicated the scale.

3.2.4.2 *264M08_CYP450*

FGENESH and BLASTX predicted a gene with one exon (CDS: 1,557 bp; genomic: 1,557 bp), which encodes a member of the Cytochrome P450 (PLN02169) family. A BLASTN query of the CDS against the wheat survey sequence database yielded three hits with 99.9% similarity (identities: 1,555/1,557 for all three hits) from 2BS (contigs 5201145, 5230436 and 5245086), two contigs with 95.1% (2DS contig 5389765) and 95.9% (2DS contig 27654) similarity (identities: 562/591 and 610/636, respectively) from 2DS, but no contigs from 2AS. The three contigs from 2BS likely represent one gene that was assembled into three different contigs, as the two SNPs between each of them and *268M08_CYP450* are at the exact same positions.

Two primer pairs were designed to amplify the CDS (Figure 3.13). PCR products of expected size were amplified in Langdon, Shango, and Shamrock. The sequence of the F19+R19 amplicon from Langdon was 100% (identities: 652/652) identical to *264M08_CYP450*. The sequence trace files of the F19+R19 amplicons from Shango and Shamrock are heavily contaminated. A BLASTN query of the Shango sequences against the wheat survey sequence database yields the three 2BS contigs (contigs 5201145, 5230436 and 5245086) as the top hits, while for the Shamrock sequences 2DS contig 5389765 is the top hit. This suggested that *264M08_CYP450* was not amplified in Shamrock.

The sequence of the F20+R20 amplicons from Langdon, Shango and Shamrock was 99.9% (identities: 1,030/1,031), 99.7% (identities: 1,016/1,019) and 96.6% (identities: 978/1,012) similar to *264M08_CYP450*, respectively. The sequence of the F20+R20 amplicons from Shango and Shamrock are 95.8% (identities: 609/636) and 99.9% (identities: 635/636) similar to 2DS contig 27564, indicating that the F20+R20 primer pair did not amplify *264M08_CYP450* in Shamrock.

No PCR products could be amplified from TTD140 repeatedly and the *264M08_CYP450* sequence was not identified in any of the sequenced TTD140 BACs, although this could be due to the fact that the TTD140 physical map is not yet complete. Several more primer pairs will be designed to try and amplify the gene from TTD140 and Shamrock, but it appears that *264M08_PKc* is not *Iw1* as it is absent from both non-glaucous varieties which contradicts the dominant nature of the *Iw1* phenotype.

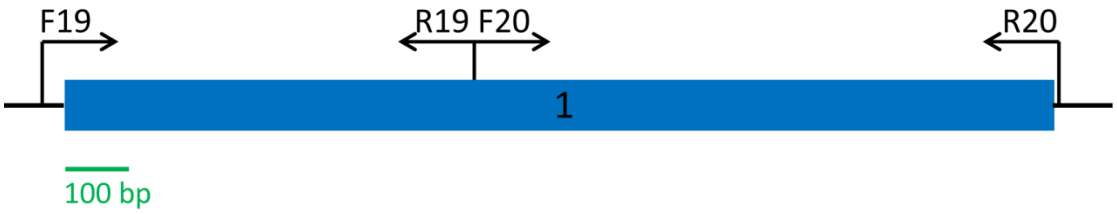


Figure 3.13 Genomic interval of *264M08_CYP450*

The figure represents the genomic interval of the predicted gene *264M08_CYP450*. The blue rectangle represents the single exon, while arrows indicate the position of primer pairs used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

3.2.4.3 *618L11_CYP450* and *1326P06_CYP450*

FGENESH and BLASTX predicted a gene with one exon (CDS: 1,557 bp; genomic: 1,557 bp), which encodes a member of the Cytochrome P450 (P450) family (PLN02169) on the two BACs 618L11 and 1326P06 (gene model identical to Figure 3.13). The sequence of these predicted P450s is 99.9% (identities: 1,555/1,557) and 99.7% (identities: 1,553/1,557) similar to *264M08_CYP450*. No SNPs could be detected between the sequences of the primer pairs F19+R19 and F20+R20 used to amplify *264M08_CYP450* from Langdon, Shango, TTD140 and Shamrock. As a result it can be hypothesized that *618L11_CYP450* and *1326P06_CYP450* are not present in TTD140 or Shamrock because they would have been amplified otherwise beforehand. In addition, as with *264M08_CYP450*, none of these two P450s was detected in any of the sequenced TTD140 BACs, although this could be due to the fact that the TTD140 physical map is not yet complete. However, the three RSL65 P450s (*264M08_CYP450*, *618L11_CYP450* and *1326P06_CYP450*) are 89.2% (identities: 1,389/1,557) similar to a P450 from barley contig 46434 (Section 2.2.5), which indicated a syntenic relationship.

3.2.4.4 *618L11_HlyIII* and *1326P06_HlyIII*

FGENESH and BLASTX predicted a gene with two exons (CDS: 936 bp; genomic: 1,019 bp), which encodes a member of the Haemolysin III family (pfam03006) on the two BACs 618L11 and 1326P06 (Figure 3.14). The CDS of *618L11_HlyIII* is 99.9% similar (identities: 935/936) to *1326P06_HlyIII*, while the complete *618L11_HlyIII* genomic contig is 99.7% similar (identities: 10,979/11,021) to the *1326P06_HlyIII* genomic contig. The two genes and the contigs around them are very similar between the two BACs, but not as similar as would be expected (99.9%), suggesting a very recent duplication.

A BLASTN query against the wheat survey sequence database yielded one hit with 99.7% similarity (identities: 1,016/1,019) from wheat 2BS (contig 5189448); three more 2BS contigs were hit, but they were all truncated. Only a small contig (217bp) was hit from 2DS, while no 2AS contigs were hit. The CDS of both *618L11_HlyIII* and *1326P06_HlyIII* is also 91.8% similar to *Hv_HlyIII*, a gene from barley 2HS contig 46434 (Section 2.2.5). This suggested a syntenic relationship between the wheat and barley genes. *618L11_HlyIII* and *1326P06_HlyIII* were not tested yet for presence/absence in Shango, TTD140 and Shamrock.

Figure 3.14 Genomic interval of *618L11_HlyIII*
The figure represents the genomic interval of the predicted gene *618L11_HlyIII*. The blue boxes represent the two exons, while the red line represents the single intron. The gene model is identical to *1326P06_HlyIII* and *Hv_HlyIII*, a gene from the 2HS barley contig 46434. The green bar indicated the scale.

3.2.5 Analysis of *Ta_1180*

The distal end of the short arm of chromosome 2B has been anchored to *Brachypodium* and rice (Section 2.21.2) and the wheat homolog of *Bradi5g01180* has been linked in all screens to the *Iw1* phenotype. A full length EST contig was obtained for *Ta_1180* from the Triticeae Full-Length CDS Database (TriFLDB (Mochida et al. 2009)). The gene consists of five exons (CDS: 666) and is predicted to encode a CBS_pair domain (cd02205) in the second half of its protein.

A BLASTN query of the full-length EST contig against the wheat survey sequence database yielded two hits to 2BS with 95.6% (identities: 453/474) and 98.4% (identities: 189/192) similarity. One of these hits covers the first two exons (contig 5007649), while the second hit covers the last three exons (contig 5182000). In addition one hit from 2AS (contig 5304717, 94.7% similarity, identities: 631/666) and one hit from 2DS (contig 5208300, 99.8% similarity, identities: 471/472) were obtained; the 2DS contig is truncated inside the first exon.

The combined CDS of the two 2BS contigs (*CS_2BS_1180*) is 99.5% similar (identities: 663/666), 96.9% (identities: 646/666) and 94.6% (630/666) similar to the CDS of *264M08_1180* (RSL65), the CDS of *784K20_1180* (TTD140) and the CDS of *Hv_1180*, respectively. The similarity between the Chinese Spring/RSL65 and TTD140 copies is much lower than would be expected for true homologues from different varieties (>99%). A possible explanation could be the insertion of a large (>7 kb) DNA transposon (DHH_Helios_42j2-1) into the sequence of both *CS_2BS_1180* and *264M08_1180*. The insertion site is 2,324 bp downstream of the second exon of *Ta_1180* and it explains why two separate 2BS hits were obtained from the wheat survey sequence database (Figure 3.15A). It is possible that this insertion disrupted the function of the gene, which is now no longer being conserved in Chinese Spring and RSL65 (Langdon). Both *784K20_1180* (Figure 3.15B) and the barley homologue of *Ta_1180* lack this insertion event, the latter suggesting that the insertion happened after the divergence of wheat and barley.

We have used the InDel marker JIC009 (identical with F21+R21), which was used to identify BAC 784K20 and BAC 264M08 from the TTD140 and RSL65 BAC libraries respectively, to correlate the phenotypes of the TTD diversity panel with the genetic polymorphism seen between TTD140 and Langdon. The phenotypic scores did not correlate with the InDel polymorphism between TTD140 and Langdon. Only twelve accessions shared the TTD140 polymorphism, eight of which originate from the southern TTD population and four with no passport information, but presumably also from the southern population based on their non-glaucous phenotype. The majority of accessions (87) share the Langdon polymorphism but show no correlation regarding phenotype or population. This

suggested that *Ta_1180* is not correlated with the *Iw1* phenotype. However, the polymorphism that JIC009 is based on is not present in Shamrock, thus we cannot conclude that *Ta_1180* is not linked with the phenotype of the TTD diversity panel, but only that marker JIC009 is not linked with it. We will clone the entire *Ta_1180* gene from all members of the TTD diversity panel to elucidate whether it is a valid candidate of *Iw1*. Interestingly we have discovered four accessions that show heterozygosity for marker JIC009. We will test the progeny to elucidate whether these were merely heterozygous plants or whether these accessions possess the two hypothesised haplotypes of the *Iw1* interval.

Table 3.5: Screen of TTD diversity panel using marker JIC009

The InDel marker JIC009 (F21+R21) was used to screen the TTD diversity panel. The phenotypic scores did not correlate with the InDel polymorphism between TTD140 and Langdon. Six accessions were found to be heterozygous. Note that genotypic scores were not obtained for all accessions.

Population	Phenotype	Langdon polymorphism	TTD140 polymorphism	Heterozygotes
northern	*Glaucous*	17	-	2
	Non-glaucous	-	-	-
southern	*Glaucous*	14	2	1
	Non-glaucous	38	6	3
no passport	*Glaucous*	16	-	
information	*Non-glaucous*	2	4	
	Total	87	12	6

A translation of the CDS from *CS_2BS_1180*, *264M08_1180*, *784K20_1180* and *Hv_1180* revealed that the predicted *CS_2BS_1180* and *264M08_1180* (RSL65) proteins are identical (identities: 221/221), while the translated *784K20_1180* (TTD140) sequence has three non-synonymous SNPs with both of them (identities: 218/222). None of these three substitutions is a positive exchange of amino acids (positives: 218/221), but since none of them fall inside the CBS_pair domain there is no evidence that the substitutions would disrupt the protein's function. The translated sequence from *Hv_1180* shows a number of amino acid substitutions in the first third of the predicted protein, but the CBS_pair domain is completely conserved between barley and the three predicted proteins from wheat. This conservation of the key domain (CBS_pair) suggested that the protein can fulfil its normal function in all analysed genotypes, if it is expressed.

We have tested expression of *Ta_1180* via RT-PCR of cDNA gained from flag leaf tissue of Shango and Shamrock after anthesis (after the appearance of the glaucous phenotype in Shango). Bands of correct size were amplified, but sequencing revealed that these bands consisted solely of the A-

genome homoeologue of *Ta_1180*. We will repeat this experiment and include the TTD diversity panel to obtain stronger evidence for the expression pattern of *Ta_1180*.

The results obtained so far do not allow us to rule out *Ta_1180* as a candidate for *Iw1;* however they also do not provide over-whelming evidence for its involvement in the phenotype. We will perform additional experiments as described above to elucidate the role of this gene.

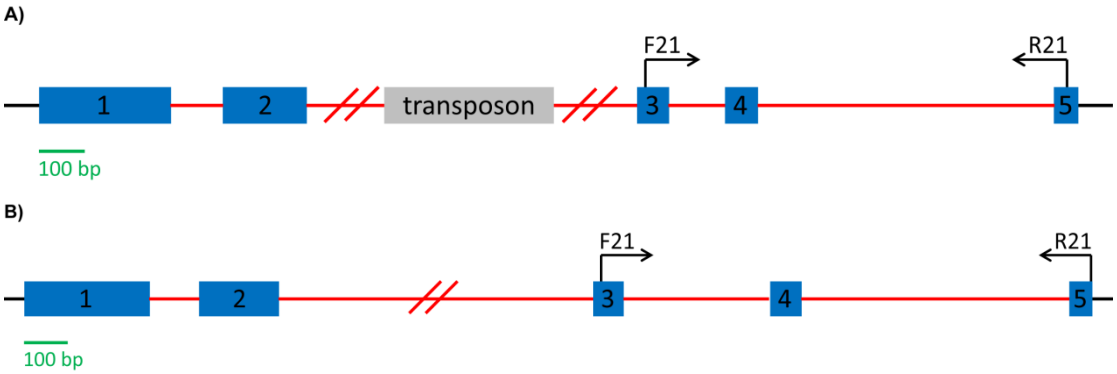


Figure 3.15 Genomic intervals of *Ta_1180* in RSL65 and TTD140

The figure represents the two different genomic intervals of the gene *Ta_1180*. (A) A DNA transposon inserted inside the second intron of *Ta_1180* in RSL65, while (B) the ORF of *Ta_1180* in TTD140 is uninterrupted. Labelled blue boxes represent the exons, red lines represent introns and arrows indicate the position of a primer pair used to identify BAC 784K20 in the TTD140 BAC library and BAC 264M08 in the RSL65 BAC library. The sequence between exons 2 and 3 was shortened here for the purpose of clarity. The green bar indicated the scale.

3.2.6 Analysis of *774P06_WPK*

As was mentioned in the initial analysis of *774P06_WPK*, the gene is predicted to encode two domains that were not found together on one sequence in any public nucleotide or protein database. This uniqueness was reminiscent of the *Iw1* locus, which seems to be missing from other grass and cereal species. Interestingly, *774P06_WPK* is surrounded by an old DNA transposon (DTC_Vincent_42j2-1; CACTA) that has less than 90% similarity to the best hit in TREP. This could explain the unique combination of the two domains. Some transposable elements are known to "acquire" additional sequence during a transposition, which can become fixed if the new combination proves to be beneficial: This is especially true for DNA transposons of the CACTA family (Takahashi et al. 1999, Finn et al. 2010). The two domains predicted to be encoded by *774P06_WPK* are a partial SRPBCC_11 domain (cd08866) and a partial FR_SDR_e domain (cd08958); both domains lack a part of their 3' ends, hence they are being called partial here.

The SRPBCC_11 domain is part of the uncharacterised group of START/RHO_alpha_C/PITP/Bet_v1/CoxG/CalC (SRPBCC) domain superfamily. SRPBCC domains have a deep hydrophobic ligand-binding pocket and bind diverse ligands. The SRPBCC domain here has a high similarity to a Polyketide cyclase / dehydrase and lipid transport family, which is involved in polyketide synthesis as well as lipid binding and transport. The FR_SDR_e domain is a member of the NADB_Rossmann superfamily. The NADB domain is present in numerous dehydrogenases of metabolic pathways and NADB domains often occur together with other domains that are responsible for specifically binding a substrate and catalyzing a particular enzymatic reaction. These are all plausible functions for a protein that involved in the biosynthetic pathway of β-diketones (a class of lipids) or their transport to the plant surface. We hypothesise that one or both of the partial domains of *774P06_WPK* act in a dominant-negative manner by competing for ligands or forming non-functional homo- or heterodimers.

3.2.6.1 Cloning *774P06_WPK* from genomic DNA

Our initial analysis of the *774P06_WPK* gene has revealed that it is present in the non-glaucous varieties TTD140 and Shamrock, but absent from the glaucous cultivars Langdon and Shango; we could not detect homoeologous copies of *774P06_WPK*, so we conclude that there is only one copy of this gene present among our germplasm. However, a paralogue of *774P06_WPK* is present on the long arm of chromosome 6B in Langdon and Shango. We have cloned the entire *774P06_WPK* from TTD140 and Shamrock and also the entire 6BL paralogue from Langdon and Shango using two primer pairs (F22+R22 and F23+R23)(Figure 3.16). The 6BL paralogue could not to be amplified in TTD140 or Shamrock, suggesting that it is absent from these genotypes. The 6BL paralogue contains a STOP codon (TAG) 27 bp before the end of exon 1, indicating a truncated and most likely non-functional protein. The *774P06_WPK* gene from TTD140 and Shamrock contains no premature STOP codon.

The genomic sequence of *774P06_WPK* is 97.7% (identities: 2,182/2,234) similar between TTD140 and Shamrock, while the CDS is 97.9% (identities: 734/750) similar between the two. Shamrock has a 31 bp insertion 147 bp upstream of the transcriptional start site predicted by FGENESH; this 31 bp insertion is also present in the 6BL paralogue of Langdon and Shango. In addition, the *774P06_WPK* sequence from Shamrock indicated a different splice site for the 4th exon (AC instead of AG, Figure 3.16), thus eliminating eleven base pairs from the putative Shamrock transcript. This in turn leads to a frameshift and a premature STOP codon which is predicted to abolish the FR_SDR_e domain, but to leave the partial SRPBCC_11 domain intact. Of the 26 SNPs between the CDS of *774P06_WPK* in TTD140 and Shamrock six SNPs are predicted to cause a non-synonymous exchange of amino acids (identities: 126/132), with three of these amino acid changes being positive substitutions (positives: 129/132). None of the amino acid changes occurs at a predicted ligand-bigand site though, suggesting that the function of the partial SRPBCC_11 domain remains unchanged. The premature STOP codon in Shamrock suggested that if *774P06_WPK* is *Iw1*, then its function stems solely from the SRPBCC domain.

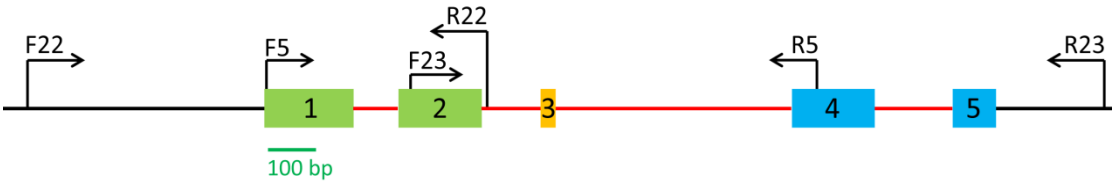


Figure 3.16 Genomic interval of *774P06_WPK*
The figure represents the extended genomic interval of the predicted gene *774P06_WPK*. The green boxes represent the SRPBCC_11 domain encoding exons, while the blue boxes represent the FR_SDR_e domain encoding exons. The orange box represents an exon that does not seem to belong to either of the two domains. The red lines represent introns and arrows indicate the position of primers used to clone the gene in Shango, Langdon, TTD140 and Shamrock. The green bar indicated the scale.

3.2.6.2 RT-PCR expression analysis of *774P06_WPK*

We have tested expression of *774P06_WPK* in TTD140, Shamrock, Shango and Langdon via RT-PCR using primer pair F5+R5 (Figure 3.16) using cDNA generated from leaf sheath tissue after anthesis (Materials and Methods 3.1.2 – 3.1.4). TTD140, Shamrock and Shango showed expression, while Langdon did not, even after repeated attempts. The TTD140, Shamrock and Shango bands were excised and sequenced. This revealed that exon 3, which was predicted by FGENESH, did not exist in the transcript (Figure 3.17 and Figure 3.18). The consequence was that the TTD140 protein was predicted to lose its FR_SDR_e domain, while the same domain was now restored to the predicted Shamrock protein. Furthermore, the absence of "exon 3" has caused STOP codon in Shamrock to disappear; it is unclear where the Shamrock transcript ends now. However, these results suggest that if *774P06_WPK* is *Iw1*, then its function stems solely from the SRPBCC domain.

The Shango transcript appeared to lack the STOP codon inside the first exon, which was identified in the genomic sequence. Furthermore, the Shango transcript had a 99 bp insertion between exons 2 and 3, which truncated exon 2 (Figure 3.17). This insertion led to a premature STOP codon in the Shango transcript, which was predicted to abolish the SRPBCC domain. This Shango transcript was unexpected and its sequence did not match that of the 6BL paralogue.

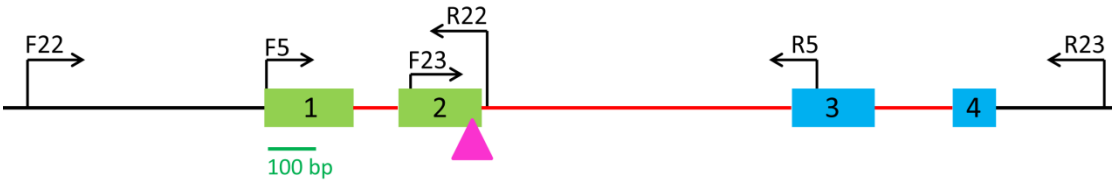


Figure 3.17 Updated genomic interval of *774P06_WPK*
After RT-PCR expression analysis it became clear that the predicted gene model from FGENESH was wrong and that *774P06_WPK* has only four exons. The green boxes represent the SRPBCC_11 domain encoding exons, while the blue boxes represent the FR_SDR_e domain encoding exons. The pink triangle represents an insertion in the transcript of Shango. The red lines represent introns and arrows indicate the position of primers used to amplify the gene in Langdon, Shango, TTD140 and Shamrock. The green bar indicated the scale.

Primer pair F5+R5 was used to amplify this new sequence from genomic Shango DNA. The PCR products were subcloned and two different products were discovered. One was the already known 6BL paralogue, while the other product was similar to *774P06_WPK* from Shamrock, but contained the 99 bp insertion seen in the Shango transcript. This suggested that Shango has two paralogues of *774P06_WPK* (Figure 3.18). We hypothesise that this new paralogue was not discovered before because primer R22 falls inside the 99 bp insertion, the sequence of which is completely dissimilar to the R22 primer sequence. We further hypothesise that the similarity of this new paralogue ends at some point downstream of exon 3 and that primer R23 could not amplify it either.

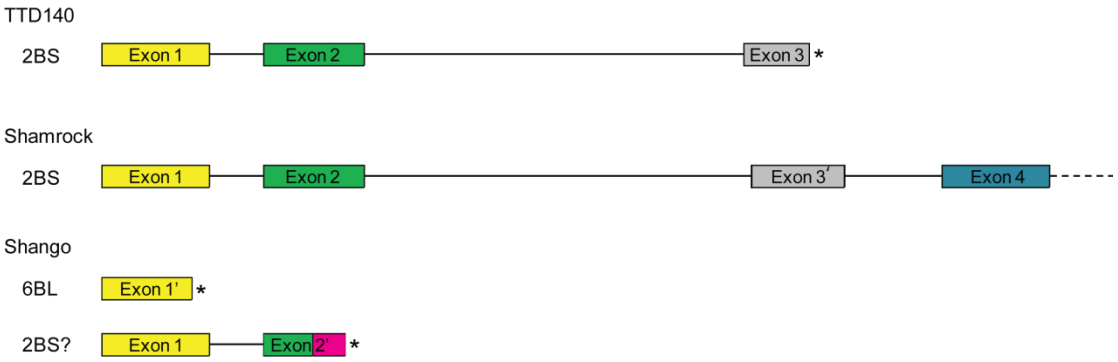


Figure 3.18 Transcripts in TTD140, Shamrock and Shango
This figure represents the transcripts discovered in TTD140, Shamrock and Shango. TTD140 and Shamrock transcripts diverge after exon 2, indicating that the first two exons are essential, but not the remaining exons. Shango has two transcripts; a very short one with a premature STOP codon inside exon 1 and a second transcript with a 99 bp insertion in exon 2, also leading to a premature STOP codon. Phenotypes and expression status are indicated. Note that transcripts were cut off at their STOP codons, represented by (*).

3.2.6.3 Transcriptional start site of *774P06_WPK*

The gene model initially predicted by FGENESH had turned out to be wrong as "exon 3" (Figure 3.16) could not be detected in TTD140, Shamrock, or Shango transcripts. The sequence upstream of the transcriptional start site predicted by FGENESH contains eight ATG codons, which could be alternative transcriptional start sites not correctly predicted by FGENESH. However, seven of these upstream ATG codons are out of frame with the *774P06_WPK* ORF. The remaining ATG codon though, which is located 252 bp upstream of the predicted start site in TTD140, would cause a premature STOP codon in Shamrock and Shango transcripts due to a 31-bp insertion in these two genotypes (Figure 3.19)(section 3.2.4.1). We have tried to design primers upstream of the FGENESH ATG codon, but due to a very high GC content around the alternative ATG, we could not design primers adjacent to it. Thus, we would not be able to distinguish the 5 'UTR of the transcript from an alternative "real" transcript.

Figure 3.19 Alternative transcriptional start site of *774P06_WPK*
An alternative start site to the one predicted by FGENESH would not affect the TTD140 transcript but lead to a premature STOP codon in Shamrock and Shango transcripts. The yellow box indicated the putative sequence belonging to exon 1 if the alternative start site is the correct one. The red triangle represents the location of the premature STOP codon in Shamrock and Shango using the alternative transcriptional start site.

Instead, we used the first two exons of *774P06_WPK*, which encode the partial SRPBCC_11 domain, to query the IPK barley database using BLASTN. The idea was to find genes with a similar domain and exon/intron structure and compare the transcriptional start sites between wheat and barley. We identified one expressed sequence on chromosome 3HL with 93.3% (identities: 336/360) similarity that is predicted to encode a complete SRPBCC_11 domain. Using this information, we also identified the corresponding homologues in wheat on chromosomes 3AL (96.1% similar, identities: 346/360), 3BL (96.1% similar, identities: 346/360) and 3DL (95% similar, identities: 342/360). The alternative transcriptional start site was not conserved in any of these four genes. The transcriptional start site predicted by FGENESH however was also the predicted and conserved start site in all these genes. This suggested that the transcriptional start site predicted by FGENESH for *774P06_WPK* is the correct start site.

3.2.6.4 Allelic diversity of *774P06_WPK*

To assess the allelic diversity of *774P06_WPK* in the TTD diversity panel we used primers F5+R5 (Figure 3.16) to clone the first two exons of the gene in all TTD accessions. Nine different alleles of *774P06_WPK* were detected in the TTD diversity panel (Figure 3.20). Note that the terms TTD and DIC are used interchangeably in this section, but refer to the same species (*Triticum turgidum* ssp. *dicoccoides*).

The majority of accessions (51) shared the allele present in Shamrock, whereas the TTD140 allele is much more rare (present in 13 accessions), although this result could be biased by the composition of the population (Table 3.6). Four alleles were similar to the 6BL paralogue present in Shango and Langdon, with a premature STOP codon at the end of exon 1, but also with characteristic SNPs (DIC55, DIC64, DIC117 and DIC119). The second allele present in Shango was also found in the diversity panel (DIC21), but none of these accessions possessed the 6BL paralogue as well. A single accession with an allele very similar to Shamrock, but not identical with it, was discovered (DIC28b).

However, no clear pattern between the distribution of alleles and the affiliation with either of the two TTD populations (northern and southern) can be observed. The Shamrock allele for example is mostly present in non-glaucous accessions (46), but also in five glaucous accessions (Table 3.6). The same is true for the TTD140 allele. This indicated that allelic diversity alone was not enough to validate *774P06_WPK*, but that we need to look at its expression pattern as well.

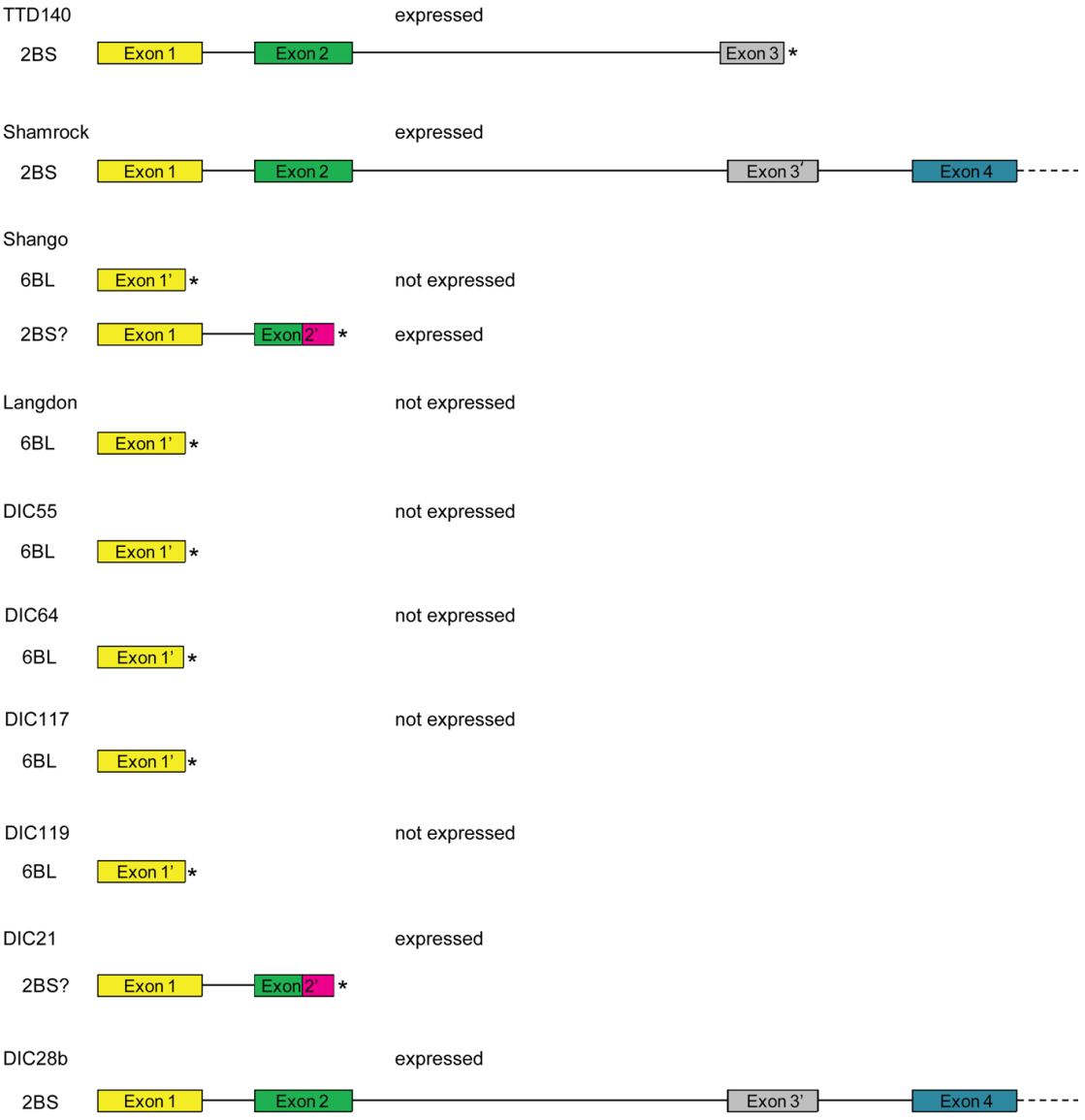


Figure 3.20 ***774P06_WPK*** **alleles**

This figure represents the alleles discovered in 117 accessions of the TTD diversity panel. Only the first two exons were sequenced for each accession. The allele from Langdon is identical with the 6BL allele in Shango. Expression status for each allele is indicated. Note that alleles were cut off at their STOP codons, represented by (*).

Table 3.6: Allelic diversity of *774P06_WPK* in the TTD diversity panel

The first two exons of *774P06_WPK* were cloned in 117 TTD accessions. Nine different alleles were detected, with the Shamrock allele being the most widely distributed one. The phenotype of accessions is indicated at the top, while affiliation to the different TTD populations is indicated at the bottom. Note that Shango type refers to the 6BL paralogue only.

	northern		southern		no passport		
Phenotype	*Glaucous*	*Non-glaucous*	*Glaucous*	*Non-glaucous*	*Glaucous*	*Non-glaucous*	Total
TTD140 tpye	1	-	1	8	-	3	13
Shamrock type	1	-	2	42	2	1	48
Shango type	5	-	-	-	7	-	12
DIC 2 type	-	-	5	-	3	-	8
DIC 21 type	-	-	1	-	6	-	7
DIC 28b type	-	-	-	-	-	1	1
DIC 55 type	13	-	2	1	3	-	20
DIC 64 type	-	-	1	-	-	-	1
DIC 117 type	-	-	2	-	-	-	2
DIC 119 type	2	-	4	-	-	-	6

To test the expression pattern of the various *774P06_WPK* alleles we performed RT-PCR on cDNA generated from leaf sheath tissue after anthesis. To reduce background noise from other transcripts and from genomic DNA, we used a new primer pair (WPK_01) that was designed at the junctions between exons 1 and 2 and between exons 3 and 4. The different *774P06_WPK* alleles were grouped into two categories: Functional alleles have no STOP codon within the first two exons (TTD140, Shamrock, DIC28b), while non-functional alleles do (Shango, DIC2, DIC21, DIC55, DIC64, DIC117, DIC119). The term “functional” and “non-functional” are used loosely to describe the putative status of the SRPBCC_11 domain (exons 1 and 2).

Table 3.7: RT-PCR expression analysis of *774P06_WPK* alleles

Expression was tested on cDNA generated from leaf sheath tissue after anthesis using RT-PCR. *774P06_WPK* alleles were grouped into two categories: Functional alleles have no STOP codon within the first two exons, while non-functional alleles do. (+) indicated expression while (-) indicated lack of expression.

Phenotype	Allele	Expression	Total
Glaucous	Non-functional	-	44
	Functional	-	7
Glaucous	Non-functional	+	3
	Functional	+	-
Non-glaucous	Non-functional	-	-
	Functional	-	7
Non-glaucous	Non-functional	+	1
	Functional	+	39

The correlation between phenotype, allele and expression status was very good (Table 3.7). Out of 54 glaucous accessions analysed, 44 showed no expression of their non-functional alleles and more importantly seven glaucous lines with functional alleles lacked expression. Three glaucous lines showed expression of their non-functional alleles; all three accessions carried the DIC21 allele (Figure 3.20). Of the 47 non-glaucous accessions, 39 accessions with functional alleles showed expression, while seven accessions with functional alleles lacked expression. This correlation is almost perfect with the exception of a single non-glaucous accession that showed expression of a non-functional allele. No glaucous accession with an expressed copy of a functional allele was discovered. We hypothesise that the seven non-glaucous accessions with functional alleles lacking expression are likely also lacking wax-producing genes. We are in the process of making test crosses with these accessions (and all other non-glaucous accessions); by crossing these accessions to a glaucous cultivar like Langdon we should be able to tell by the phenotype of the F_1 generation whether these accessions carry *Iw1* or lack wax-producing genes.

3.2.6.5 Genetic complementation of Fielder with *774P06_WPK*

Because of the strong body of evidence obtained via the allelic diversity and RT-PCR expression analysis we decided to validate *774P06_WPK* via genetic complementation. The glaucous spring wheat cultivar Fielder is used regularly and with high efficiency for transformation experiments at NIAB. A set of vectors exist at NIAB into which constructs can be inserted via homologous recombination (Gateway cloning). We decided to send two constructs for complementation. One would be the genomic construct with as much sequence upstream of the transcriptional start site as was at our disposal (1,115 bp); this was to improve our chances of including the native promoter within the construct. The second construct would be the CDS of *774P06_WPK*, which would be recombined behind the actin promoter of rice in the transformation vector. Primers were designed to amplify the required DNA/cDNA sequences, with five extra base pairs in front of the CDS construct to ensure that it is in frame with the actin promoter. The constructs were sent to NIAB where the team led by Dr. Emma Wallington performed the transformation and the subsequent steps to recover primary transformants from transformed callus cells.

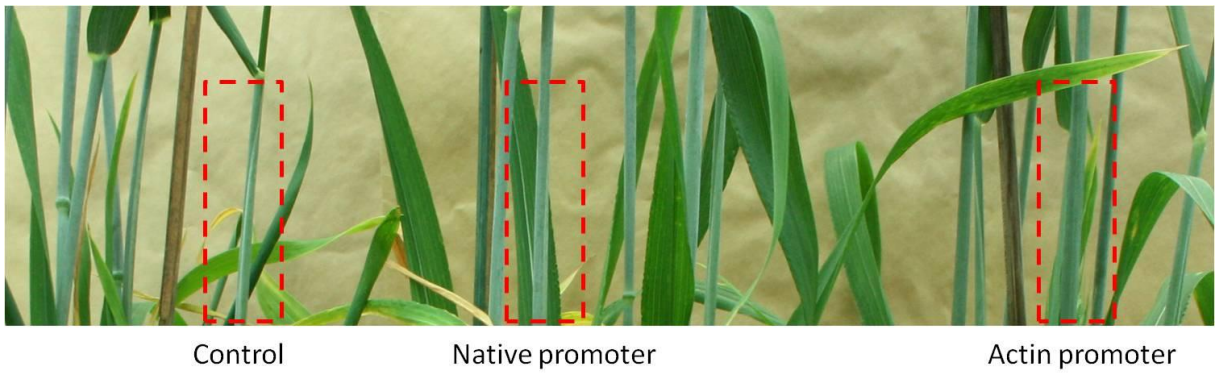


Figure 3.21 Primary transformants of *774P06_WPK*
Two constructs, one genomic and one CDS construct, were sent for transformation into glaucous spring wheat cultivar Fielder at NIAB. Primary transformants were sent back to JIC. All transformants displayed a glaucous phenotype on the leaf sheath (marked by red box), the exposed of the peduncle and the spike. From left to right in this picture: Primary transformant with no insert ("Control"), primary transformant with genomic construct ("Native promoter"), primary transformant with CDS construct ("Actin promoter").

Approximately 20 primary transformants for each construct were generated and sent back to us. The insert copy number was assessed by qPCR at NIAB and all transformants had a least one copy of the respective insert. The plantlets were grown up in a controlled environment room (CER). However, after anthesis all plants were displaying a strong glaucous phenotype (Figure 3.21). We collected leaf sheath and flag leaf tissue; the leaf sheath tissue for RNA extraction and the flag leaf tissue for wax extraction. Using the same primers as for the RT-PCR analysis of the TTD diversity panel (WPK_01), we tested the expression of twelve primary transformants from each construct. Only two transformants carrying the native construct did not show expression, while all other transformants did show expression (Figure 3.22). One out of seven control plants (no insert as determined by qPCR) also displayed expression of *774P06_WPK* (Figure 3.22, Lane 29). We are going to extract leaf wax from these transformants in the coming weeks to test the composition of the surface wax layer. This is done to ensure that all β-diketones aliphatics and other waxes, which are present on wild type Fielder plants, are also present on the transformants. However, the lack of a non-glaucous phenotype coupled with expression of the constructs in most cases indicated that *774P06_WPK* is not *Iw1*.

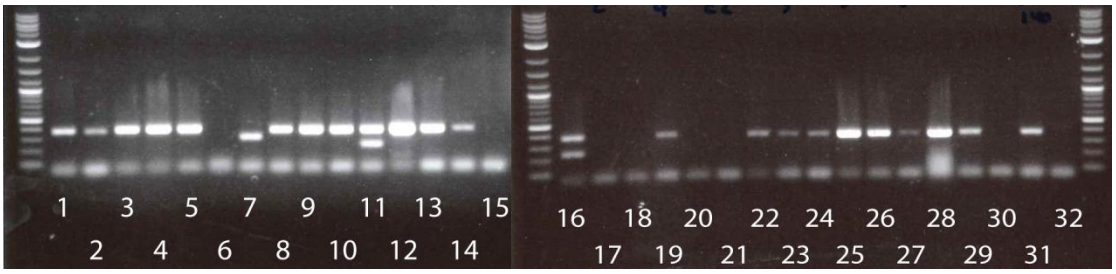


Figure 3.22 Expression of *774P06_WPK* in primary transformants
Expression of *774P06_WPK* was tested in the primary transformants using RT-PCR. CDS-construct (1-5, 8-14), genomic construct (16-19, 22-28), Control without insert (6-7, 15, 20-21, 29-30), TTD140 (31), Langdon (32).

3.2.7 Candidate gene discussion

This section will revisit the predicted genes from the *Iw1* interval that are still considered as candidate genes and elucidate their potential to fit in with the role of *Iw1*.

3.2.7.1 *305N15_CYP450*

305N15_CYP450 is predicted to encode a member of the Cytochrome P450 (P450) monooxygenase superfamily, specifically a flavonoid 3'-monooxygenase. P450s are present in all living organisms and are known to be active in a wide range of biosynthetic and detoxicative pathways. P450s in plants participate in the synthesis of lignins, UV protectants, pigments, defense compounds, fatty acids, hormones, and signaling molecules (Schuler and Werck-Reichhart 2003). Several hundred P450 species are known to be present in mono- and dicot plants, which has hampered the identification and functional characterization of P450s due to protein redundancies (Nelson et al. 2004).

P450s in plants have been discovered to form operon-like clusters that are involved in species-specific metabolic functions like the production of benzoxazinoids in maize (Frey et al. 1997) or avenacin in oat (Qi et al. 2004). Due to the wide range of metabolic functions performed by P450s it is easy to envisage *Iw1* to be a P450.

The most likely scenario to inhibit β-diketone production or transport would be the synthesis of a signaling molecule or hormone that triggers the inactivation of the genes/proteins involved with β-diketone production or transport. Alternatively, mutations at key ligand binding sites or within the catalytic domain of the enzyme might change or abolish its function or it could compete for precursor molecules and ligands with the genes/proteins involved in β-diketone production, although

based on the dominant effect of *Iw1* and the absolute lack of β-diketone aliphatics in non-glaucous *Iw1* lines these scenarios seem less likely.

We will perform more experiments to validate or rule out this candidate in the coming weeks. It is also possible to imagine this enzyme as a target of *Iw1*.

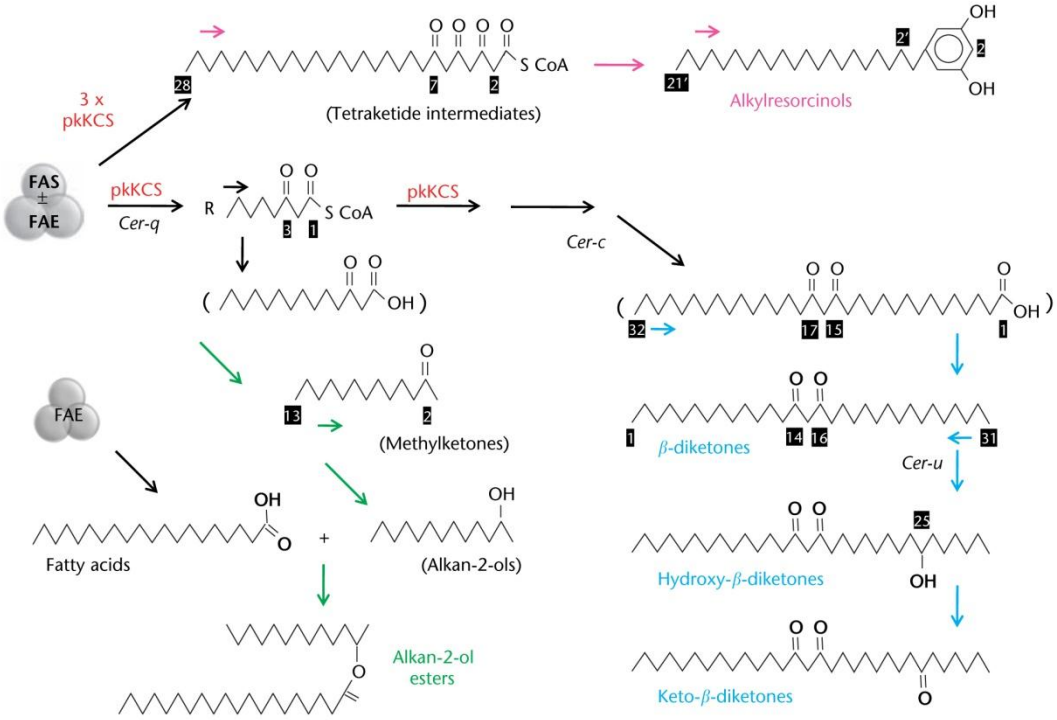


Figure 3.23 Biosynthetic pathways of lipid components
Acyl chains from the FAS pathway get exported to the ER where they become targets for various types of FAE complexes. This figure displays a pathway present in many *Gramineae* species, producing amongst others β- and hydroxy-β-diketones. Figure taken from von Wettstein-Knowles (2012).

3.2.7.2 *305N15_CHS*

305N15_CYP450 is predicted to encode a member of the Chalcone and stilbene synthases (CHS) family, which are plant-specific polyketide synthases (KCSs), also called type III PKSs. This is a class of proteins produces different compounds dependent on the starter molecule. They act by performing carboxylative condensation reactions to extend polyketide chains.

This type of enzyme was mentioned in the Introduction to chapter 1, where they were described to perform the same function as fatty acid elongase (FAE) multienzyme complexes, but that in doing so they were taking 'shortcuts' resulting in carbon chains with double bonds, hydroxy- or keto-groups remaining inside the carbon chain. These molecules act as precursors for a three-branch pathway,

which either leads to alkylresorcinols, alkan-2-ols, alkan-2-ol-esters or to β-diketones and hydroxy-β-diketones (Figure 3.23). The type of 'shortcut' taken by individual KCSs and the resulting change to the carbon chain (double bond, hydroxy- or keto-group) depends on the enzyme (von Wettstein-Knowles 2012).

While this gene is likely involved in the biosynthetic pathway of β-diketone aliphatics, it is difficult to imagine how this gene could be *Iw1*. A change in expression or the mutation of a ligand binding or catalytic site would not explain the dominant nature of the *Iw1* phenotype or the complete absence of β-diketone aliphatics in cuticular waxes. KCSs are single enzymes that do not form homo- or heterodimers, so we also can exclude the possibility of the formation of non-functional dimers. This gene is in fact more likely to be a target of *Iw1* function rather than *Iw1* itself. We will test this candidate further over the coming weeks.

3.2.7.3 *23N19_CYP450*

23N19_CYP450 is predicted to encode a member of the Cytochrome P450 (P450) superfamily, specifically a member of the fatty acid (omega-1)-hydroxylase/midchain alkane hydroxylase family. This protein here predicted to have a homologous function to the *MAH1* gene from *Arabidopsis thaliana*. *MAH1* is involved in converting alkanes into secondary alcohols and possibly even ketones (von Wettstein-Knowles 2012). The *MAH1* pathway is not directly related to β-diketone synthesis; in fact β-diketones have not been identified in Arabidopsis at all. However, we do not know the targets of this P450, which could be involved in a different biosynthetic pathway in wheat than in Arabidopsis. This has been observed for the *VRN* genes in wheat, which are homologous to genes in Arabidopsis but function in a different manner (Yan et al. 2003, Yan et al. 2004). Thus *23N19_CYP450* is still a valid and good candidate for *Iw1*.

As was mentioned for *305N15_CYP450*, mutations at key ligand binding sites or within the catalytic domain of the enzyme might change or abolish its function in non-glaucous lines. Likewise, this enzyme could compete with the genes/proteins involved in β-diketone production for ligands. However, these scenarios are not very likely considerin the dominant effect of *Iw1* and the absolute lack of β-diketone aliphatics in non-glaucous *Iw1* lines. Again, this gene seems more likely to be a target for *Iw1* function rather than *Iw1* itself.

3.2.7.4 *23N19_FAE1*

23N19_FAE1 is predicted to encode a CHS-like domain similar to *305N15_CHS*. The CHS-like domain consists of two domains; a FAE1_CUT1_RppA domain and an ACP_syn_III_C domain. The insertion of a LTR retrotransposon of the *Copia* family disrupts the CDS and is predicted to abolish the ACP_syn_III_C domain, but leaving the FAE1_CUT1_RppA domain intact. The latter is a member of the type III polyketide synthases, fatty acid elongases and fatty acid condensing enzymes and likely to be involved in the biosynthetic pathway of β-diketone aliphatics.

But as was mentioned for *305N15_CHS* it seems unlikely that an enzyme of the β-diketone pathway is *Iw1* itself, as mutations in the catalytic domain or ligand binding sites cannot explain the dominant nature of the gene. However, the insertion of a LTR retrotransposon inside *23N19_FAE1* might explain the *Iw1* phenotype. We would hypothesise that expression of *23N19_FAE1* would lead to transcription of the reverse transcriptase encoded in the retrotransposon. These transcripts might be targeted by small RNAs, resulting in transcriptional gene silencing (TGS). We would further hypothesise that this leads to the silencing of homologous promoters, which control genes of the β-diketone biosynthetic pathway, via trans-TGS (Mette et al. 2000).

We would need to test expression and the methylation status of *23N19_FAE1* to determine whether it is active or not and whether it might be inactive due to TGS-induced methylation. We could also test the TTD diversity panel for presence or absence of the inserted retrotransposon inside *23N19_FAE1* to try and correlate the geno- and phenotypes. We will pursue this gene more over the coming weeks.

3.2.7.5 *23N19_LRR*

23N19_LRR1 is predicted to encode a partial Leucine Rich Repeat (LRR) domain, which is usually associated with Nucleotide Binding Site (NBS) domains in disease resistance proteins (DeYoung and Innes 2006). These proteins detect conserved pathogen-associated molecular patterns (PAMPs) of pathogens and elicit a number of defense response including hypersensitive response (HR) triggered cell death.

It is difficult to imagine how such a protein is involved in the non-pathogen elicited production or inhibition of β-diketone aliphatics. In addition, this predicted gene only encodes a partial LRR domain, suggesting that it might be a pseudogene with no function. We will perform a few more

experiments, but it seems that this sequence can only be amplified from TTD140, so it seems likely that this gene is not *Iw1*.

3.2.7.6 *784K20_PFF*

784K20_PFF is predicted to encode a protein of unknown no similarity to other proteins, which makes it impossible to say anything about the putative function of this gene. However, since we hypothesise that since *Iw1* is unique to wheat but absent from modern cultivars we will not necessarily be able to find homologous proteins in any public database. However, it seems that *784K20_PFF* is only present in TTD140, but not in Shamrock, which suggests that it is not *Iw1*. We will test this hypothesis in the coming weeks.

3.2.7.7 *Ta_1180*

Ta_1180 is predicted to encode a cystathionine-β-synthase (CBS)_pair or Bateman domain. This domain consists of two single CBS domains that associate and form a cleft for ligand binding of adenosyl containing molecules like adenosyl monophospate (AMP) or adenosyl triphosphate (ATP) (Scott et al. 2004, Ignoul and Eggermont 2005).

CBS domains usually occur in conjunction with other functional domains, but five CBS-pair encoding genes with no other attached domains have been reported in rice (Kushwaha et al. 2009). The function of these CBS_pair domains in plants remains unclear. Single CBS domains have been associated with the homeostasis of the cell's redox status by regulating H_2O_2 levels inside the cell (Yoo et al. 2011).

It is unclear how this gene might be affecting the synthesis or inhibition of β-diketone aliphatics. A possible explanation might be the binding of an adenosyl containing signal molecule or the triggering of a signal molecule cascade by binding adenosyl containing molecules. Nonetheless, this gene was always linked to the *Iw1* phenotype.

We have performed a set of experiments but we cannot rule out *Ta*_1180 yet. We will clone the gene from the TTD diversity panel and perform a thorough analysis of its expression pattern via RT-PCR in the coming weeks.

3.2.8 EMS mutagenesis of TTD140 and Shamrock

To create an alternative approach for identifying *Iw1*, 3,000 seeds of both TTD140 and Shamrock were treated with 0.85% Ethyl methanesulfonate (EMS, Materials and Methods 3.1.5). Plants were developed to the M_2 generation, but due to severe sterility in the M_1 of TTD140, no mutants were detected in the few M_2 families. We have obtained seed from 1,985 M_2 families of the mutagenised Shamrock population. These were sown out in the field in 2012 and phenotyped in the summer of 2013 for the presence of glaucous plants.

We could detect glaucous plants in 35 M_2 families. Tissue was collected from these putative mutants for RNA and DNA extraction. Tissue from non-glaucous siblings that would act as control plants was also collected. We used the two presence/absence markers JIC018 and JIC019 (Table 2.6) to test whether the putative mutants are truly of Shamrock genotype or whether they are contaminants. Only family 33, which consists of a single mutant plant, showed amplification for both markers, while no products from either marker could be amplified from the remaining 34 putative mutants. We have performed RT-PCR on cDNA extracted from flag leaf tissue of the single family 33 plant using marker WPK_01 and confirmed expression of this Shamrock specific gene. We will now have to wait for the phenotype of the putative mutant for the final confirmation.

If the mutant is glaucous it will be crossed to Shamrock to generate a F_2 population that is segregating for glaucousness. We will use the EMS-induced mutations that should occur at an approximate interval of 40 kb (Uauy et al. 2009) to generate markers and map the glaucous trait. This should allow us to reduce the size of the *Iw1* interval and thus reduce the size of candidate genes that we need to test. Parallel to this, the sequence of candidate genes inside the *Iw1* interval will be obtained from the mutant line to compare with the TTD140 and Shamrock sequences and identify potential EMS mutations that could underlie the loss of the non-glaucous phenotype.

3.2.9 RNA-seq of BC_4 Robigus NILs

Based on the physical map so far and the dominant nature of *Iw1*, we hypothesise that the gene is either not expressed (or only to a low level) or absent from glaucous germplasm. By sequencing the transcriptome of highly isogenic BC_4 NILs we should be able to detect differences in transcript abundance between a glaucous and non-glaucous NIL pair to address the question of expression in our hypothesis. A BC_4 NIL pair of Robigus was selected and tissue from the leaf sheath and peduncle was collected four weeks after anthesis. Total RNA was extracted from these tissues (Materials and

Methods 3.1.2) and handed to TGAC for library preparation and subsequent sequencing on an Illumina HiSeq to generate 100 bp paired end reads (Materials and Methods 3.1.6).

The run finished only recently and we did not have time to analyse the data before submission of this thesis. We plan to map the reads of the glaucous and non-glaucous lines to the predicted ORFs and genomic sequences inside the *Iw1* interval. This relatively simple analysis should reveal whether the predicted genes are expressed at all (in these tissues and at that point in time) and should also reveal differences in expression. We have used three biological samples per genotype/tissue allowing us to test these differences statistically. As the physical map expands new gene models will be analysed using these criteria.

We also plan to map our RNA-seq reads to transcriptome data from durum wheat cultivar Kronos (Krasileva et al. 2013) with a high specificity (>99%) to distinguish reads from the A- and B-genome: Kronos is a glaucous variety and following our hypothesis from above its transcriptome should not contain *Iw1* transcripts. The unmapped reads can be mapped to Kronos a second time, but with only 95% similarity requirement to eliminate homoeologous reads from the D-genome. The remaining unmapped reads will be assembled *de novo* for both glaucous and non-glaucous NILs, followed by reciprocal mapping of reads against the assemblies. This should hopefully identify transcripts that are either unique to the non-glaucous NIL or that are significantly more abundant in the non-glaucous NIL than in the glaucous NIL. These transcripts represent candidates of *Iw1*, which will then be mapped and further validated.

3.2.10 Discussion

This chapter described the continuing efforts in the identification and validation of candidate genes for *Iw1*. A panel of 124 *Triticum turgidum* ssp. *dicoccoides* (TTD) accessions was phenotyped for glaucousness. For 96 of these accessions we have information on their collection site. Out of these, all 24 accessions from the northern wild emmer population were glaucous, whereas the 72 accessions from the southern population were either glaucous or non-glaucous. This suggested that non-glaucousness controlled by *Iw1* emerged in the southern wild emmer population only. Glaucousness was defined here as showing a waxy bloom on at least one aerial plant organ. It is possible that weak alleles of *Iw1*, which inhibit glaucousness only in a defined tissue, exist in the northern and southern populations, although it is equally possible that these partial phenotypes are caused by misexpression of *W1* or genes of the network. Test crosses of 22 non-glaucous TTD accessions with glaucous cultivar Langdon were made already and the F_1 generation is currently growing. The remaining non-glaucous lines will be crossed to Langdon as well. By analysing the F_1 phenotype we will be able to tell which accession carries the dominant inhibitor *Iw1* and which ones are mutants of genes controlling glaucousness. The latter will be useful for studying cuticular wax of non-glaucous plants lacking *Iw1*, as well as for the cloning and functional assessment of glaucousness producing genes like *W1*.

The TTD diversity panel will serve as a validation tool for the genes predicted inside the *Iw1* interval of TTD140 and RSL65. We know that both TTD140 and Shamrock contain *Iw1*, so all candidate genes are first tested for presence/absence in these two genotypes. We have possibly ruled out two genes with this method and hope to rule out two more using more independent primer pairs. However, six genes could not be ruled out by this method and we need to perform additional experiments to assess their validity as candidate genes.

One of these six candidate genes, *774P06_WPK*, was already tested further. The juxtaposition of two domains that could not be found together in any public database was reminiscent of the apparent uniqueness of *Iw1* among grass and cereal species. In addition, these two domains were likely to influence the biosynthetic pathway of β-diketones. The *774P06_WPK* gene was less well conserved between TTD140 and Shamrock (97.8%) than expected (>99%). However, by comparing the putative proteins the SNPs were shown to be outside of any ligand-binding or catalytic sites. An initial expression analysis via RT-PCR revealed that the initial gene model predicted by FGENESH was wrong, which ruled out the second predicted domain of *774P06_WPK* to be essential for its function. An alternative transcriptional start site was ruled out by comparing the remaining domain (SRPBCC_11) to homologous sequences from wheat and barley. The allelic diversity of *774P06_WPK*

among the TTD diversity panel correlated very well with the phenotypes (Figure 3.20 and Table 3.6), as did the expression analysis via RT-PCR (Table 3.7). This was a strong body of evidence so we decided to test *774P06_WPK* via genetic complementation and sent two constructs to NIAB to be transformed into the glaucous spring wheat cultivar Fielder. One of these constructs consisted of the genomic *774P06_WPK* sequence. This construct also contained 1,115 bp of sequence upstream of the transcriptional start site to encompass the native promoter. The second construct consisted of the *774P06_WPK* CDS, which would be expressed the actin promoter of rice. Both constructs were successfully transformed at NIAB by the team of Dr. Emma Wallington. Twenty primary transformants for each construct were sent back to JIC. However, all of the transformants displayed a glaucous phenotype, suggesting that *774P06_WPK* is not *Iw1*. This was further confirmed by RT-PCR expression analysis, which showed the constructs to be expressed in 22 out of 24 primary transformants tested. We will extract cuticular waxes from these transformants to elucidate whether they produce the entire set of β-diketone aliphatics observed in 'regular' glaucous lines or maybe just a subset. This is to ensure that we transformed the fully functional *774P06_WPK* gene and not just one part of it which only inhibits production of a single wax class. This is also important based on the hypothesis that *Iw1* might be composed of several genes, similar to the *cer-cqu* locus from barley but with opposite function, which would require two or more transgenic events.

The apparent failure of *774P06_WPK* as a candidate gene of *Iw1* was very disappointing, especially after all the evidence that was collected beforehand. This highlights a specific problem with the *Iw1* interval. It was proposed in chapter 2 that two distinct haplotypes exist for the *Iw1* interval. This was hypothesised based on a lack of recombination across the *Iw1* interval and an apparent divergence in gene content between the TTD140 and RSL65 physical maps. A consequence of the haplotype theory is that allelic diversity and expression profiles of genes in non-glaucous TTD accessions will probably be highly similar because of linkage to the haplotype and vice versa for glaucous accessions. This problem has already been identified in genes from the TTD140 physical map that are absent from the glaucous Langdon and Shango cultivars, while genes from the RSL65 map are absent from the non-glaucous TTD140 and Shamrock (Table 3.3 and 3.4).

One way to circumvent this problem was the mutagenesis of 3,000 seeds each from TTD140 and Shamrock using EMS. The idea was to knock out *Iw1* and recover glaucous plants in the M_2 generation. However, the TTD140 population suffered from severe sterility and did not yield any glaucous M_2 mutants. In the Shamrock population 35 glaucous M_2 families were detected out of 1,985 M_2 families. However, only a single family, which also consisted of a single plant, showed Shamrock specific amplification of two presence / absence markers inside the *Iw1* interval. This

suggested that either the dosage of EMS was too low (0.85%), which is unlikely as we observed multiple developmental phenotypes inside the population, or that something about the structure of *Iw1* makes its' mutagenesis more difficult; the gene could be quite small, making it harder to "hit" or alternatively, *Iw1* could be tandem-duplicated. The latter would require a mutation event in two adjacent genes in a single plant, which has a very low chance of occurring. The lack of glaucous mutants was observed before with another EMS mutagenised Shamrock population. However, that population only consisted of 1,200 M_2 families so the failure of the experiment was assigned to the low number of families. Whatever the reason for the low number of glaucous mutants in these two EMS populations, we seem to have identified one glaucous mutant. This mutant will be crossed to Shamrock to generate a F_2 population that is segregating for glaucousness. We will use the EMS-induced mutations that should occur at an approximate interval of 40 kb (Uauy et al. 2009) to generate markers and map the glaucous trait. Candidate genes will be sequenced from the mutant line and compared with sequences from TTD140 and Shamrock to identify EMS mutations that might be causative of the glaucous phenotype.

Seven genes are still candidates for *Iw1*, although in two cases (*23N19_LRR* and *784K20_PFF*) it is unclear how these genes could function as inhibitors of β-diketone production. The function of *Ta_1180*, which encodes a CBS-pair domain, is also unclear as there is little information about the putative role of these proteins in plant. The CBS_pair domain is known to bind adenosyl containing molecules, so it might be implicated in some form of signaling cascade that triggers the inactivation of the β-diketone biosynthetic pathway.

Two of the other candidates are predicted to encode Cytochrome P450 (P450) proteins. This class of proteins is involved in a wide range of metabolic pathways. One of these P450s (*23N19_CYP450*) has similarity to a known protein from Arabidopsis (*MAH1*) which is involved in synthesis of secondary alcohols and ketones from alkanes. This pathway is not related to the β-diketone pathway, making this P450 a less likely candidate. Little information exists for the other P450 (*305N15_CYP450*), but a possible mechanism could be the synthesis of a signal molecule or hormone that triggers the deactivation of the β-diketone biosynthetic pathway.

The remaining two candidate genes are predicted to encode CHS-like domains, but only *305N15_CHS* encodes a complete domain. This gene is likely to be involved somewhere in the β-diketone pathway, as it encodes a type III polyketide synthase (PKS). These enzymes extend carbon chains, but take 'shortcuts' in doing so, which results in double bonds, hydroy- or keto-groups to be left inside the chain, which leads, amongst others, to the production of β-diketones. However, it is unclear how a

mutation in this gene could cause the dominant *Iw1* phenotype as it has no apparent regulatory functions.

The other candidate gene predicted to encode a CHS-like domain (*23N19_FAE1*) has a LTR retrotransposon inserted. This insertion is predicted to abolish a sub-domain of the CHS-like domain (ACP_syn_III_C), while the other sub-domain (FAE1_CUT1_RppA) is predicted to stay intact. The FAE1_CUT1_RppA domain is a member of the type III polyketide synthases, fatty acid elongases and fatty acid condensing enzymes and likely to be involved in the biosynthetic pathway of β-diketone aliphatics. However, the same applies for this protein as for the one encoded by *305N15_CHS*: It lacks a regulatory domain that sufficiently explains the dominant *Iw1* phenotype. However, it is possible that not the gene itself is performing that function, but rather the retrotransposon inserted in it. Expression of the *23N19_FAE1* promoter would lead to transcription of the reverse transcriptase gene encoded in the retrotransposon, which could trigger a transcriptional gene silencing (TGS) response in which the locus becomes methylated and silenced. Furthermore, promoters homologous to the *23N19_FAE1* promoter could be silenced via trans-TGS (Mette et al. 2000). Since *23N19_FAE1* is likely involved in the β-diketone pathway based on its PKS encoding domain, homologous promoters would belong to other PKSs which would trigger the methylation of these specific loci and thus the inhibition of β-diketone synthesis. This hypothesis would explain the dominant nature of *Iw1* and its uniqueness among other grass and cereal species. To test this hypothesis we would need to investigate the pattern of transcripts from glaucous (*iw1*) and non-glaucous (*Iw1*) plants.

With the creation of highly isogenic BC_4 NILs it is now possible for us to analyse the transcriptome of glaucous and non-glaucous NIL pairs, which should only differ in the *Iw1* interval. Various different ways of analysing this vast amount of data are conceivable and some have been outlined in section 3.2.4. The highly isogenic background of the NILs would make it even possible to “simply” perform a *De-novo* assembly of the reads from glaucous and non-glaucous lines (also separated by tissues) and identify unique transcripts by reciprocal mapping of reads to the assemblies. This would also allow us to test the TGS-hypothesis for *23N19_FAE1*, as transcripts with a 5’ UTR homologous to the one from *23N19_FAE1* should be less abundant in the *Iw1* NIL.

To summarize, we have identified and started to validate predicted candidate genes from the TTD140 and RSL65 physical maps and will continue to do so as the maps are further expanding. We have examined one particular gene that seemed promising based on its allelic diversity and expression profile more closely, but which did not confer a non-glaucous phenotype upon

transformation into glaucous wheat. This highlights the need for additional validation methods. We have likely obtained a single EMS-induced mutant of *Iw1* in hexaploid wheat that, should it be proven to be true, will be of immense value for future validation efforts of candidate genes and to reduce the size of the *Iw1* interval via recombination. In addition, the transcriptome of a pair of highly isogenic BC_4 NILs has just been sequenced and will be assembled and analysed over the coming weeks to predict and validate candidate genes of *Iw1*.

The combination of these approaches should lead to the successful cloning of *Iw1* in the near future.

General Discussion

The aim of this study was to clone and validate the *Inhibitor of wax 1* (*Iw1*) locus and to characterize its effects on the composition of surface waxes in wheat, as well as its effects on yield, green-canopy duration and several other physiological traits.

Aerial surfaces of *Iw1* lines lack tubular waxes

We have studied the effects that *Iw1* has on the composition and quantity of the epicuticular wax layer using scanning electron microscopy combined with a detailed biochemical analysis. The flag leaf and peduncle surfaces of glaucous varieties like Shango or Langdon are covered with tubular shaped wax crystals, which almost completely obscure the stomata and other types of wax crystals. There is also some variation in the distribution of wax crystals on the adaxial side of the flag leaves, where alternating strips of thicker and thinner layers of wax, mostly consisting of a platelet shaped wax, are visible. This is reminiscent of the arrangement of mesophyll and vascular tissue, which are also organized in alternating strips along the long axis of the leaf (Esau 1953). This suggests that the thick layers of platelet shaped wax coincide with the vascular tissue running underneath, possibly providing added insulation. In contrast, flag leaves and peduncles of non-glaucous varieties like Shamrock and TTD140 completely lack the tubular shaped wax crystals seen so abundantly on glaucous plant surfaces. Instead, the non-glaucous plant surfaces are almost completely devoid of visible wax crystals apart from the adaxial side of flag leaves. Here, as in the glaucous varieties, alternating strips of thick and thin layers of platelet shaped wax can be observed running along the leaf axis parallel to the rows of stomata and thus to the vascular tissue of the leaf.

Iw1 prevents formation of β-diketone aliphatics

The tubular and platelet shaped wax crystals had been observed already 50 years ago and the presence of the tubular shaped wax was linked to the occurrence of β-diketones (Hall et al. 1965). The platelet shaped wax on the other hand was hypothesized to be linked to the presence of primary alcohols (Netting and Wettstein-Knowles 1973). We have performed a preliminary analysis of flag leaf wax extracts obtained from glaucous (Shango) and non-glaucous (Shamrock) varieties using thin layer chromatography (TLC) and spectrophotometry; we have found a significant and severe reduction in the amount of β-diketones in wax extracts from non-glaucous flag leaves. This was followed by an extensive gas chromatography/mass spectrometry (GC-MS) analysis of flag leaf and peduncle wax extracts from Shango and Shamrock. In total, 53 compounds were identified in the wax extracts and 26 of these were quantified. Wax extracts from flag leaves and peduncles of Shango consisted mostly of β-diketones and hydroxy-β-diketones, which could not be discovered, not even in traces, in the analogous wax extracts of Shamrock. Wax extracts from flag leaves also contained high

amounts of C_{28}-primary alcohols (C_{28}POHs) in both Shango and Shamrock, which was not observed in wax extracts from peduncles. These results were confirmed by analysing wax extracts from glaucous and non-glaucous pairs of BC_2 NILs using GC-MS. This indicates that the rod-like tubes seen on glaucous plant surfaces are indeed β-diketone aliphatics. The high abundance of C_{28}POHs in wax extracts from flag leaves of both varieties indicates that these compounds form the platelet shaped wax observed mainly on the adaxial side of flag leaves.

The timing of glaucousness

We performed a time-course experiment where we followed wax deposition in pairs of glaucous and non-glaucous BC_2 NILs over the course of their growing cycle. This analysis showed that prior to stem extension wax of vegetative leaves is primarily composed of C_{28}POHs. After stem extension and before the emergence of the flag leaf the composition of wax diversifies and already all compounds observed in the previous GC-MS experiment could be detected. The composition of the wax extracts did not change qualitatively from this point in the wheat growth cycle, but quantitative changes were observed with all compounds increasing in amount over time. This suggests a possible activation of wax producing genes like *W1* after the transition from the vegetative to the reproductive growth cycle. This in turn suggests that *Iw1* is also activated after the transition to the reproductive phase, although a constitutive activity of *Iw1* or a switch-like mechanism, which inactivates wax producing genes for example via chromatin remodeling, would be also feasible. Based on the complete lack of β-diketone aliphatics in Shamrock it seems unlikely that *Iw1* and wax producing genes are competing for substrates or otherwise engaged in a competitive interaction, as *Iw1* is completely dominant.

***Iw1* does not affect the thickness or composition of the cuticle**

We have compared the thickness of the cuticle and the cuticular conductance of *Iw1* and *iw1* germplasm and could not detect significant differences in either trait, suggesting that the structure and thickness of the cuticle is not affected by *Iw1*. The lack of differences in cuticular conductance between *Iw1* and *iw1* germplasm ties in with the presence of platelet shaped wax (C_{28}POHs) on the flag leaf surface of both glaucous and non-glaucous lines and their hypothesised function of insulating vascular tissue. To test this hypothesis we would need to compare the cuticular conductance of a wheat variety with a severe reduction or complete lack of the platelet shaped wax to glaucous and/or non-glaucous lines. Alternatively, gum arabic could be used to remove all epicuticular wax, thus producing an artificial 'waxless mutant' which can be used to assess these effects (Vogg et al. 2004).

Non-glaucousness reduces reflectance of light

We have analysed the light reflectance properties of flag leaves from glaucous and non-glaucous BC_2 NILs. The ratio of incident to reflected light was significantly lower (on average 13.73%) in all non-glaucous lines. Flag leaves of two pairs of BC_2 NILs were also tested for their light transmission properties, with the non-glaucous Alchemy NILs showing a higher transmission of light than their glaucous counterparts. However, this effect was not observed in the other NIL pair. These effects are very interesting as light is generally sparser in a UK-like environment compared with Mediterranean conditions. By reflecting less light non-glaucous plants could utilize more of the incoming light, especially if more of it can pass through the canopy where it can be intercepted by other photosynthetically active tissue. In addition, the reduced reflection of incoming light will likely increase the temperature inside the canopy leaves, which could be beneficial given the relatively low temperatures in the UK even in summer (~20°C in July and August (Met_Office 2012a)), as RUBISCO activity is optimal at ~25°C in wheat (Kobza and Edwards 1987). The reduced light reflection and suspected increase in leaf temperature does not lower water-use efficiency (WUE) in non-glaucous lines, as was demonstrated by carbon isotope discrimination.

The effects of *Iw1* on yield and green-canopy duration

One of the major aims of this study was to reproduce the increased yield and green-canopy duration effects seen in the DH population of Shango and Shamrock (Simmonds et al. 2008) in a set of NILs. If the yield and green-canopy duration effects could be observed in the NILs the causal genes would be have to originate from the distal 2BS segment introduced from *T. dicoccoides*, rather than being caused by a unique juxtaposition of alleles from Shamrock and Shango. In the end we could not observe a significant increase in yield in any of the NILs. The erratic weather over the last years did not help to achieve consistent field results either. However, out of the six NIL pairs tested one pair did show an effect on yield in every year, although this effect was not always significant. The non-glaucous Hereward NILs repeatedly out-competed their glaucous siblings for yield with a mean effect of ~4% across all years and locations. The size of the effect is reminiscent of the one observed in the DH population. We have created recombinant lines from the Hereward NILs and will fine map the yield effect using these lines. This analysis should reveal whether the yield effect coincides with *Iw1*. As for green-canopy duration, we did not observe a significant effect in the BC_2 NILs, but we did detect a significant effect in four out of six BC_4 NIL pairs in 2011/2012. This highlights the gain in accuracy obtained this highly isogenic material. Whether this increase in green-canopy duration has an effect on yield or not is not clear. Reports on the effects of the functional *GPC-B1* allele suggest that these two traits are not necessarily linked (Uauy et al. 2006, Waters et al. 2009), as plant senescence and grain-filling are two uncoupled processes.

On the usefulness of NILs

A set of six near isogenic lines was developed prior to the start of this project and continually developed until the BC_4 stage. These NILs have been used in all physiological and field-related experiments and have proven to be an extremely useful tool. The phenotype of an organism is determined by genotype, the environment and the interaction between them. We cannot control the latter but we can control the first. Over the last few years we have experienced a number of extreme weather events here in Norfolk and the UK; from the driest winter and spring on record to the wettest and dullest summer on record and unto the coldest spring on record. Despite this we managed to obtain reliable field data for most seasons with a low coefficient of variation, which enabled us to assign the observed effects to specific genetic intervals or dismiss them as environmental noise. This was all due to the isogenic nature of the NILs, which allowed us to make reliable comparisons in an unreliable environment. The production of NILs takes up a lot of time in wheat, but based on the results from the previous years it is time and effort well spent. Collecting data from six NIL pairs, plus their biological replicates, constituted a lot of work, but it is prudent to have more than one NIL pair at hand as long as the genetic interactions between the respective cultivar and the introgression are not clear. This is especially true when analysing complex traits like yield or plant-water relations.

The *Iw1* interval

At the start of my PhD the *Iw1* interval was approximately 2.3-cM in size. After screening 10,454 gametes from two mapping populations for recombinants the interval was reduced in size to 0.42-cM. This is still a comparatively large genetic interval considering the amount of gametes screened. Other map-based cloning projects achieved genetic intervals of 0.04-cM (6,190 gametes, *VRN1* (Yan et al. 2003)), 0.07-cM (9,831 gametes, *Vrs1* (Komatsuda et al. 2007)) or 0.34-cM (4,044 gametes, *mlo* (Büschges et al. 1997)) with comparable or even lower amounts of gametes. This comparatively large size is likely partly due to a lack of markers on the distal side of the *Iw1* interval, where most of the recombination events are located. But after starting the construction of physical maps in the non-glaucous variety TTD140 and the glaucous recombinant substitution lines RSL65 (cultivar Langdon) we started to suspect the presence of two distinct haplotypes across the *Iw1* interval. This was concluded from the lack of similarity between the two physical maps in terms of gene content and would explain the lack of recombination inside the *Iw1* interval. However, neither of the two physical maps is complete and especially the RSL65 map is still a lot smaller than the TTD140 map. Thus we cannot exclude the possibility that the observed differences in gene content are merely the result of incomplete maps rather than the presence of two distinct haplotypes. Likewise, the lack of

recombination could be due to a recombination cold-spot inside the *Iw1* interval (Lichten and Goldman 1995, Sidhu and Gill 2005). We will shortly receive the sequences of a great number of BACs from the physical maps, especially from RSL65. These sequences were unfortunately delayed by technical difficulties and could not be incorporated in this thesis. These BAC sequences should give us a better understanding of the relationship between TTD140 and RSL65 in the *Iw1* interval and will help us decide on the next steps.

Deletion mutants to break the recombination dead-lock

Whichever of the two above-mentioned hypotheses is true, the lack of recombination inside the *Iw1* interval cannot be ignored. We have started to predict genes from sequenced BACs of the physical map and will validate them with a mix of expression analysis and allelic diversity, followed by genetic complementation. However, with no recombination inside the *Iw1* interval we will likely be faced with a plethora of candidate genes that are expressed in the correct pattern and will also show a matching pattern of allelic diversity, which would make genetic complementation too expensive. A solution for this would be the creation of deletion mutants. This approach has been used successfully for the *Ph1* locus, which was also suffering from a lack of recombination (Griffiths et al. 2006). By using for example γ-irradiation a set of deletion mutants, missing defined segments of the *Iw1* interval, could be developed and phenotyped. Candidate genes mapping to deletion bins essential for *Iw1* function could then be further tested by genetic complementation.

Knocking out *Iw1* via EMS mutagenesis

We have also explored other routes for the validation of candidate genes. One of these is the attempt to knock out *Iw1* via EMS mutagenesis in Shamrock and TTD140. Unfortunately however the mutagenised TTD140 population suffered from extreme sterility, which resulted in a very low number of M_2 plants, none of which was glaucous as would be expected. For the mutagenised Shamrock population we discovered a single plant out of 1,985 M_2 families to be glaucous. We have confirmed this putative mutant to be of Shamrock genotype using Shamrock-specific PCR and RT-PCR markers. However, the low number of glaucous mutants inside the EMS population is worrying, especially since many other defects and mutant phenotypes have been observed in the population, suggesting that the lack of glaucous mutants is not due an insufficient dose of EMS. Possible explanations could be that *Iw1* is a small coding region that is hard to mutagenise due to its intrinsic small size or has a relatively low GC content which would make it less amenable to the G>A/C>T transitions characteristic of EMS in wheat. Alternatively, *Iw1* could be encoded by a pair of tandemly duplicated genes in which the probability of simultaneously changing amino acids critical to their function is essentially non-existent. Whatever the reasons, if the progeny of the putative mutant

display a glaucous phenotype it means that we have essentially tagged *Iw1*. This would allow us to create a new mapping population between Shamrock and the EMS-induced *iw1* Shamrock mutants. This population should be able to recombine inside the *Iw1* interval, assuming that the lack of recombination observed in the other populations is due to the presence of two distinct haplotypes. Mutations induced by the EMS treatment could serve as the basis for markers between the otherwise isogenic lines (Abe et al. 2012). These EMS-induced SNPs should be present at ~40 kb intervals based on the observed mutation frequencies using endonuclease digestion (Uauy et al. 2009). Alternatively, we could sequence the mutant line using Illumina HiSeq: A paired-end 100 bp library run on three sequencing lanes should produce more than 300 million reads, which would result in a 3.75-fold coverage of the genome. This would allow us to compare our gene models with the genomic sequence and identify SNPs that cause deleterious effects.

Candidate gene evaluation

The last two examples for the identity of *Iw1* are extreme scenarios, while it is just as likely that *Iw1* is a 'regular' gene. We have identified a number of possible candidate genes inside the TTD140 physical map that we are currently pursuing; these genes could be implicated in the inhibition of β-diketone production based on the domains they are predicted to encode. One gene in particular (*774P06_WPK*) had caught our interest because of its unique juxtaposition of domains, which were not found together in any nucleotide or protein databases. This tied in with the uniqueness of the *Iw1* locus, for which no homologues in other grass species are known. After acquiring a great body of evidence for this gene we decided to test it by genetic complementation. Two constructs, one with a native promoter and one CDS construct regulated by an *Actin* promoter, were transformed into the glaucous spring wheat cultivar Fielder. However, all primary transformants remained glaucous even though the constructs were expressed. We will extract cuticular wax from the transformants to assess whether they contain all wax classes associated with glaucousness. We are also currently crossing Shamrock to Fielder to ensure that *Iw1* is active in this genetic background, but it appears that *774P06_WPK* is not *Iw1*.

Using RNA-seq to identify *Iw1*

Another strategy for the validation, or even identification, of candidate genes is RNA-seq. By sequencing total RNA from a pair of BC_4 NILs, which are ~96.9% isogenic, we hope to identify transcripts that are unique to the non-glaucous NILs or highly overrepresented in them. These transcripts can be mapped by PCR to verify that they originate from the *Iw1* interval. We have received this data only recently and will start to analyse it over the coming weeks. However, the prerequisite for success with this approach lies in the expression pattern and mode of action of *Iw1*.

We would expect *Iw1* to be expressed after the switch to the reproductive phase when the wax producing genes are supposed to become active as well. However, we cannot exclude that *Iw1* is only expressed in a discrete developmental window at the top of a regulatory network where it sets of a self-regulating gene cascade. Likewise, we cannot exclude the possibility that *Iw1* is in fact not a gene but a regulatory non-coding RNA. However, the multitude of approaches that we are performing should cover these eventualities.

Identifying *Iw1* via its 2D homoeologue

Because of the difficulties in cloning *Iw1*, we have recently started to collaborate with another research group that has created a mapping population between a glaucous and a non-glaucous *Aegilops tauschii* accession. The F_1 generation was completely non-glaucous, suggesting that the phenotype is caused by *Iw2*. We will provide markers for the F_2 progeny of this population obtained from our physical maps. This approach allows us to work in a diploid organism and offers the possibility to identify *Iw1* via its 2D homoeologue.

Identifying possible targets of *Iw1*

Another possibility to clone *Iw1* is to identify some of its putative targets, specifically the wax producing gene *W1* which is located closely to *Iw1*. For this we have obtained a non-glaucous tetraploid line with a non-functional copy of *W1*. This line, called AUS2499 (Johnson et al. 1983), was crossed with glaucous tetraploid cultivars Kofa+*Lr19*, Langdon and Cappelli. The F_1 progeny of all crosses was completely glaucous, indicating that AUS2499 has indeed genotype *w1*. We have started screening a few hundred F_2 lines of the Kofa+*Lr19* x AUS2499 cross using KASPar markers and could confirm that the glaucous trait maps close to *Iw1* on to the short arm of chromosome 2B. We have to increase the marker density next to elucidate whether *W1* maps inside the *Iw1* interval. Our hypothesis is that *W1* maps inside the *Iw1* interval, which would possibly allow us to identify *W1* using the steadily growing RSL65 physical map. We further hypothesise that *W1* might give us a clue as to the identity of *Iw1*, based on the assumption that the first is regulated by the latter.

In addition, we have detected a previously undescribed locus for glaucousness on the long arm of chromosome 7A in a small F_2 population from CIMMYT. This locus is homologous to the *wa1* locus from rye on chromosome 7RL (Korzun et al. 1997) and is a dominant elicitor of glaucousness. This locus is present in a number of tetraploid spring wheat cultivars from CIMMYT, which we have crossed to the non-glaucous accession AUS2499. We will fine-map the gene using KASPar markers once we can harvest the F_2 seeds of our crosses. We hypothesise that this locus is also a potential target for *Iw1*, although we have not made that cross yet.

Identifying quantitative regulators of glaucousness

Most loci described in wheat that affect glaucousness are dominant, either producing or inhibiting it. A single locus on chromosome 3A was reported to be quantitatively affecting the glaucosity on flag leaves (Bennett et al. 2012). We have crossed the glaucous cultivar Shango with the UK winter wheat cultivar Stigg. The latter is completely glaucous, apart from the spike which is completely non-glaucous. Surprisingly, the F_1 generation showed an easily distinguishable intermediate level of glaucousness on the spike. This dosage-dependent phenotype is very interesting as it was not described before and thus might constitute a previously unknown player in the β-diketone pathway. We will extract cuticular wax from the spikes of F_1 plants to analyse the composition of waxes and we will map this trait in the F_2 generation using KASPar markers.

Personal statement

Over the course of my PhD I have witnessed an unprecedented rise in the number of tools and resources available to cereal geneticists. Having worked for several years on the model species *Arabidopsis thaliana* with its readily accessible genome sequence, large collection of knock-out mutants and easy-to-use agarose gel based markers, my start into wheat research was equal to a step back in time. Sequence databases held only EST data, knock-out mutants were few in number and the SSCP marker system required to be run over night.

But already shortly afterwards the KASPar marker system was introduced, followed by the release of a 5x coverage shotgun sequence of the wheat genome, which enabled us to convert an EST sequence into a full-length sequence of a (small) gene for all three genomes. Having access to genome-specific intron sequence, which harbours many SNPs, simplified and also accelerated the design of KASPar markers. Soon afterwards we could start to access the wheat survey sequence database and the barley genome sequence. TILLING populations for tetraploid and hexaploid wheat were created and released. The number of SNPs identified between different genotypes increased dramatically and so did the number of available genome-specific KASPar markers. Genetic maps of new varieties were created in a matter of weeks rather than years. Draft sequences of the A- and D-genome progenitors were released. A transcriptome database was created for tetraploid and hexaploid wheat to replace the incomplete and collated set of wheat UniGenes.

Virtually overnight wheat genetics had entered the 21st century. These changes will surely attract more young researchers that were previously put off by the unwieldy wheat genome and the lack of tools and resources for it. It was a privilege and an exciting experience to witness all these changes in such a short amount of time first hand.

References

Adamski, N. M., et al. (2013). "The Inhibitor of wax 1 locus (Iw1) prevents formation of β- and OH-β-diketones in wheat cuticular waxes and maps to a sub-cM interval on chromosome arm 2BS." The Plant Journal **74**(6): 989-1002.
Akhunov, E. D., et al. (2007). "Mechanisms and Rates of Birth and Death of Dispersed Duplicated Genes during the Evolution of a Multigene Family in Diploid and Tetraploid Wheats." Molecular Biology and Evolution **24**(2): 539-550.
Allan, R. E. and O. A. Vogel (1960). "F1 monosomic analysis involving a smooth-awn durum wheat." Wheat Information Service **11**: 3-4.
Allen, A. M., et al. (2011). "Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (Triticum aestivum L.)." Plant Biotechnol J **9**(9): 1086-1099.
Allen, A. M., et al. (2013). "Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (Triticum aestivum L.)." Plant Biotechnol J **11**(3): 279-295.
Arumuganathan, K. and E. D. Earle (1991). "Nuclear DNA content of some important plant species." Plant Molecular Biology Reporter **9**(3): 208-218.
Asnaghi, C., et al. (2000). "Application of synteny across Poaceae to determine the map location of a sugarcane rust resistance gene." Theoretical and Applied Genetics **101**(5-6): 962-969.
Badr, A., et al. (2000). "On the Origin and Domestication History of Barley (Hordeum vulgare)." Molecular Biology and Evolution **17**(4): 499-510.
Barakat, A., et al. (1997). "The distribution of genes in the genomes of Gramineae." Proceedings of the National Academy of Sciences **94**(13): 6857-6861.
Barber, H. N. and A. G. Netting (1968). "Chemical genetics of β-diketone formation in wheat." Phytochemistry **7**(12): 2089-2093.
Bartoš, J., et al. (2012). "Intraspecific sequence comparisons reveal similar rates of non-collinear gene insertion in the B and D genomes of bread wheat." BMC Plant Biology **12**(1): 1-12.
Beecher, B., et al. (2002). "Expression of wild-type pinB sequence in transgenic wheat complements a hard phenotype." Theoretical and Applied Genetics **105**(6-7): 870-877.
Bennett, D., et al. (2012). "Identification of novel quantitative trait loci for days to ear emergence and flag leaf glaucousness in a bread wheat (Triticum aestivum L.) population adapted to southern Australian conditions." Theoretical and Applied Genetics **124**(4): 697-711.
Bennett, M. D. and J. B. Smith (1976). "Nuclear DNA amounts in angiosperms." Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **274**: 227-274.
Bennett, M. D. and J. B. Smith (1991). "Nuclear DNA Amounts in Angiosperms." Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **334**(1271): 309-345.
Bennetzen, J. L. and M. Freeling (1997). "The Unified Grass Genome: Synergy in Synteny." Genome Research **7**(4): 301-306.
Bennetzen, J. L., et al. (1998). "Grass genomes." Proceedings of the National Academy of Sciences **95**(5): 1975-1978.
Blum, A. (1975a). "Effect of the bm gene on epicuticular wax and the water relations of Sorghum bicolor." Israel Journal of Botany **24**: 50-ss.
Blum, A. (1975b). "Effect of the bm gene on epicuticular wax deposition and the spectral characteristics of Sorghum leaves." SABRAO Journal **7**: 45-52.
Blum, A. (1988). Plant Breeding for stress environments, CRC Press.
Blum, A. (2005). "Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive?" Australian Journal of Agricultural Research **56**(11): 1159.
Börner, A., et al. (2002). "Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (Triticum aestivum L.)." Theoretical and Applied Genetics **105**(6-7): 921-936.
Borojevic, K. and K. Borojevic (2005). "The transfer and history of "reduced height genes" (Rht) in wheat from Japan to Europe." J Hered **96**(4): 455-459.

Buckler, E. S. t., et al. (2001). "Molecular diversity, structure and domestication of grasses." Genet Res **77**(3): 213-218.
Büschges, R., et al. (1997). "The barley Mlo gene: a novel control element of plant pathogen resistance." Cell **88**(5): 695-705.
Carver, B., et al. (1989). "Genetic analysis of photosynthetic variation in hexaploid and tetraploid wheat and their interspecific hybrids." Photosynthesis Research **20**(2): 105-118.
Casacuberta, J. M. and N. Santiago (2003). "Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes." Gene **311**(0): 1-11.
Cenci, A., et al. (2003). "Construction and characterization of a half million clone BAC library of durum wheat (Triticum turgidum ssp. durum)." Theoretical and Applied Genetics **107**(5): 931-939.
Cervantes, D. E., et al. (2002). "Oviposition Responses by Hessian Fly, Mayetiola destructor, to Wheats Varying in Surfaces Waxes." Journal of Chemical Ecology **28**(1): 193-210.
Chalupska, D., et al. (2008). "Acc homoeoloci and the evolution of wheat genomes." Proceedings of the National Academy of Sciences **105**(28): 9691-9696.
Chao, S., et al. (1989). "RFLP-based genetic maps of wheat homoeologous group 7 chromosomes." Theoretical and Applied Genetics **78**(4): 495-504.
Chao, S. and D. Somers. "Wheat and Barley DNA Extraction in 96-well Plates." from http://maswheat.ucdavis.edu/protocols/general_protocols/DNA_extraction_003.htm.
Chatterton, N. J., et al. (1975). "Photosynthesis and Transpiration of Bloom and Bloomless Sorghum." Canadian Journal of Plant Science **55**(2): 641-643.
Chopra, S., et al. (1999). "Molecular characterization of a mutable pigmentation phenotype and isolation of the first active transposable element from Sorghum bicolor." Proceedings of the National Academy of Sciences **96**(26): 15330-15335.
CIMMYT (2004). Wild Wheat Relatives Help Boost Genetic Diversity.
Devos, K. M., et al. (1992). "RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye." Theoretical and Applied Genetics **83**(8): 931-939.
Devos, K. M. and M. D. Gale (1992). "The use of random amplified polymorphic DNA markers in wheat." Theoretical and Applied Genetics **84**(5-6): 567-572.
DeYoung, B. J. and R. W. Innes (2006). "Plant NBS-LRR proteins in pathogen sensing and host defense." Nat Immunol **7**(12): 1243-1249.
Doležel, J., et al. (1998). "Plant Genome Size Estimation by Flow Cytometry: Inter-laboratory Comparison." Annals of Botany **82**(suppl 1): 17-26.
Dreisigacker, S., et al. (2005). "Genetic Diversity among and within CIMMYT Wheat Landrace Accessions Investigated with SSRs and Implications for Plant Genetic Resources Management." Crop Sci. **45**(2): 653-661.
Driscoll, C. J. and N. F. Jensen (1964). "Chromosomes Associated with Waxlessness, Awnedness and Time of Maturity of Common Wheat." Canadian Journal of Genetics and Cytology **6**(3): 324-333.
Dubcovsky, J. and J. Dvorak (2007). "Genome Plasticity a Key Factor in the Success of Polyploid Wheat Under Domestication." Science **316**(5833): 1862-1866.
Dubcovsky, J., et al. (1997). "Seed-storage-protein loci in RFLP maps of diploid, tetraploid, and hexaploid wheat." Theoretical and Applied Genetics **95**(7): 1169-1180.
Durrett, R. T., et al. (2002). "A Simple Formula Useful for Positional Cloning." Genetics **160**(1): 353-355.
Dvořák, J., et al. (1984). "The C-band pattern of a Ph− mutant of durum wheat." Canadian Journal of Genetics and Cytology **26**(3): 360-363.
Dvorak, J., et al. (1998). "The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat." Theoretical and Applied Genetics **97**(4): 657-670.
Edwards, D., et al. (1982). "The cuticle of early vascular plants and its evolutionary significance." The plant cuticle. Linnean Society Symposium Series: 341-361.
Edwards, D., et al. (1998). "Stomata in early land plants: an anatomical and ecophysiological approach." Journal of Experimental Botany **49**(Special Issue): 255-278.
Esau, K. (1953). Plant anatomy. New York, Wiley.

Fahima, T., et al. (1998). "Microsatellite DNA polymorphism divergence in Triticum dicoccoides accessions highly resistant to yellow rust." Theoretical and Applied Genetics **96**(2): 187-195.
FAO (2009). How to Feed the World in 2050.
Faris, J. D., et al. (2003). "A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene." Genetics **164**(1): 311-321.
Faris, J. D., et al. (2000). "Saturation Mapping of a Gene-Rich Recombination Hot Spot Region in Wheat." Genetics **154**(2): 823-835.
Farquhar, G. D., et al. (1989). "Carbon Isotope Discrimination and Photosynthesis." Annual Review of Plant Physiology and Plant Molecular Biology **40**(1): 503-537.
Febrero, A., et al. (1998). "Yield, carbon isotope discrimination, canopy reflectance and cuticular conductance of barley isolines of differing glaucousness." Journal of Experimental Botany **49**(326): 1575-1581.
Fedak, G. (1999). "Molecular aids for integration of alien chromatin through wide crosses." Genome **42**(4): 584-591.
Feldman, M. (2001). The Origin of Cultivated Wheat. The Wheat Book: A History of Wheat Breeding. A. P. Benjean and W. J. Angus. Paris, Lavoisier Publishing**:** 3-56.
Feldman, M., et al. (1995). Wheats. Evolution of Crop Plants. J. Smartt and N. W. Simmonds. London, Longman Scientific**:** 184-192.
Feuillet, C., et al. (2008). "Cereal breeding takes a walk on the wild side." Trends in Genetics **24**(1): 24-32.
Finn, R. D., et al. (2010). "The Pfam protein families database." Nucleic Acids Res **38**(suppl 1): D211-D222.
Flavell, R. B., et al. (1974). "Genome size and the proportion of repeated nucleotide sequence DNA in plants." Biochemical Genetics **12**(4): 257-269.
Frey, M., et al. (1997). "Analysis of a Chemical Plant Defense Mechanism in Grasses." Science **277**(5326): 696-699.
Fu, D., et al. (2005). "Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat." Molecular Genetics and Genomics **274**(4): 442-443.
Gadaleta, A., et al. (2009). "Genetic and physical mapping of new EST-derived SSRs on the A and B genome chromosomes of wheat." Theoretical and Applied Genetics **118**(5): 1015-1025.
Gale, M. D. and K. M. Devos (1998). "Comparative genetics in the grasses." Proceedings of the National Academy of Sciences **95**(5): 1971-1974.
Gepts, P. (2003). Ten thousand years of crop evolution. Plants, Genes, and Crop Biotechnology. M. Chrispeels and D. Sadava. Sudbury, MA, Bartlett and Jones**:** 328-359.
Gerechter-Amitai, Z. K. and R. W. Stubbs (1970). "A valuable source of yellow rust resistance in Israeli populations of wild emmer,Triticum dicoccoides Koern." Euphytica **19**(1): 12-21.
Ghimire, S. K., et al. (2005). "Genetic Diversity and Geographical Differentiation in Asian Common Wheat (Triticum aestivum L.), Revealed by the Analysis of Peroxidase and Esterase Isozymes." Breeding Science **55**(2): 175-185.
Gray, J. and A. J. Boucot (1977). "Early vascular land plants: proof and conjecture." Lethaia **10**(2): 145-174.
Griffiths, S., et al. (2006). "Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat." Nature **439**(7077): 749-752.
Gupta, P., et al. (2002). "Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat." Theoretical and Applied Genetics **105**(2-3): 413-422.
Gupta, P. K., et al. (2001). "Single nucleotide polymorphisms (SNPs): a new paradigm in molecular marker technology and DNA polymorphism detection with emphasis on their use in plants." Current Science **80**(4): 524-535.
Hall, D. M., et al. (1965). "Infra-Specific Variation in Wax on Leaf Surfaces." Australian Journal of Biological Sciences **18**(2): 323-332.
Hammer, K. (1984). "Das Domestikationssyndrom." Die Kulturpflanze **32**(1): 11-34.
Harberd, N. P., et al. (1987). "Identification of a transposon-like insertion in a Glu-1 allele of wheat." Molecular and General Genetics MGG **209**(2): 326-332.

Harwood, W. A. (2012). "Advances and remaining challenges in the transformation of barley and wheat." Journal of Experimental Botany **63**(5): 1791-1798.
Haudry, A., et al. (2007). "Grinding up Wheat: A Massive Loss of Nucleotide Diversity Since Domestication." Molecular Biology and Evolution **24**(7): 1506-1517.
Hedden, P. (2003). "The genes of the Green Revolution." Trends Genet **19**(1): 5-9.
Hegde, Y. and P. E. Kolattukudy (1997). "Cuticular waxes relieve self-inhibition of germination and appressorium formation by the conidia ofMagnaporthe grisea." Physiological and Molecular Plant Pathology **51**(2): 75-84.
Heredia, A. (2003). "Biophysical and biochemical characteristics of cutin, a plant barrier biopolymer." Biochimica et Biophysica Acta (BBA) - General Subjects **1620**(1–3): 1-7.
Heun, M., et al. (1997). "Site of Einkorn Wheat Domestication Identified by DNA Fingerprinting." Science **278**(5341): 1312-1314.
HGCA (2006). Recommended List. H. R. l. 2006/2007. www.hgca.com, HGCA.
HGCA (2008). "The Wheat Growth Guide." from http://www.hgca.com/cms_publications.output/2/2/Publications/On-farm%20information/The%20Wheat%20Growth%20Guide.mspx?fn=show&pubcon=4444.
Holloway, P. J. (1982). Structure and Histochemistry of Plant Cuticular Membranes: An Overview. The plant cuticle. D. F. Cutler, K. L. Alvin and C. E. Price. London, Academic Press**:** 1-32.
Huang, L., et al. (2003). "Map-Based Cloning of Leaf Rust Resistance Gene Lr21 From the Large and Polyploid Genome of Bread Wheat." Genetics **164**(2): 655-664.
Huang, X., et al. (2002). "Assessing genetic diversity of wheat (Triticum aestivum L.) germplasm using microsatellite markers." Theoretical and Applied Genetics **105**(5): 699-707.
Huehn, M. (1993). "Harvest index versus grain/straw-ratio. Theoretical comments and experimental results on the comparison of variation." Euphytica **68**(1-2): 27-32.
Ignoul, S. and J. Eggermont (2005). "CBS domains: structure, function, and pathology in human proteins." American Journal of Physiology - Cell Physiology **289**(6): C1369-C1378.
Initiative, W. (2013). "Objectives of the wheat Initiative." from http://www.wheatinitiative.org/about/objectives.
International Brachypodium Initiative (2010). "Genome sequencing and analysis of the model grass Brachypodium distachyon." Nature **463**(7282): 763-768.
International Wheat Genome Sequencing Consortium (2011). "Wheat Survey Sequence."
Islam, A. K. M. R. and K. W. Shepherd (1991). Alien genetic variation in wheat improvement. Chromosome Engineering in Plants: Genetics, Breeding, Evolution (Developments in Plant Genetics and Breeding). P. K. Gupta and T. Tsuchiya. Amsterdam, Netherlands, Elsevier Science Publishers.
Jakob, S. S., et al. (2004). "The Considerable Genome Size Variation of Hordeum Species (Poaceae) Is Linked to Phylogeny, Life Form, Ecology, and Speciation Rates." Molecular Biology and Evolution **21**(5): 860-869.
Jantasuriyarat, C., et al. (2004). "Identification and mapping of genetic loci affecting the free-threshing habit and spike compactness in wheat (Triticum aestivum L.)." Theoretical and Applied Genetics **108**(2): 261-273.
Jeffree, C. E. (1996). Structure and Ontogeny of Plant Cuticles. Plant cuticles: an integrated functional approach. G. Kerstiens. Oxford, Bios Scientific Publishers**:** 33-82.
Jensen, N. F. and C. J. Driscoll (1962). "Inheritance of the Waxless Character in Wheat." Crop Sci. **2**(6): 504-505.
Jetter, R., et al. (2007). Composition of Plant Cuticular Waxes. Annual Plant Reviews Volume 23: Biology of the Plant Cuticle, Blackwell Publishing Ltd**:** 145-181.
Johnson, D. A., et al. (1983). "Yield, Water Relations, Gas Exchange, and Surface Reflectances of Near-Isogenic Wheat Lines Differing in Glaucousness1." Crop Sci. **23**(2): 318-325.
Kalendar, R., et al. (2000). "Genome evolution of wild barley (Hordeum spontaneum) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence." Proceedings of the National Academy of Sciences **97**(12): 6603-6607.
Kawahara, Y., et al. (2013). "Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data." Rice **6**(1): 1-10.

Kawaura, K., et al. (2005). "Expression Profile of Two Storage-Protein Gene Families in Hexaploid Wheat Revealed by Large-Scale Analysis of Expressed Sequence Tags." Plant Physiology **139**(4): 1870-1880.
Kerstiens, G. (1996). "Cuticular water permeability and its physiological significance." Journal of Experimental Botany **47**(12): 1813-1832.
Kihara, H. (1944). "Discovery of the DD-analyser, one of the ancestors of Triticum vulgare." Agricultural Horticulture **19**: 889-890.
Kilian, A., et al. (1997). "Towards map-based cloning of the barley stem rust resistance genes Rpg1 and rpg4 using rice as an intergenomic cloning vehicle." Plant Molecular Biology **35**(1-2): 187-195.
King, R. W. and P. von Wettstein-Knowles (2000). "Epicuticular waxes and regulation of ear wetting and pre-harvest sprouting in barley and wheat." Euphytica **112**(2): 157-166.
Kloeckener-Gruissem, B. and M. Freeling (1995). "Transposon-induced promoter scrambling: a mechanism for the evolution of new alleles." Proceedings of the National Academy of Sciences **92**(6): 1836-1840.
Kobza, J. and G. E. Edwards (1987). "Influences of Leaf Temperature on Photosynthetic Carbon Metabolism in Wheat." Plant Physiology **83**(1): 69-74.
Koch, K., et al. (2006). "Structural analysis of wheat wax (Triticum aestivum, c.v. 'Naturastar' L.): from the molecular level to three dimensional crystals." Planta **223**(2): 258-270.
Koch, K. and H.-J. Ensikat (2008). "The hydrophobic coatings of plant surfaces: Epicuticular wax crystals and their morphologies, crystallinity and molecular self-assembly." Micron **39**(7): 759-772.
Komatsuda, T., et al. (2007). "Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene." Proceedings of the National Academy of Sciences **104**(4): 1424-1429.
Konishi, S., et al. (2006). "An SNP Caused Loss of Seed Shattering During Rice Domestication." Science **312**(5778): 1392-1396.
Korzun, V., et al. (1997). "RFLP-based mapping of three mutant loci in rye (Secale cereale L.) and their relation to homoeologous loci within the Gramineae." Theoretical and Applied Genetics **95**(3): 468-473.
Kozubek, A. and J. H. P. Tyman (1995). "Cereal grain resorcinolic lipids: mono and dienoic homologues are present in rye grains." Chemistry and Physics of Lipids **78**(1): 29-35.
Krasileva, K., et al. (2013). "Separating homeologs by phasing in the tetraploid wheat transcriptome." Genome Biology **14**(6): R66.
Kulwal, P. L., et al. (2003). "QTL mapping for growth and leaf characters in bread wheat." Plant Science **164**(2): 267-277.
Kunst, L. and L. Samuels (2009). "Plant cuticles shine: advances in wax biosynthesis and export." Current Opinion in Plant Biology **12**(6): 721-727.
Kushnir, U. and G. M. Halloran (1984). "Transfer of high kernel weight and high protein from wild tetraploid wheat (Triticum turgidum dicoccoides) to bread wheat (T. aestivum) using homologous and homoeologous recombination." Euphytica **33**(1): 249-255.
Kushwaha, H., et al. (2009). "Genome wide expression analysis of CBS domain containing proteins in Arabidopsis thaliana (L.) Heynh and Oryza sativa L. reveals their developmental and stress regulation." BMC Genomics **10**(1): 200.
Lagudah, E. S., et al. (2001). "Wheat genomics." Plant Physiology and Biochemistry **39**(3–4): 335-344.
Li, C., et al. (2006). "Rice Domestication by Reducing Shattering." Science **311**(5769): 1936-1939.
Lichten, M. and A. S. H. Goldman (1995). "Meiotic Recombination Hotspots." Annual Review of Genetics **29**(1): 423-444.
Limagrain (2011) Nickerson variety breaks world wheat yield record.
Lodwig, E. M., et al. (2005). "Role of Polyhydroxybutyrate and Glycogen as Carbon Storage Compounds in Pea and Bean Bacteroids." Molecular Plant-Microbe Interactions **18**(1): 67-74.
Lukaszewski, A. J. and C. A. Curtis (1993). "Physical distribution of recombination in B-genome chromosomes of tetraploid wheat." Theoretical and Applied Genetics **86**(1): 121-127.
Lundqvist, U. and A. Lundqvist (1988). "Mutagen specificity in barley for 1580 eceriferum mutants localized to 79 loci." Hereditas **108**(1): 1-12.

Lundqvist, U. and P. Wettstein-Knowles (1982). "Dominant mutations at Cer-yy change barley spike wax into leaf blade wax." Carlsberg Research Communications **47**(1): 29-43.
Luo, M. C., et al. (2007). "The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication." Theoretical and Applied Genetics **114**(6): 947-959.
Marchler-Bauer, A., et al. (2011). "CDD: a Conserved Domain Database for the functional annotation of proteins." Nucleic Acids Res **39**(suppl 1): D225-D229.
Matthews, D. E., et al. (2008). "The Triticeae Repeat Sequence Database." from http://wheat.pw.usda.gov/ITMI/Repeats/.
Mayer, K. F. X., et al. (2011). "Unlocking the Barley Genome by Chromosomal and Comparative Genomics." The Plant Cell Online **23**(4): 1249-1263.
McCallum, C. M., et al. (2000). "Targeting Induced Local Lesions IN Genomes (TILLING) for Plant Functional Genomics." Plant Physiology **123**(2): 439-442.
Merah, O., et al. (2000). "Effect of Glaucousness on Carbon Isotope Discrimination and Grain Yield in Durum Wheat." Journal of Agronomy and Crop Science **185**(4): 259-265.
Met_Office (2011). "UK Climate in spring 2011." from http://www.metoffice.gov.uk/climate/uk/2011/spring.html.
Met_Office (2012a). 1981-2010 UK climate averages.
Met_Office (2012b). "UK climate in summer 2012." from http://www.metoffice.gov.uk/climate/uk/2012/summer.html.
Mette, M. F., et al. (2000). "Transcriptional silencing and promoter methylation triggered by double-stranded RNA." EMBO J **19**(19): 5194-5201.
Mikkelsen, J. D. and P. von Wettstein-Knowles (1978). "Biosynthesis of β-diketones and hydrocarbons in barley spike epicuticular wax." Archives of Biochemistry and Biophysics **188**(1): 172-181.
Mochida, K., et al. (2009). "TriFLDB: A Database of Clustered Full-Length Coding Sequences from Triticeae with Applications to Comparative Grass Genomics." Plant Physiology **150**(3): 1135-1146.
Monneveux, P., et al. (2004). "Relationships between Grain Yield, Flag Leaf Morphology, Carbon Isotope Discrimination and Ash Content in Irrigated Wheat." Journal of Agronomy and Crop Science **190**(6): 395-401.
Moore, G., et al. (1995). "Cereal Genome Evolution: Grasses, line up and form a circle." Current biology : CB **5**(7): 737-739.
Moose, S. P. and P. H. Sisco (1996). "Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity." Genes & Development **10**(23): 3018-3027.
Morris, B. D., et al. (2000). "Identification of 1-Octacosanal and 6-Methoxy-2-Benzoxazolinone from Wheat as Ovipositional Stimulants for Hessian Fly, Mayetiola destructor." Journal of Chemical Ecology **26**(4): 859-873.
Morris, C. (2002). "Puroindolines: the molecular genetic basis of wheat grain hardness." Plant Molecular Biology **48**(5-6): 633-647.
Nalam, V. J., et al. (2006). "Map-based analysis of genes affecting the brittle rachis character in tetraploid wheat (Triticum turgidum L.)." Theoretical and Applied Genetics **112**(2): 373-381.
Nelson, D. R., et al. (2004). "Comparative Genomics of Rice and Arabidopsis. Analysis of 727 Cytochrome P450 Genes and Pseudogenes from a Monocot and a Dicot." Plant Physiology **135**(2): 756-772.
Nesbitt, M. and D. Samuel (1996). From staple crop to extinction? The archeaology and history of hulled wheats. Hulled Wheats. S. Padulosi, K. Hammer and J. Heller.
Netting, A. G. and P. Wettstein-Knowles (1973). "The physico-chemical basis of leaf wettability in wheat." Planta **114**(4): 289-309.
Neuffer, M., et al. (1997). Mutants in Maize. Cold Spring Harbor Laboratory press, Cold Spring Harbor, NY.
Nevo, E., et al. (1991). "Resistance of wild emmer wheat to stem rust: Ecological, pathological and allozyme associations." Euphytica **53**(2): 121-130.
Nevo, E., et al. (1986). "Resistance of wild wheat to stripe rust: Predictive method by ecology and allozyme genotypes." Plant Systematics and Evolution **153**(1-2): 13-30.

Nevo, E., et al. (1985). "Patterns of resistance of Israeli wild emmer wheat to pathogens I. Predictive method by ecology and allozyme genotypes for powdery mildew and leaf rust." Genetica **67**(3): 209-222.
Ni, X., et al. (1998). "Influence of cereal leaf epicuticular wax on Diuraphis noxia probing behavior and nymphoposition." Entomologia Experimentalis et Applicata **89**(2): 111-118.
Nussbaumer, T., et al. (2013). "MIPS PlantsDB: a database framework for comparative plant genome research." Nucleic Acids Res **41**(D1): D1144-D1151.
Orata, F. (2012). Derivatization Reactions and Reagents for Gas Chromatography Analysis. Advanced Gas Chromatography - Progress in Agricultural, Biomedical and Industrial Applications. M. A. Mohd, InTech**:** 83-108.
Ozkan, H., et al. (2002). "AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey." Mol Biol Evol **19**(10): 1797-1801.
Paux, E., et al. (2006). "Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B." Plant J **48**(3): 463-474.
Peleg, Z. V. I., et al. (2008). "Allelic diversity associated with aridity gradient in wild emmer wheat populations." Plant, Cell & Environment **31**(1): 39-49.
Peng, J., et al. (2000). "Molecular Genetic Maps in Wild Emmer Wheat, Triticum dicoccoides: Genome-Wide Coverage, Massive Negative Interference, and Putative Quasi-Linkage." Genome Research **10**(10): 1509-1531.
Pflieger, S., et al. (2001). "The candidate gene approach in plant genetics: a review." Molecular Breeding **7**(4): 275-291.
Premachandra, G. S., et al. (1994). "Epicuticular wax load and water-use efficiency in bloomless and sparse-bloom mutants of Sorghum bicolor L." Environmental and Experimental Botany **34**(3): 293-301.
Qi, X., et al. (2004). "A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants." Proceedings of the National Academy of Sciences of the United States of America **101**(21): 8233-8238.
Reicosky, D. A. and J. W. Hanover (1978). "Physiological Effects of Surface Waxes." Plant Physiol **62**(1): 101-104.
Richards, R. A. (1984). Glaucousness in wheat, its effect on yield and related characteristics in dryland environments, and its control by minor genes. Proceedings of the sixth International Wheat Genetics Symposium, Kyoto University.
Richards, R. A., et al. (1986). "Glaucousness in Wheat: Its Development and Effect on Water-use Efficiency, Gas Exchange and Photosynthetic Tissue Temperatures*." Functional Plant Biology **13**(4): 465-473.
Riederer, M. and L. Schreiber (2001). "Protecting against water loss: analysis of the barrier properties of plant cuticles." Journal of Experimental Botany **52**(363): 2023-2032.
Röder, M. S., et al. (1998). "A Microsatellite Map of Wheat." Genetics **149**(4): 2007-2023.
Rong, J. K., et al. (2000). "A new powdery mildew resistance gene: Introgression from wild emmer into common wheat and RFLP-based mapping." Euphytica **115**(2): 121-126.
Rostoks, N., et al. (2005). "Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress." Mol Genet Genomics **274**(5): 515-527.
Salamov, A. A. and V. V. Solovyev (2000). "Ab initio Gene Finding in Drosophila Genomic DNA." Genome Research **10**(4): 516-522.
Salse, J., et al. (2009). "Reconstruction of monocotelydoneous proto-chromosomes reveals faster evolution in plants than in animals." Proceedings of the National Academy of Sciences **106**(35): 14908-14913.
Samuels, L., et al. (2008). "Sealing Plant Surfaces: Cuticular Wax Formation by Epidermal Cells." Annu Rev Plant Biol **59**(1): 683-707.
SanMiguel, P., et al. (1998). "The paleontology of intergene retrotransposons of maize." Nat Genet **20**(1): 43-45.
Schuler, M. A. and D. Werck-Reichhart (2003). "Functional genomics of P450s." Annu Rev Plant Biol **54**: 629-667.

Scott, J. W., et al. (2004). "CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations." The Journal of Clinical Investigation **113**(2): 274-284.
Shirasu, K., et al. (2000). "A Contiguous 66-kb Barley DNA Sequence Provides Evidence for Reversible Genome Expansion." Genome Research **10**(7): 908-915.
Sidhu, D. and K. Gill (2005). "Distribution of genes and recombination in wheat and other eukaryotes." Plant Cell, Tissue and Organ Culture **79**(3): 257-270.
Simmonds, J. R., et al. (2008). "Mapping of a gene (Vir) for a non-glaucous, viridescent phenotype in bread wheat derived from Triticum dicoccoides, and its association with yield variation." Euphytica **159**(3): 333-341.
Simons, K. J., et al. (2006). "Molecular Characterization of the Major Wheat Domestication Gene Q." Genetics **172**(1): 547-555.
Slade, A. J., et al. (2005). "A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING." Nature Biotechnology **23**(1): 75-81.
Soderlund, C., et al. (1997). "FPC: a system for building contigs from restriction fingerprinted clones." Computer applications in the biosciences : CABIOS **13**(5): 523-535.
Sorrells, M. E., et al. (2003). "Comparative DNA Sequence Analysis of Wheat and Rice Genomes." Genome Research **13**(8): 1818-1827.
States, D. J. and W. Gish (1994). "Combined use of sequence similarity and codon bias for coding region identification." J Comput Biol **1**(1): 39-50.
Stein, N., et al. (2000). "Subgenome chromosome walking in wheat: A 450-kb physical contig in Triticum monococcum L. spans the Lr10 resistance locus in hexaploid wheat (Triticum aestivum L.)." Proceedings of the National Academy of Sciences **97**(24): 13436-13441.
Sternberg, J. C., et al. (1962). "The Mechanism of Response of Flame Ionization Detectors." Gas Chromatography: 231-267.
Sturaro, M., et al. (2005). "Cloning and characterization of GLOSSY1, a maize gene involved in cuticle membrane and wax production." Plant Physiol **138**(1): 478-489.
Tacke, E., et al. (1995). "Transposon tagging of the maize Glossy2 locus with the transposable element En/Spm." The Plant Journal **8**(6): 907-917.
Takahashi, S., et al. (1999). "Capture of a genomic HMG domain sequence by the En/Spm-related transposable element Tpn1 in the Japanese morning glory." Molecular and General Genetics MGG **261**(3): 447-451.
Tanksley, S. D. and S. R. McCouch (1997). "Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild." Science **277**(5329): 1063-1066.
Tatusov, R. L., et al. (2000). "The COG database: a tool for genome-scale analysis of protein functions and evolution." Nucleic Acids Res **28**(1): 33-36.
Teshima, K. M. and H. Innan (2008). "Neofunctionalization of Duplicated Genes Under the Pressure of Gene Conversion." Genetics **178**(3): 1385-1398.
Tsuba, M., et al. (2002). "Chemical factors of the leaf surface involved in the morphogenesis of Blumeria graminis." Physiological and Molecular Plant Pathology **60**(2): 51-57.
Tsunewaki, K. (1964). "Genetic Studies of A 6x-Derivative from an 8x Triticale." Canadian Journal of Genetics and Cytology **6**(1): 1-11.
Tsunewaki, K. (1966). "Comparative gene analysis of common wheat and its ancestral species. II. Waxiness, growth habit and awnedness." Japanese Journal of Botany **19**: 175-254.
Tsunewaki, K. and K. Ebana (1999). "Production of near-isogenic lines of common wheat for glaucousness and genetic basis of this trait clarified by their use." Genes & Genetic Systems **74**(2): 33-41.
Tulloch, A. P. and L. L. Hoffman (1973). "Leaf wax of Triticum aestivum." Phytochemistry **12**(9): 2217-2223.
Tulloch, A. P. and L. R. Hogge (1978). "Gas chromatographic—mass spectrometric analysis of β-diketone-containing plant waxes : Use of trimethylsilyl ethers." Journal of Chromatography A **157**(0): 291-296.
Uauy, C., et al. (2006). "A NAC Gene Regulating Senescence Improves Grain Protein, Zinc, and Iron Content in Wheat." Science **314**(5803): 1298-1301.

Uauy, C., et al. (2009). "A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat." BMC Plant Biology **9**(1): 115.
Uauy, C. and D. University of California (2007). Positional Cloning of Gpc-B1, a Wheat Quantitative Trait Loci Affecting Senescence and with Pleiotropic Effects on Grain Protein, Zinc and Iron Concentration, University of California, Davis.
Uzarowska, A., et al. (2009). "Validation of candidate genes putatively associated with resistance to SCMV and MDMV in maize (Zea mays L.) by expression profiling." BMC Plant Biology **9**(1): 15.
Van Deynze, A. E., et al. (1998). "Anchor probes for comparative mapping of grass genera." Theoretical and Applied Genetics **97**(3): 356-369.
Vogg, G., et al. (2004). "Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional characterization of a mutant deficient in a very-long-chain fatty acid β-ketoacyl-CoA synthase." Journal of Experimental Botany **55**(401): 1401-1410.
von Wettstein-Knowles, P. (1972). "Genetic control of β-diketone and hydroxy-β-diketone synthesis in epicuticular waxes of barley." Planta **106**(2): 113-130.
Von Wettstein-Knowles, P. (1974). "Gene mutation in barley inhibiting the production and use of C26 chains in epicuticular wax formation." FEBS Letters **42**(2): 187-191.
von Wettstein-Knowles, P. (1976). "Biosynthetic relationships between β-diketones and esterified alkan-2-ols deduced from epicuticular wax of barley mutants." Molecular and General Genetics MGG **144**(1): 43-48.
von Wettstein-Knowles, P. (2012). Plant Waxes. eLS, John Wiley & Sons, Ltd.
von Wettstein-Knowles, P. and B. Søgaard (1980). "The cer-cqu region in barley: Gene cluster or multifunctional gene." Carlsberg Research Communications **45**(2): 125-141.
Vos, P., et al. (1995). "AFLP: a new technique for DNA fingerprinting." Nucleic Acids Res **23**(21): 4407-4414.
Vrána, J., et al. (2000). "Flow Sorting of Mitotic Chromosomes in Common Wheat (Triticum aestivum L.)." Genetics **156**(4): 2033-2041.
Wang, T. L., et al. (2012). "TILLING in extremis." Plant Biotechnol J **10**(7): 761-772.
Waterhouse, P. M. and C. A. Helliwell (2003). "Exploring plant genomes by RNA-induced gene silencing." Nat Rev Genet **4**(1): 29-38.
Waters, B. M., et al. (2009). "Wheat (Triticum aestivum) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain." Journal of Experimental Botany **60**(15): 4263-4274.
Wessler, S. R. (1998). "Transposable elements associated with normal plant genes." Physiologia Plantarum **103**(4): 581-586.
Wessler, S. R., et al. (1995). "LTR-retrotransposons and MITEs: important players in the evolution of plant genomes." Current Opinion in Genetics & Development **5**(6): 814-821.
Wettstein-Knowles, P. and A. G. Netting (1976). "Composition of epicuticular waxes on barley spikes." Carlsberg Research Communications **41**(5): 225-235.
Wicker, T., et al. (2003). "Rapid Genome Divergence at Orthologous Low Molecular Weight Glutenin Loci of the A and Am Genomes of Wheat." The Plant Cell Online **15**(5): 1186-1197.
Wilkinson, P., et al. (2012). "CerealsDB 2.0: an integrated resource for plant breeders and scientists." BMC Bioinformatics **13**(1): 219.
Williams, J. G., et al. (1990). "DNA polymorphisms amplified by arbitrary primers are useful as genetic markers." Nucleic Acids Res **18**(22): 6531-6535.
Winfield, M. O., et al. (2012). "Targeted re-sequencing of the allohexaploid wheat exome." Plant Biotechnol J **10**(6): 733-742.
Wingen, L. U., et al. (2012). Mining for Useful Variation in the AE Watkins Wheat Landrace Collection. 19th EUCARPIA GENERAL CONGRESS. Budapest, Hungary.
Xu, X., et al. (1997). "Sequence Analysis of the Cloned glossy8 Gene of Maize suggests that it may code for a b-ketoacyl reductase required for the biosynthesis of cuticular waxes." Plant Physiol **115**: 501-510.
Yan, L., et al. (2004). "The Wheat VRN2 Gene Is a Flowering Repressor Down-Regulated by Vernalization." Science **303**(5664): 1640-1644.

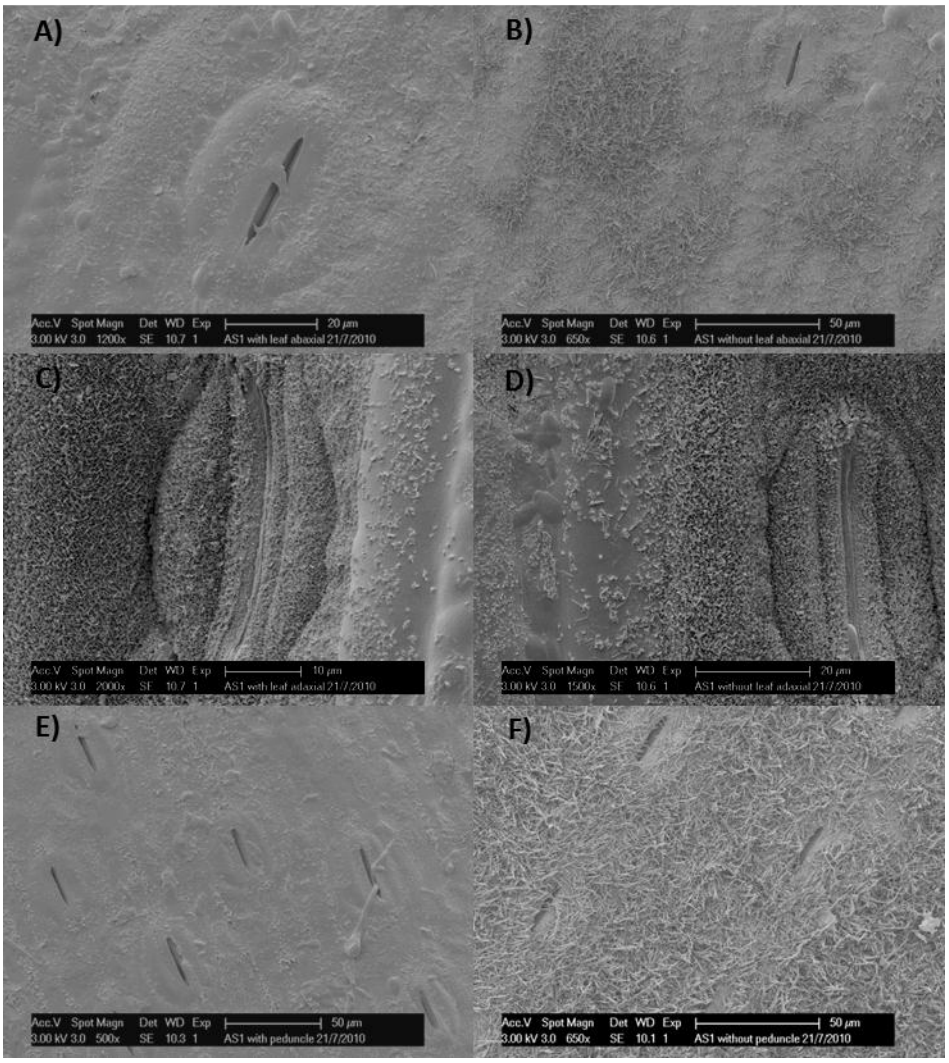
Yan, L., et al. (2003). "Positional cloning of the wheat vernalization gene VRN1." Proceedings of the National Academy of Sciences **100**(10): 6263-6268.
Yoo, K. S., et al. (2011). "Single Cystathionine β-Synthase Domain–Containing Proteins Modulate Development by Regulating the Thioredoxin System in Arabidopsis." The Plant Cell Online **23**(10): 3577-3594.
Zadoks, J. C., et al. (1974). "A decimal code for the growth stages of cereals." Weed Research **14**(6): 415-421.
Zhang, P., et al. (2006). "Genetic Diversity and Relationships of Wheat Landraces from Oman Investigated with SSR Markers." Genetic Resources and Crop Evolution **53**(7): 1351-1360.
Zhang, W., et al. (2005). "Molecular characterization of durum and common wheat recombinant lines carrying leaf rust resistance (Lr19) and yellow pigment (Y) genes from Lophopyrum ponticum." Theoretical and Applied Genetics **111**(3): 573-582.
Zhang, Z. and M. H. Saier, Jr. (2009). "A Novel Mechanism of Transposon-Mediated Gene Activation." PLoS Genet **5**(10): e1000689.

Appendix

A1 Additional cryo-SEM images of BC_2F_3 NILs

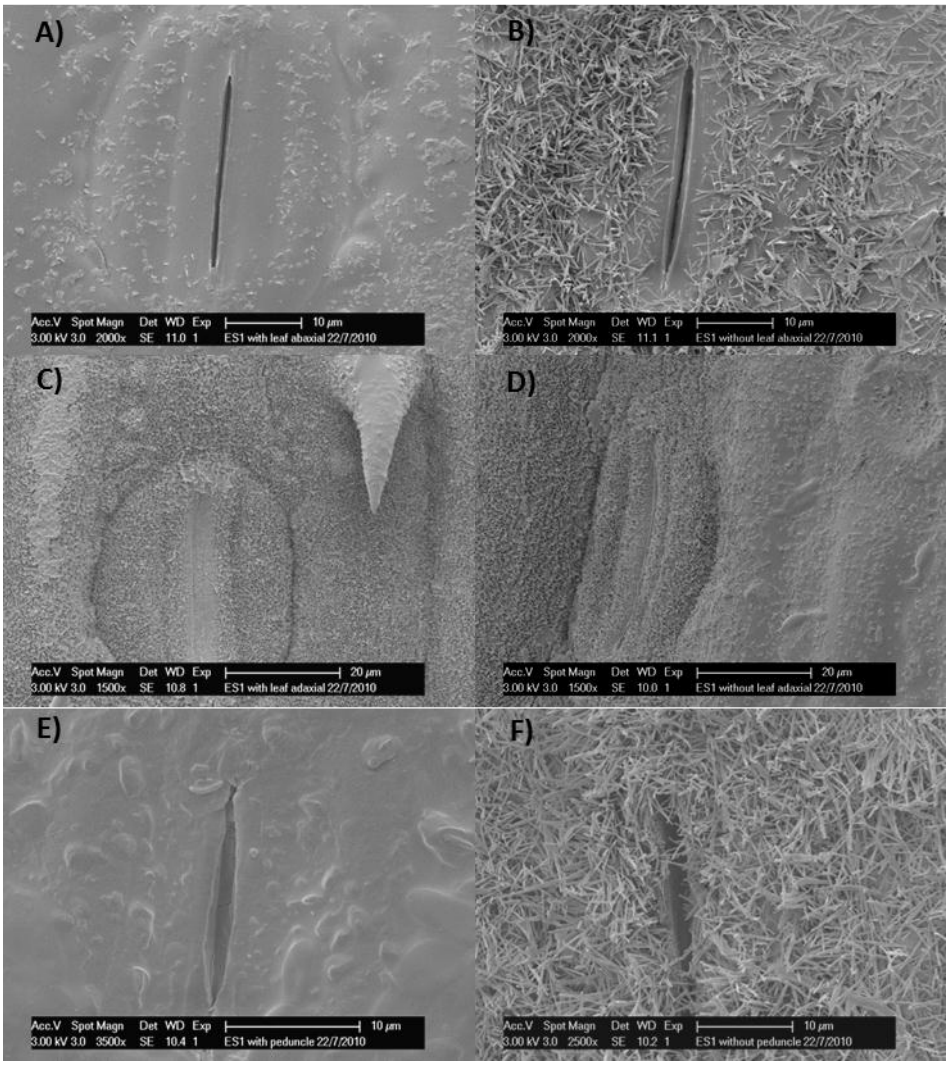
A1.1: SEM images of Alchemy NILs

A NIL pair of Alchemy background, with genotype *Iw1* (A, C, E) and genotype *iw1* (B, D, F), was analysed using cryo-SEM. Tubular shaped wax crystals are absent from *Iw1* lines, mimicking the Shamrock phenotype.



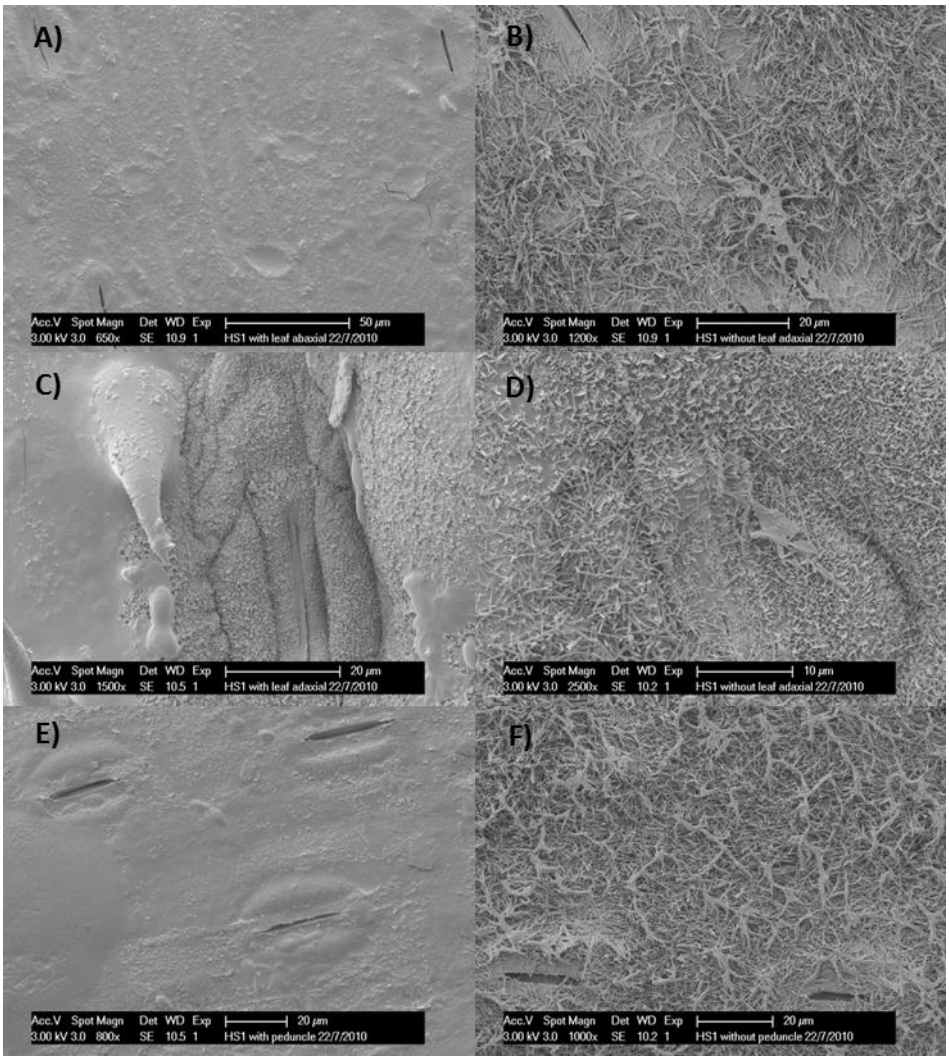
A1.2: SEM images of Einstein NILs

A NIL pair of Einstein background, with genotype *Iw1* (A, C, E) and genotype *iw1* (B, D, F), was analysed using cryo-SEM. Tubular shaped wax crystals are absent from *Iw1* lines, mimicking the Shamrock phenotype.



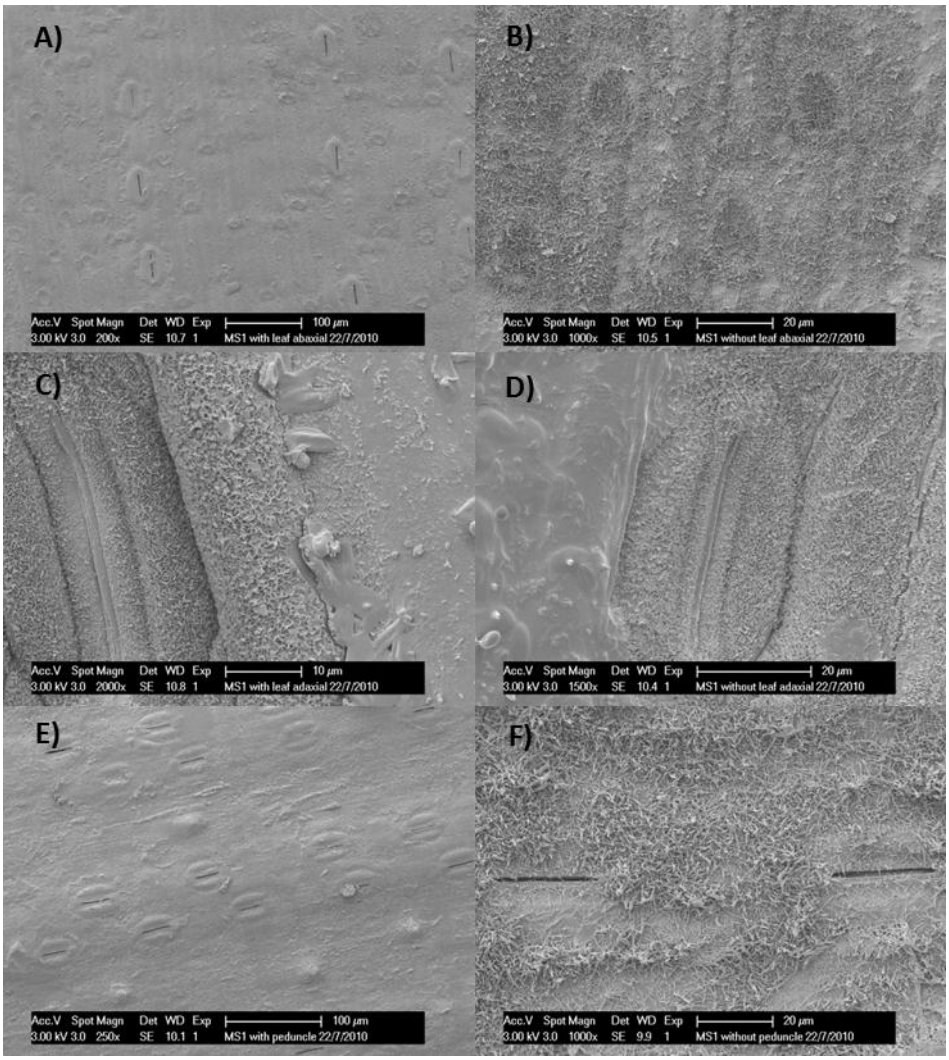
A1.3: SEM images of Hereward NILs

A NIL pair of Hereward background, with genotype *Iw1* (A, C, E) and genotype *iw1* (B, D, F), was analysed using cryo-SEM. Tubular shaped wax crystals are absent from *Iw1* lines, mimicking the Shamrock phenotype.



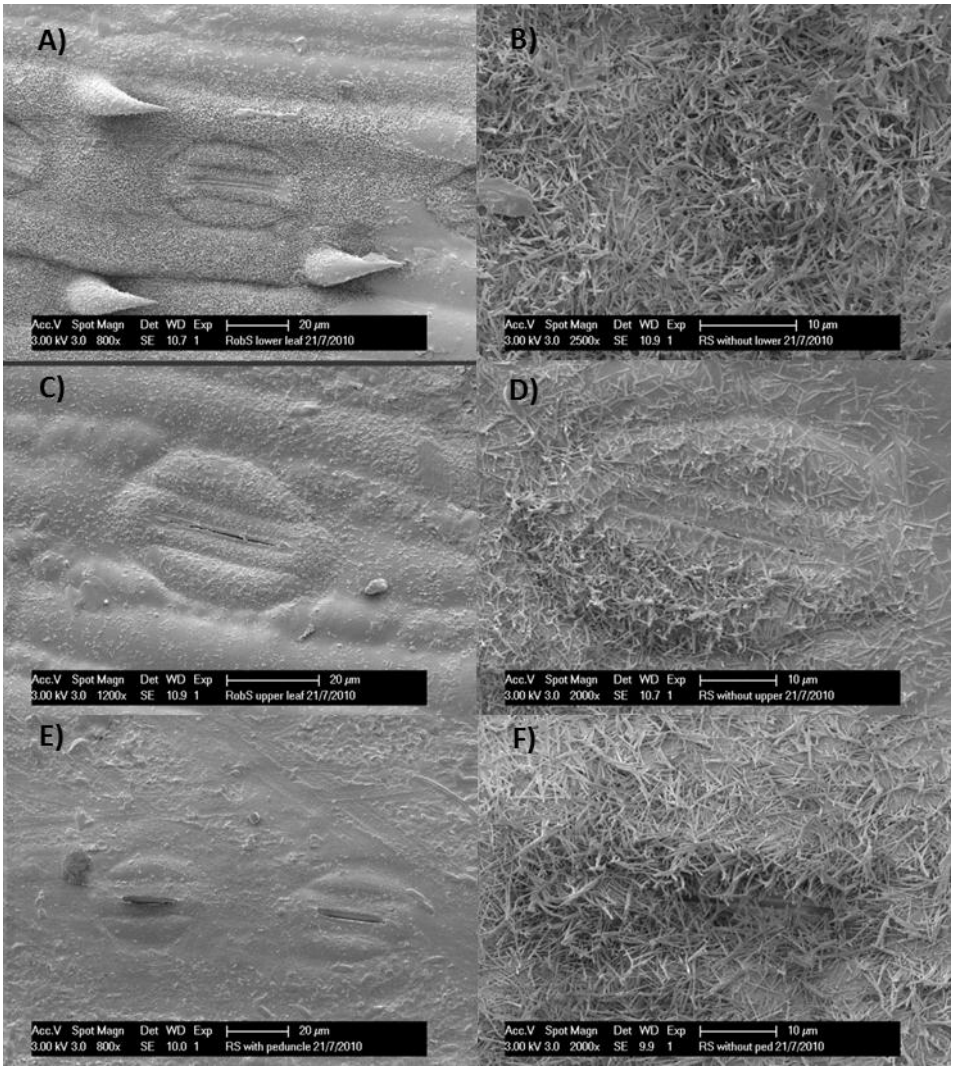
A1.4: SEM images of Malacca NILs

A NIL pair of Malacca background, with genotype *Iw1* (A, C, E) and genotype *iw1* (B, D, F), was analysed using cryo-SEM. Tubular shaped wax crystals are absent from *Iw1* lines, mimicking the Shamrock phenotype.



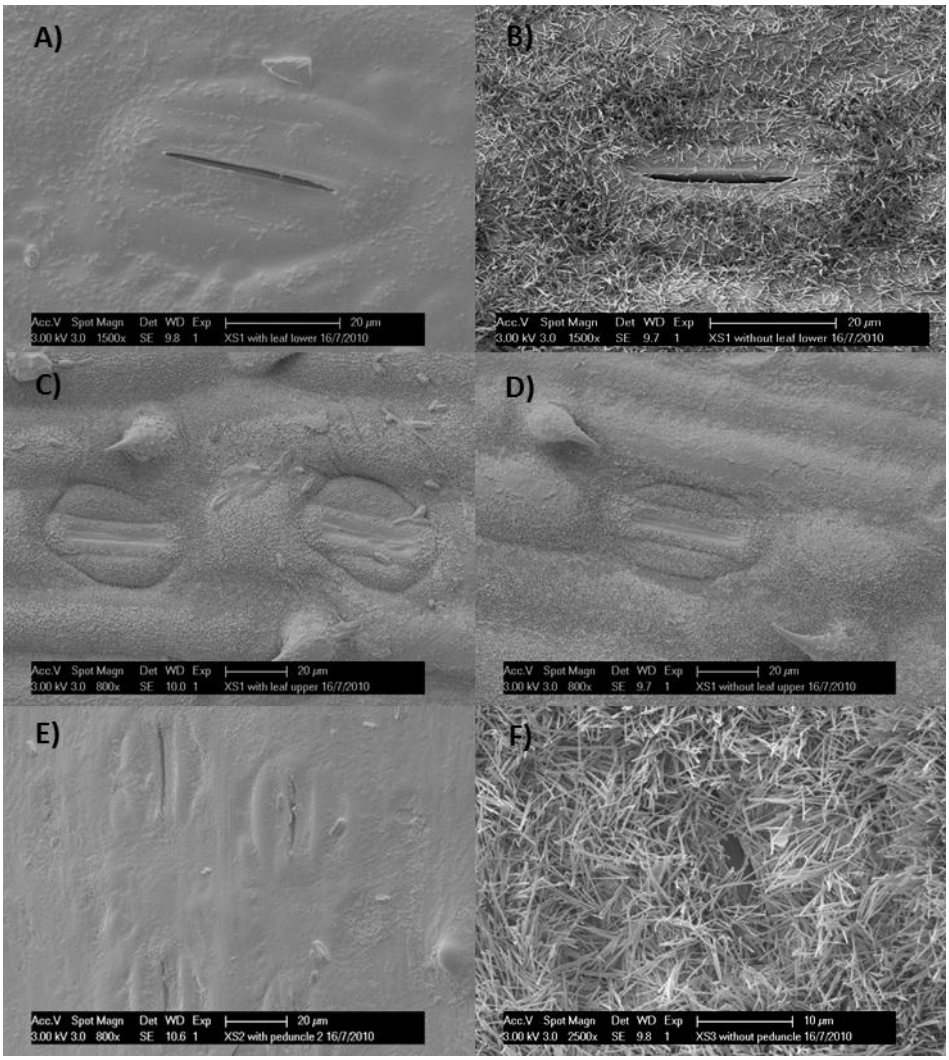
A1.5: SEM images of Robigus NILs

A NIL pair of Robigus background, with genotype *Iw1* (A, C, E) and genotype *iw1* (B, D, F), was analysed using cryo-SEM. Tubular shaped wax crystals are absent from *Iw1* lines, mimicking the Shamrock phenotype.



A1.6: SEM images of Xi19 NILs

A NIL pair of Xi19 background, with genotype *Iw1* (A, C, E) and genotype *iw1* (B, D, F), was analysed using cryo-SEM. Tubular shaped wax crystals are absent from *Iw1* lines, mimicking the Shamrock phenotype.



A2 Wax components identified in the GC-MS analysis of Shango and Shamrock

Number	Wax component	Quantified	Number	Wax component	Detected
1	C_{27} Alkane	Quantified	27	C_{22} Alkane	Detected
2	C_{29} Alkane	Quantified	28	C_{23} Alkane	Detected
3	C_{31} Alkane	Quantified	29	C_{24} Alkane	Detected
4	C_{24}POH	Quantified	30	C_{25} Alkane	Detected
5	C_{26}POH	Quantified	31	C_{26} Alkane	Detected
6	C_{28}POH	Quantified	32	C_{28} Alkane	Detected
7	C_{30}POH	Quantified	33	C_{33} Alkane	Detected
8	C_{32}POH	Quantified	34	C_{18}POH	Detected
9	C_{16} FA	Quantified	35	C_{20}POH	Detected
10	C_{18} FA	Quantified	36	C_{22}POH	Detected
11	C_{22} FA	Quantified	37	C_{34}POH	Detected
12	C_{24} FA	Quantified	38	C_{17} FA	Detected
13	C_{26} FA	Quantified	39	C_{20} FA	Detected
14	C_{28} FA	Quantified	40	C_{21} FA	Detected
15	C_{30} FA	Quantified	41	C_{23} FA	Detected
16	C_{32} FA	Quantified	42	C_{25} FA	Detected
17	β-DK	Quantified	43	C_{24} aldehyde	Detected
18	OH-β-DK	Quantified	44	C_{26} aldehyde	Detected
19	C_{28} aldehyde	Quantified	45	C_{32} aldehyde	Detected
20	C_{30} aldehyde	Quantified	46	C_{34} aldehyde	Detected
21	C_{23} AR	Quantified	47	C_{19} AR	Detected
22	MAR_{19}	Quantified	48	C_{21} AR	Detected
23	MAR_{21}	Quantified	49	C_{25} AR	Detected
24	MAR_{23}	Quantified	50	C_{27} AR	Detected
25	MAR_{25}	Quantified	51	unknown-1	Detected
26	MAR_{27}	Quantified	52	unknown-2	Detected
			53	unknown-3	Detected

A3 Markers used for fine mapping *Iw1* in Shango x Shamrock and Langdon x TTD140

Marker name	Forward Primer (5'-3')	Reverse Primer (5'-3')
JIC001	AACTTGAGGAGACTGAAGAGC	GCTTTGCATCTCTTTCATTTC
JIC002	GCCAGTGAGCAGATCAAACA	CGCTCCGATCTTCAGCTAAT
JIC003	GGCAAGGAGAGGTGAAGGA	CGAGGTTCTGCTCGATGATG
JIC004	TGGACGCACATGATGGAC	TCCTGCGACCGATAACATACT
JIC004-KASPar	TTGTTCCATTCCAGAACGG[A/C]	GAAGAAATCCCTGACGACATCG
JIC005	GCAAGAAGGCTGCTCAATTT	CCAACCATACTAGGAAGGTGAA
JIC006	GGATACATCTGGAGCCCTCA	TCGAGGAAGCAAAATCCAAC
JIC007	TGGAAAGCATTAGTCAGTCTTCTTT	AAAGCCATGCTTGAACTGGT
JIC007-KASPar	TCGTGTCCAGGCTAGGAAACT[T/G]	AACGTGGTCTCCAGGAGCGT
JIC008	CTCCATCCGGCACAAGAA	CTTCCCCGTCAGCACCTC
JIC009	CGTCTCCAAGAAGGACAAGG	TTGAGCATCAATGCAGCAG
JIC010	TCCTGAAGAAACACATGCAGA	CCTTTTCTCCAGCTCAATCG
JIC010-KASPar	GCTGCCGATTGCTTCTGCTAA[T/C]	ATATCACCTTGAACGGTGGTG
JIC011	GGAACACAGAAAACCAAAGGA	TCTTGCCAGTTGTTCATCCTC
JIC011-KASPar	GAAACAGGTTAGTATTGTATCTATTACA[T/C]	GAAGAATATAAGAGGATAAACACCG
JIC012	ACTTCACCCACCACAATGAG	CTGCATTGCCTTGTCCTTG
JIC012-KASPar	CCGAAATGGTTGAAAGCAAT[G/T]	AGGTGATGCTCCTGGGTCA
JIC013	CATTTCTGCATCAGGCCTTT	GGCAATGAGTTGGTGGTCTT
JIC014	CAAATTCGTCAGGGGAGAAG	TATCCACATCAAGCCGCATT
JIC015	GAGAAGGATGAGCCCTTTGT	AATCATCCACGTGCAGTTTCT

JIC015-KASPar	AGCGACAAATCTGAAGAATCC[A/G]	TAATGACATCCTGTGGTGTGGTGT
JIC016	CAAATTCATACCCTTGGAGAGCA	CTTCCAACTGTACCATACATATTTCA
Xgwm614	GATCACATGCATGCGTCATG	TTTTACCGTTCCGGCCTT
Xwmc25	TCTGGCCAGGATCAATATTACT	TAAGATACATAGATCCAACACC

A4 Markers used in physical map construction

Marker name	Product size	Forward Primer (5'-3')	Reverse Primer (5'-3')
JIC009[a]	TTD140: 695 bp, Langdon: 724 bp	TCACAGGATACCAGTTGTGAATG	CAAGGCCTGGAACACGTC
JIC011[a]	TTD140: 115 bp, Langdon: 127 bp	GCTTTGTCATCGGTGTTTATCCTCT	CCAAGAATAATCTGCAATAAACTAGC
JIC016[a]	Langdon: 437 bp	CAAATTCATACCCTTGGAGAGCA	CTTCCAACTGTACCATACATATTTCA
JIC017[b]	TTD140: 178 bp	CATGAGGTGAGGAAGCTGGAG	ATTACTATGTTCCCCAGCAAACG
JIC018[b]	TTD140: 89 bp	CAGGACTCCCTCAGAGACCTGT	TGTGCCAAAGTAAAAGAAAAATG
JIC019[b]	TTD140: 539 bp	GGCTCCATCCATACATATTTCATT	GAGAGACAATCACAATCATAATCATCAT
JIC020[b]	TTD140: 430 bp	ACCCCTTAATCCAGGACTCCCTCACC	AGTTACTGTCCTCTGAGTCCTCGTA
JIC021[c]	263 bp	GAAATCTCATAAGCTCCGTTAGCG	GATACTAGACAAATTCCATGAAGTTTG
JIC022[b]	TTD140: 503 bp	GTCCTGCCAGAGTCTATCGCTAT	CGGAGAGCACCTCAAGTATGTAG
JIC023[b]	TTD140: 266 bp	CACCCTCGTGGACACGATGT	GTGTGTGCATGGAAGGTTGATG
JIC024[b]	TTD140: 452 bp	TATACAAAAATGGAGGGAGTATTG	GCCGTACACCATGCCTATCT
JIC025[b]	TTD140: 486 bp	ACTCCCTCACCTACCAATCCCT	CAGGATCAATCTCTTCAGTGATG
JIC026[b]	TTD140: 355 bp	TGGGGACGCTAACAGAAGTC	ACGTGATGAACCAGAAGACAAG

JIC027[a]	TTD140: 600 bp	AATTTAAGCACGCCTGATCG	GGAGCTTCCTTGAATACCAACC
JIC028[b]	TTD140: 482 bp	ACGCCGCTGGGAACAGCACA	AGTAGGCATTTTCCTAATAGTGTCATC
JIC029[a]	TTD140: 379 bp	CAAGTTAGCATGAAGAATGAATTG	CCACTGATAAAAGGAAAGTTTTCTTG
JIC030[a]	TTD140: 399 bp	TCCCAACAAGTAGACCACATATCAC	CATGTCTGGTACACCATTAGTGGC
JIC031[a]	Langdon: 1,009 bp	GATCGGCAACCAGGACAAT	ATGACCTGGTTCGCTGGTTCTA
JIC032[b]	Langdon: 307 bp	TGACAAGCATCCTCAAGGTAGTCT	GAAGCAAACATGATGATTCGG
JIC033[a]	Langdon: 529 bp	TCTTGGGAGAACAGAATTGACAT	ATATGTTCACGCTAACGGGTAGCA
JIC034[a]	Langdon: 465 bp	CGAGGATACACGCTACTGGTAAGT	AGGACCAATGGTAACAATGGAGAAG
JIC035[a]	Langdon: 405 bp	GTACCAACTATCATCGGATACAAGG	ATAGGGCTCAAACATAACGGAAGA
JIC036[a]	TTD140: 490 bp, Langdon: 515 bp	TGTAACATACATTTCTTTTGAATGGTAC	CAGAACTCTCTCGTTACTAGGCTG

A5 Markers used to creates probes for radioactive hybridization

Marker name	5' - 3' sequence	Orientation	Product size	Designed on BAC
Iw614	TATCGTCATCATCATCGTCGCAG	Left primer	TTD140: 305 bp	329M10
Iw666	CAAAGTTTGAGGACCAGGCTC	Right primer		
Iw663	TGATGGCGAACACGGCTCAG	Left primer	TTD140: 450 bp	329M10
Iw625	CAATGAGTTTGCACCGTTCATA	Right primer		
Iw626	CACCTTAATCTCATAATTGCGGT	Left primer	TTD140: 361 bp	329M10
Iw627	ATCCATTTTGAAGACAAGTATTTCC	Right primer		
Iw1086	GTCCTGCCAGAGTCTATCGCTAT	Left primer	TTD140: 503 bp	170H13
Iw1087	CGGAGAGCACCTCAAGTATGTAG	Right primer		
Iw1107	CCATAGGTGCGGCTCTGACT	Left primer	TTD140: 383 bp	170H13
Iw994	AGAATAGGTAGCTCCGACGTGC	Right primer		

Iw1107	CCATAGGTGCGGCTCTGACT	Left primer	TTD140: 213 bp	170H13
Iw993	ATGGACGTATCACGATCTACATCAAC	Right primer		
Iw1108	TCCTGAGACTTGGTTCCAATGTG	Left primer	TTD140: 373 bp	170H13
Iw993	ATGGACGTATCACGATCTACATCAAC	Right primer		
Iw614	TATCGTCATCATCATCGTCGCAG	Left primer	TTD140: 305 bp	329M10
Iw666	CAAAGTTTGAGGACCAGGCTC	Right primer		
Iw663	TGATGGCGAACACGGCTCAG	Left primer	TTD140: 450 bp	329M10
Iw625	CAATGAGTTTGCACCGTTCATA	Right primer		
Iw626	CACCTTAATCTCATAATTGCGGT	Left primer	TTD140: 361 bp	329M10
Iw627	ATCCATTTTGAAGACAAGTATTTCC	Right primer		
Iw497	ACGCCGCTGGGAACAGCACA	Left primer	TTD140: 482 bp	551i08
Iw447	AGTAGGCATTTTCCTAATAGTGTCATC	Right primer		
Iw1175	TTTGGATGTAACTTTTTGAGTAGATG	Left primer	TTD140: 399 bp	784K20/577P12
Iw1176	CATGTCTGGTACACCATTAGTGGC	Right primer		
Iw1174	TCCCAACAAGTAGACCACATATCAC	Left primer	TTD140: 262 bp	784K20/577P12
Iw1176	CATGTCTGGTACACCATTAGTGGC	Right primer		

A6 Markers used for the analysis of candidate genes

Marker name	5' - 3' sequence	Product size
F1	AGGACTTCACGATACCACACTAACTC	984 bp
R1	TCATCTGAGCTACACCAAGGG	
F2	ACTTCTCCATTATCCACATGCA	709 bp
R2	CTTGTTTGATGAGCTTGCTTCC	
F3	CGATAATGTAATACTGCCGAGTACA	701 bp

R3	CACTTTGTGATGCCTTCCTTG	
F4	ACCACTCGGAGGAACTACTCAAC	989 bp
R4	TGTTCTTATTCTCCCATTTGTGCT	
F5	CAGGGAAGAAGAACCACCG	1,225 bp
R5	AAGGGAGAAGCCGTGCGGAACACA	
F6	GTTCAAGTTCAACGCCAAGG	1,176 bp
R6	ACATTGAGTGTTCCATTGACTGCTG	
F7	CACCCGAATTGTCTTTCACCTC	1,096 bp
R7	ATTGTCTTACACGGGCAAGCTC	
F8	CCATCGCTTAGGCAGAGGTT	1,149 bp
R8	CTACACTGCAATTATGGGGTCA	
F9	CAAAACAGCTAACGGAGAACACTG	904 bp
R9	CAAACTCGTAATGCTCTCTTGAGAT	
F10	ATCTCAAGAGAGCATTACGAGTTTG	1,504 bp
R10	GCTCCAGTATGATGTCGTAGGATG	
F11	ATCAGCCTCTCACCTAGCCTG	1397 bp
R11	ATGTGGTCAAGATCATCATTGC	
F12	ATTAGCAGAACCGACATTGTTG	1,200 bp
R12	GCGAAAGTACCCTTGTGAGCT	
F13	GATTTTTAAGAGGATGCCATCAAG	763 bp
R13	GGAGTTGCAGATGAAGAAGG	
F14	GTGTCAGCAGGAATATAAGCAC	1,234 bp
R14	GAAAGAACCTCCGACAAACTG	
F15	TATGCTTGTGGTGGATGGCG	884 bp
R15	TCGAGCTGACAAGGCTTAATACG	
F16	GAACTACTAAAGAAGCTGAAGGTTGG	917 bp
R16	CAAATATGGAGGTCAGAAAGAGGA	

F17	CAAATGGATTGCCTATCAAGTATGAC	1,085 bp
R17	GTATGTGATTACACCATTACCAAATTC	
F18	CTCACTAATTCAATTCACAGGTAGTTC	949 bp
R18	TGTGTTGGAATATGACTAAATCAGTTC	
F19	CATGGCTCGCTGCTAGTATAGTAC	673 bp
R19	GTCAAATGGAGTGCCAGTGC	
F20	GCACTGGCACTCCATTTGAC	1,114 bp
R20	AATAGTGCGACTGTGTGTGTCCTA	
F21	TCACAGGATACCAGTTGTGAATG	TTD140: 691 bp / Langdon: 724 bp
R21	CAAGGCCTGGAACACGTC	
F22	GCGTTCTGATTGGTTCATGG	979 bp
R22	GAGAATGACCAACAGGGTATGATC	
F23	GTTCAAGTTCAACGCCAAGG	1,547 bp
R23	GCATCATTGGACTCGACCTACA	
WPK_01	TTCaCCCaCCTCTACCAGGTCG GGTCAAGTAACTCAGCCTTGGG	381 bp

Acknowledgements

The completion of this thesis and my entire PhD would not have been possible without the help and support of many people.

I would like to thank my family, especially my parents and my sister, for their continued support and belief in me. Buzka!

Likewise, I want to thank the love of my life Barbara for her support and love.

I have to thank Cristobal Uauy for introducing me to the exciting world of wheat genetics, his continued support and friendship and his lightning speed proof-reading. ¡Gracias!

I want to thank Ruth, Philippa, Pauline, Charlotte, Yang, Meluleki, Lena, Albor and Angela for great friendship, help and having some really good times!

To the guys and gals of the Uauy lab, who are not only good fun to work with but also good friends, I would also like to say: “Thank you”. It never gets boring with you!

The same goes for the people I shared and still share an office with: Emilie, Jordi, Nick, Philippa, Sarah and Sarah. Thank you for the laughs and the biscuits/cookies/cake/chocolate.

Special thanks go to Sarah and Sam (and the chicks!), Kim and Frazer (and baby Hugo), Christine for nice BBQs, good conversations, friendship, laughter and some really yummy cake.

I want to thank Matthew and Elise Tucker for mentoring and befriending a geeky kid, for “sorting me out”, for friendship and an Australian outlook on life.

Likewise I want to thank Michael Lenhard for “luring” me to JIC and the subsequent help and friendship; Holger Breuninger for friendship, many jokes, great game nights and generally god fun; I also want to thank the other “Lennies” and “Lauxies” Son, Adrien, Marko, Sven, Lore, Lena (again), Eric, Philipp, Annika, Enno, Diana, Norbert, Tobias, Sonia, Minako and the Professor for good times.

Thank you to Adrian, Max, Sarah, James, Sonia, Hélène, Joelle, Lars and everybody else who helped me in this project.

Thank you to everybody who has been (or still is) at JIC and extended good humour, friendship and help.

Thank you to the friends I had to leave behind; Florian, Sedl, Sigi, Hölli, Mirko, Bernd, Chris, Jörg, Fabian, Daniel: I shall remember the good old days. I also want to thank everybody who ever played with the SC Freibad: Thank you for a fun time!