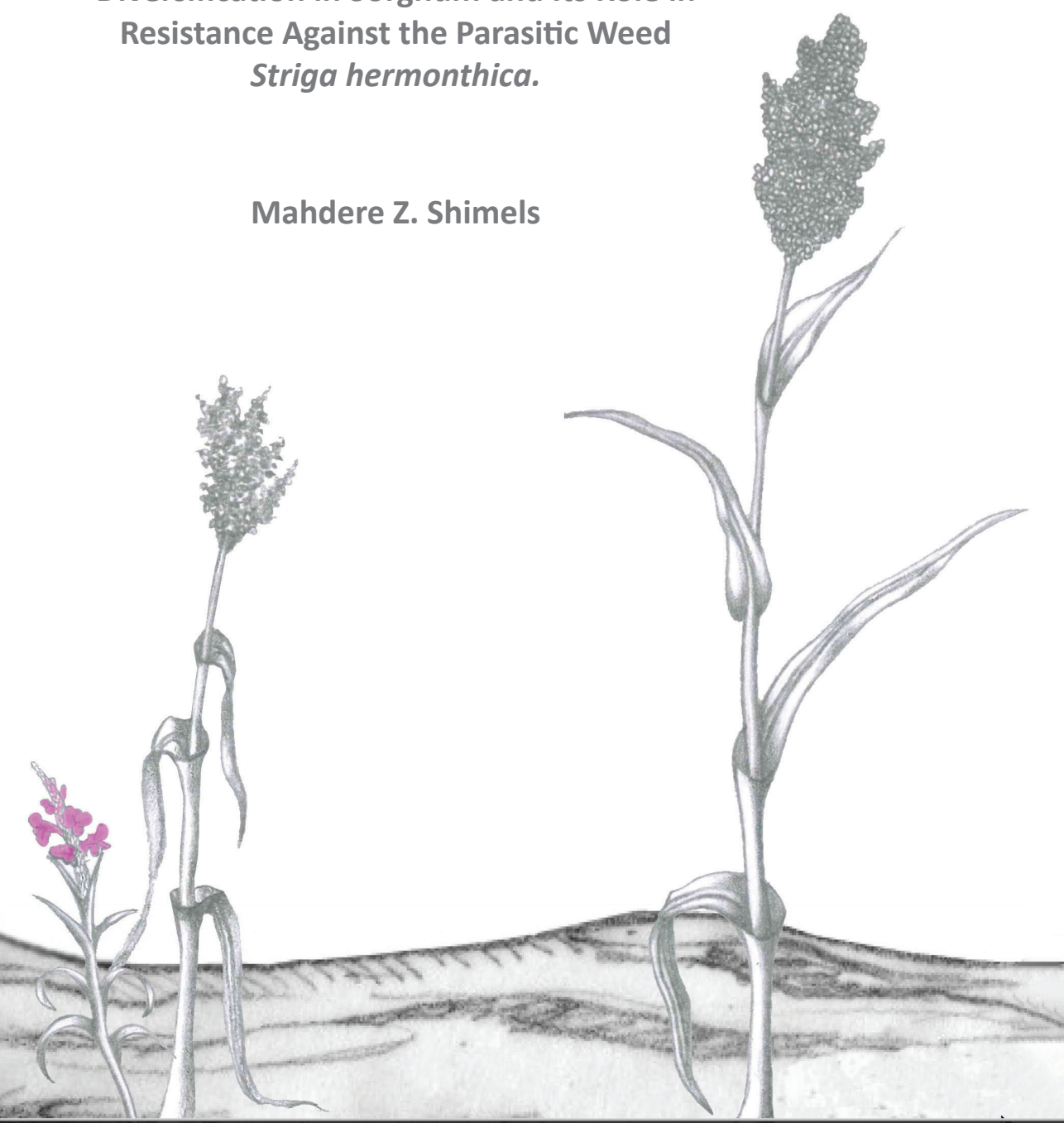
The Mechanism Underlying Strigolactone Diversification in Sorghum and Its Role in Resistance Against the Parasitic Weed *Striga hermonthica*.

Mahdere Z. Shimels



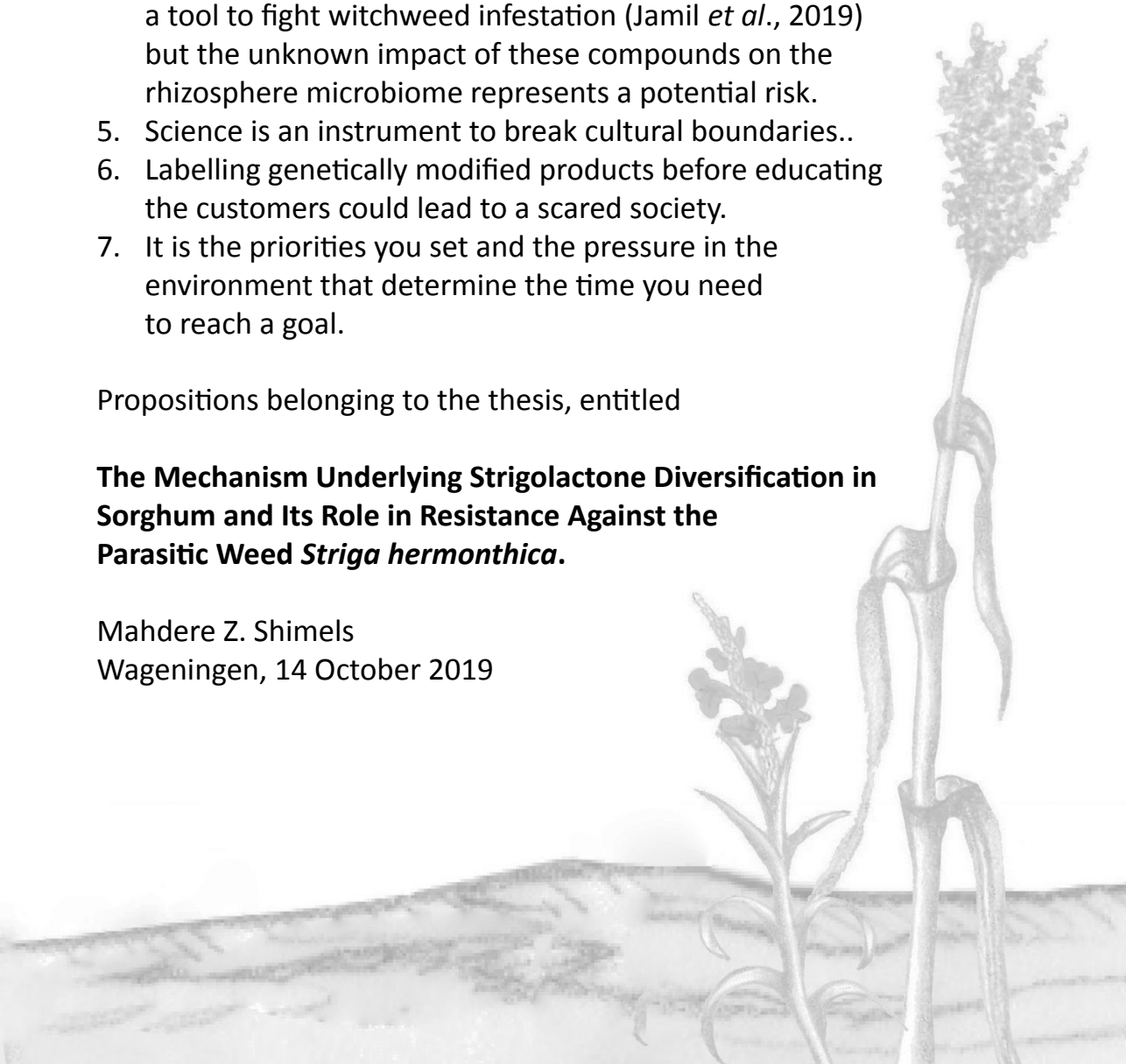
Propositions

1. The production of high levels of *Striga* germination stimulants by some sorghum genotypes does not necessarily imply they are susceptible to the parasite (this thesis).
2. Sulfotransferase, SbSOT4A, is a key enzyme in strigolactone diversification. (this thesis).
3. Despite well-known benefits of retaining “transgenerational memory” (Mozgova et al., 2019), the loss of it is also an important survival mechanism for organisms.
4. Synthetic analogs of germination stimulants have been proposed as a tool to fight witchweed infestation (Jamil *et al*., 2019) but the unknown impact of these compounds on the rhizosphere microbiome represents a potential risk.
5. Science is an instrument to break cultural boundaries..
6. Labelling genetically modified products before educating the customers could lead to a scared society.
7. It is the priorities you set and the pressure in the environment that determine the time you need to reach a goal.

Propositions belonging to the thesis, entitled

The Mechanism Underlying Strigolactone Diversification in Sorghum and Its Role in Resistance Against the Parasitic Weed *Striga hermonthica*.

Mahdere Z. Shimels
Wageningen, 14 October 2019



The Mechanism Underlying Strigolactone Diversification in Sorghum and Its Role in Resistance Against the Parasitic Weed *Striga hermonthica*

Mahdere Z. Shimels

Thesis committee

Promotors

Prof. Dr Harro J. Bouwmeester
Professor of Plant Physiology
Wageningen University & Research
Chair of Plant Hormone Biology
Swammerdam Institute for Life Sciences, University of Amsterdam

Co-promotor

Dr C.P. Ruyter-Spira
Biosafety Officer
Wageningen University & Research

Other members

Prof. Dr Paul Struik, Wageningen University & Research
Dr Jonne Rodenburg, University of Greenwich, United Kingdom
Prof. Dr Jos Raaijmakers, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen
Dr Robert Schuurink, University of Amsterdam

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

The Mechanism Underlying Strigolactone Diversification in Sorghum and Its Role in Resistance Against the Parasitic Weed *Striga hermonthica*

Mahdere Z. Shimels

Thesis

Submitted in in fulfilment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Monday 14 October 2019

at 4 p.m. in the Aula.

Mahdere Z. Shimels

The Mechanism Underlying Strigolactone Diversification in Sorghum and Its Role in Resistance Against the Parasitic Weed *Striga hermonthica*,

154 pages,

PhD thesis, Wageningen University, Wageningen, the Netherlands (2019)

With references, with summary in English

ISBN 978-94-6395-091-6

DOI https://doi.org/10.18174/498779

Contents

Chapter 1

General Introduction

Striga infestation

Presently, we rely on agriculture to produce most of the food we need to feed the world. Sorghum (*Sorghum bicolor* [L.] Moench) is among the top five cereals grown worldwide after maize, rice, wheat and barley. Globally, it is cultivated on more than 50 million hectares of land annually (FAOSTAT 2012). According to a report from 2012, Nigeria is the largest producer of sorghum in Africa followed by Ethiopia and Sudan (FAOSTAT 2012). These major production areas are generally known for low rainfall and low soil fertility, conditions that are not suitable for other cereals such as maize, wheat and rice. However, sorghum is quite resilient against abiotic stresses. Nevertheless, for farmers, growing sorghum comes with challenges such as pests and diseases. Infestation of sorghum fields by the parasitic weeds *Striga hermonthica* and, to a lesser extent, *Striga asiatica*, stands out as a major problem especially in sub-Saharan Africa (Figure 1) (Mohamed et al. 2006; Ejeta 2007a).

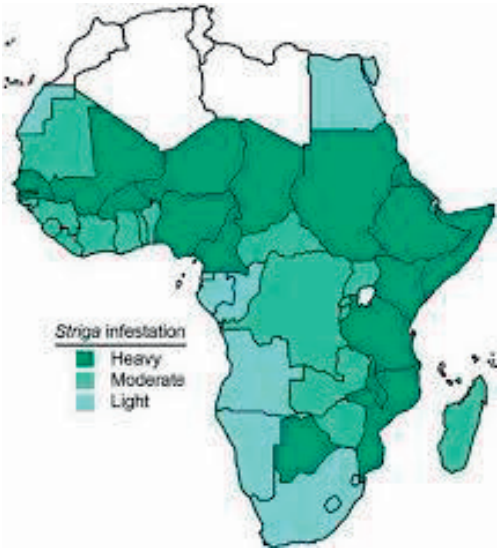


Figure 1. Distribution of *Striga hermonthica* in Africa (Ejeta 2007a)

There are around 30 *Striga* species identified and for most, the Poaceae (grasses) are their host (Spallek et al. 2013). *Striga hermonthica* and *Striga asiatica* are the two *Striga* species that affect cereals such as millet, cowpea, sorghum, maize and rice while *Striga gesnerioides* infects dicots such as cowpea, all staple foods in large parts of the world (Berner et al. 1998; Parker 2009; Sibhatu 2016). *Striga* affects over 40 countries and impacts over 300 million farmers (Runo et al. 2018; Spallek et al. 2013). A *Striga* infection causes stunted growth, early senescence and thereby often complete failure of a crop, by draining its nutrients and water (Ueda et al. 2015; Parker and Riches, 1993). The weed also triggers physiological changes in the host plant such as reduced water use efficiency. Furthermore, the balance of growth regulators in the host plant is changed, which affects its photosynthesis negatively (Rodenburg et al. 2017).

An infection with *Striga* can result in up to 70-100% yield losses (Ejeta 2007a). For *Striga* to successfully parasitize the host plant, three conditions should be fulfilled, suitable temperature, moisture and the presence of the host plant which can be detected through signaling molecules exuded by the host plant itself (Haussmann et al. 2000).

The signaling molecules referred to above, trigger *Striga* seed germination and are called strigolactones (Yoneyama et al. 2010; Bouwmeester et al. 2007). Strigolactones are carotenoid derived molecules and are mainly produced in the roots and exuded into the rhizosphere. The presence of these compounds at extremely low concentrations of μM level can stimulate the germination of the parasitic seeds (Nomura et al. 2018; Kim et al. 2010). The production of strigolactones has been linked to land colonization during the Silurian period (430 million years ago, MYA) and the emergence of the first Embryophytes (land plants). The evolutionary progress of these gametophyte bearing simple unisporangiate sporophytes continued throughout the Devonian age.

Figure 2. Sorghum field infested by *Striga hermonthica*.
Photo by Mahdere Shimels, Ethiopia

Exposed to the dryer environment, plants needed to develop mechanisms to absorb water and nutrients and move it throughout the plant, which leads to the formation of vascular plants. It was around the end of the Devonian age (360 MYA) when most of the plant organs as we know today such as vasculature, roots and shoots evolved.

Over the course of this 450 MYA of land plant evolution, strigolactones played a significant role. They were produced by charophyte algae before plants invaded the land. They were also present in several lineages of plants such as bryophytes: hornworts, liverworts and mosses. Studies suggest rhizoid elongation in Streptophyta and promoting growth of bryophytes was the primary role of strigolactones during evolution. Other studies have shown that exudates of charophyte algae did not induce germ tube branching of *Gigaspora rosea* which supports that this was not their primary role. Their role in initiating colonization of plant roots by arbuscular mycorrhizal (AM) fungi probably evolved with land colonization (Bouwmeester et al. 2007). Fossil evidence on the first mycorrhizal symbiosis with plants dates back to 407 MYA during the Devonian period (Strullu-Derrien et al. 2018). Fossils show that at that time plant roots already associated with fungi belonging to the Glomeromycota, which also today are the most common fungi to associate with plant roots in more than 80% of terrestrial plants.

AM fungi that also belong to the Glomeromycota are obligate biotrophs that cannot complete their life cycle without colonizing a host plant. Strigolactones induce the pre-symbiotic stage in AM fungi, which is characterized by hyphal branching of the germinated spores. In addition to hyphal branching, strigolactones may stimulate spore germination or act as a chemoattractant to direct the AM hyphae towards the roots. Once it finds the host root, AM fungi form a hyphopodium on the host root where its hyphae emerge from and penetrate the root cells. Then the fungus forms arbuscules - from the Latin arbusculum, small tree - the major site for nutrient exchange between the fungus and the plant that are formed in the cells of the root cortex by repeated branching of the hyphae (Bücking et al. 2012). Inorganic phosphorus is taken up from the soil by the hyphae of the fungus and accumulates in its vacuole as polyphosphates (inorganic phosphate) that are later hydrolyzed to an available form of phosphate and delivered to the host plant. Plants with AM fungi were found to contain up to 4-fold higher inorganic oxidized form of phosphate than plants that were fertilized but that did not create a symbiosis with AM fungi (Nouri et al. 2014). In return, the host plant provides carbon from photosynthesis to the AM fungi. This crucial role of strigolactones in facilitating symbiosis explains why plants produce and exude strigolactones into the rhizosphere despite its negative role in signaling their presence to parasitic plants.

Strigolactones have also been shown to regulate plant architecture. One of the most studied is its role as a branching hormone. Strigolactone deficient mutants in pea, *Arabidopsis*, rice and petunia showed enhanced shoot branching (Umehara et al. 2008; Gomez-Roldan et al. 2008). Interestingly, under optimal conditions, strigolactones can serve as both growth promoters or inhibitors of different organs of plants. For instance, above-ground, strigolactones repress bud outgrowth but can stimulate secondary growth of the stem and promote internode elongation (de Saint Germain et al. 2013; Agusti et al. 2012). Underground, strigolactones promote root hair elongation but repress lateral root formation (Ruyter-Spira et al. 2011; Wu et al. 2017). Strigolactones have also been demonstrated to play a role in biotic and abiotic stress responses. The best-studied example of this is their increased production and exudation during plant exposure to low nutrient conditions, especially phosphate (Jamil et al. 2011; Yoneyama et al. 2007). This upregulation of strigolactone production under low phosphate results in repression of shoot branching and the adaptation of root architecture (Sun et al. 2016; Gomez-Roldan et al. 2008). Under low phosphate conditions, the shoot-to-root ratio of plants changes. As a result, root hairs elongate and their number increases allowing the plant to find the soil part that is richer in phosphate (Kohlen et al. 2011; Umehara et al. 2010). Through this mechanism, plants can promptly respond to changing environmental conditions.

Strigolactone biosynthesis

Strigolactones are biosynthetically derived from the carotenoids and their biosynthesis involves a series of enzymatic reactions, which starts with the isomerization of all-trans-β-carotene to 9-cis-β-carotene by the β-carotene isomerase, DWARF27 (D27) (Sorefan et al. 2003; Alder et al. 2012). This 9-cis-β-carotene subsequently serves as substrate for Carotenoid Cleavage Dioxygenase 7 (CCD7) that converts it into the C13 β-ionone and C27 9-cis-β-apo-10`-carotenal (Alder et al. 2012; Sorefan et al. 2003; Schwartz et al. 2004). Further cleavage of 9-cis-β-apo-10`-carotenal is catalyzed by CCD8 and this cleavage reaction results in the formation of carlactone with specific 2'R configuration of the D-ring (Alder et al. 2012; Seto et al. 2014). Intriguingly, the CCD7 enzymes of different species such as *Arabidopsis* and pea both showed the highest affinity towards 9-cis-β-carotene but could also cleave 9-cis-zeaxanthin which results in the formation of a hydroxylated 9-cis-apo-10'-carotenal, which upon further cleavage results in C3-hydroxylated carlactone (Bruno et al. 2014). So far, there are no indications, however, that this C3-hydroxylated carlactone is a strigolactone precursor in vivo. Thus, it seems that carlactone is the precursor for all strigolactones, canonical as well as non-canonical. The non-canonical strigolactones such as carlactonoic acid (CLA), methyl carlactonoate (MeCLA), zealactone, zeapyranolactone, avenaol and heliolactone do not have the conventional ABC-ring that the canonical strigolactones have, but all strigolactones, including the non-canonical, have the D-ring and the enol ether bridge (Figure 3) (Xie 2017; Wang et al. 2018). Also, the non-canonical strigolactones have been shown to have strigolactone-like activity such as inducing seed germination of parasitic plants and hyphal branching in AM fungi. For instance, zealactone that was isolated from the root exudate of maize was shown to induce *Striga* seed germination (Nomura et al. 2018; Charnikhova et al. 2017). In *Arabidopsis*, the cytochrome P450 MORE AXILLARY GROWTH 1 (MAX1), was shown to catalyze the conversion of carlactone to CLA which is further converted to methyl carlactonoate (MeCLA) by an unknown methyltransferase (Figure 3) (Abe et al. 2014). LATERAL BRANCHING OXIDOREDUCTASE (LBO) can subsequently convert MeCLA into an as yet unidentified non-canonical strigolactone with 16 Da higher mass than its substrate (Brewer et al. 2016). Genes involved in the production of other non-canonical strigolactones such as in maize and black oats are yet to be discovered.

So far, more than 20 canonical strigolactones have been identified. They share a common skeleton of a tricyclic lactone (ABC ring) connected to a butenolide D-ring in 2'R configuration via an enol ether bridge. They can be grouped into two types based on the stereochemistry of the B-C ring junction; strigol-type, with β orientation, and orobanchol-type, with α orientation (Wang et al. 2018). In rice, the genes that catalyze the biosynthesis of two orobanchol type strigolactones, 4-deoxyorobanchol and orobanchol, have recently been discovered. The cytochrome P450 MAX1 homolog Os900, called carlactone oxidase, catalyzes the conversion of carlactone to 4-deoxyorobanchol which is subsequently converted to orobanchol by another MAX1 homolog Os1400, called orobanchol synthase (Figure 3) (Zhang et al. 2014). Carlactone can also be converted – though the enzymatic mechanism is as yet

unknown - to a strigol-type strigolactone, 5-deoxystrigol, which is believed to be the precursor of all other strigol-type strigolactones (Al-Babili et al. 2015). Modifications of either 4-deoxyorobanchol or 5-deoxystrigol such as hydroxylation, acetylation, oxidation and decarboxylation will result in the formation of other orobanchol- and strigol-type strigolactones, respectively. For instance, in a feeding assay, 5-deoxystrigol has been shown to be the precursor of the strigol-type strigolactone, sorgomol, likely through hydroxylation at C9 (Figure 3) (Xie et al. 2010; Motonami et al. 2013). Further oxidation of the created hydroxymethyl group of sorgomol followed by decarboxylation was hypothesized to result in sorgolactone production (Motonami et al. 2013). However, biosynthesis of most of the strigolactones is still only postulated and the genes encoding these putative enzymatic activities are still unknown (Figure 3).

It is important to note that different plant species produce and/or exude different blends of strigolactones both in amount and/or type. Several studies showed that even different cultivars of the same species can produce/exude different blends of strigolactones (Gobena et al. 2017; Yoneyama et al. 2008; Awad et al. 2006; Yoneyama et al. 2011). This may have an impact on the performance of the plant either through the hormonal function of strigolactones or because of their rhizosphere signaling role (Lopez-Obando et al. 2015; Yoneyama et al. 2015). For instance, the exudates of different sorghum genotypes have a different composition of strigolactones (Mohemed et al. 2018). Those with high levels of 5-deoxystrigol and sorgomol showed significantly higher germination stimulant activity on *Striga* seeds. This was confirmed under field conditions where these genotypes have high *Striga* infestation when compared to genotypes with exudates that contain low levels of 5-deoxystrigol and higher levels of orobanchol (Mohemed et al. 2016). This is supported by in vitro studies showing that, for example, strigol-type strigolactones have a higher *Striga* germination stimulant activity than orobanchol-type strigolactones (Mohemed et al. 2018; Nomura et al. 2013). Interestingly, this stereochemically specific response is *Striga* species-dependent. In contrast to *S. hermonthica*, *S. gesnerioides* is more sensitive to orobanchol-type than strigol-type strigolactones (Ueno et al. 2011; Vurro et al. 2019). On the other and, the induction of hyphal branching activity in AM fungi *Gigaspora margerita* was not affected by the stereochemistry of strigolactones. Rather, modification on the A- AB-ring was shown more determinant for their activity towards inducing hyphal branching activity (Akiyama et al. 2010). These different responses of different organisms to different strigolactones can also be seen in bioassays with root exudates. Different HPLC fractions of root exudates collected from rice showed contrasting activity towards *Striga* germination, but also AM fungal hyphal branching (Cardoso et al. 2014).

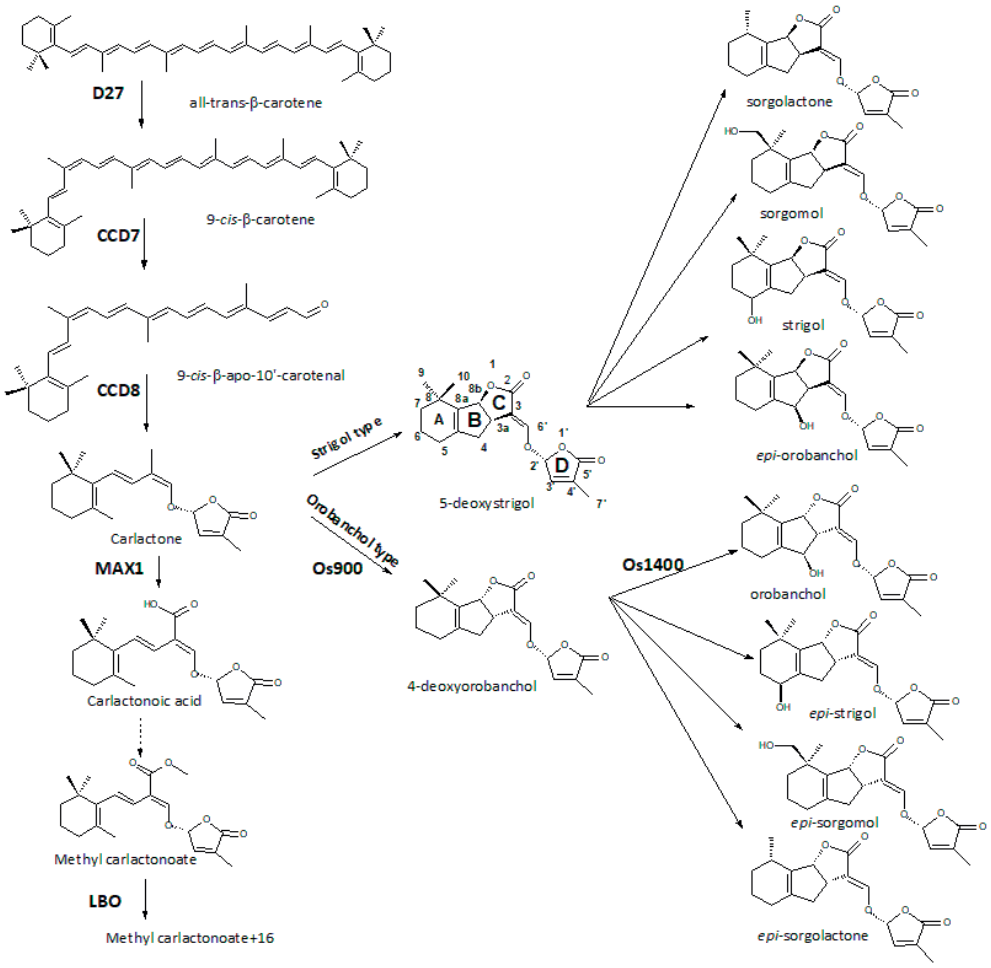


Figure 3. Schematic representation of part of the (postulated) strigolactone biosynthetic pathway.

Strigolactones as plant hormones

Because strigolactones are also plant hormones, it is of interest to review how they are involved in regulating homeostasis and how they interact with other plant hormones. The intricate coordination of the developmental, metabolic and physiological activity and responses in plants requires a well-coordinated machinery and communication within and outside the plant. Plant hormones such as abscisic acid (ABA), auxin (IAA), brassinosteroids (BR), cytokinin (CK), ethylene (ET), gibberellin (GA), jasmonate (JA), salicylic acid (SA) and strigolactones are the major regulators of this. Many studies have shown that the biosynthesis of these hormones is tightly regulated and that they are interconnected through cross-talk. For instance, IAA flux is also regulated by strigolactones and, vice versa, IAA promotes strigolactone biosynthesis (Hayward et al. 2009). Together this creates a fine-tuned balance between these two hormones to regulate shoot branching (Zhan et al. 2018). This complex

network is tightly regulated to decide the synthesis, concentration, transportation and modifications of each hormone based on the internal situation such as its developmental stage and/or in response to external circumstances such as the environment it is exposed to. The level of plant hormones can be regulated both by biosynthesis and inactivation processes. Methylation is a good example of hormone modification that changes the activity. Methylation of hormones may turn them into a volatile, making them more active as signaling molecules but less active as hormones (Spíchal et al. 2004). Methylation of GA in *Arabidopsis* and rice leads to its inactivation (Varbanova et al. 2007). Glycosylation is another form of hormone modification that can lead to an inactive storage form of several different plant hormones or degradation as is the case for CK glycosylation (Varbanova et al. 2007; Piotrowska et al. 2011). The genes that are encoding these modification enzymes are regulated by hormonal and environmental signals. Other modifications such as oxidation, hydroxylation, sulfation and amino acid conjugation are also common in plants (Sorefan et al. 2003). Interestingly, a single plant hormone can be subjected to different modifications for different purposes. For instance, SA can undergo glycosylation, methylation, amino acid conjugation, sulfation and hydroxylation (Dempsey et al. 2011). Glycosylation of SA has been shown to prevent the accumulation of toxic levels of free SA (Sorefan et al. 2003). Sulfation is another intriguing conjugation reaction. During the past two decades, several sulfated plant hormones have been discovered. In general, it is assumed that sulfate conjugation improves physiological processes such as the response to stresses, growth and development by affecting biological activities of certain compounds. In mammals, for example, sulfated compounds were found to have increased mobility since they become more water-soluble, but sulfation has also been postulated to be a detoxification mechanism (Kester et al. 1999; Visser et al. 1998; Coughtrie 1996). In plants, glucosinolate and flavonoid sulfation are best described. Sulfated flavonoids were shown to increase polar auxin transport by serving as antagonists of quercetin, which binds to the receptor that inhibits auxin polar transport (Ananvoranich et al. 1994; Varin et al. 1989). Recently, a rice sulfotransferase, OsSOT9, was shown to be induced in response to several stresses such as heat, drought and cold (Cao et al. 2016). A study on *Arabidopsis* SOTs also showed that the expression of *AtSOT12* was induced by salt stress and *sot12* mutants were hypersensitive to salt and pathogens and had increased levels of ABA (Baek et al. 2010). Though it has been considered for long that the role of hormone modification is inactivation/detoxification and storage, these recent findings show it may also serve as a crucial step in the activation of the biological role of compounds. So far, there are no indications that conjugation of strigolactones occurs.

Can we use fundamental knowledge to control Striga?

Controlling *Striga* using several approaches ranging from cultural to chemical and genetic options has been tried. Intercropping, using trap crops, crop rotation, hand weeding, using biological control agents such as *F. oxysporum*, application of herbicides and fuming the land without the presence of the host plant using chemicals such as ethylene to trigger germination were used as techniques to reduce the impact of the weed (Nzioki et al. 2016; Parker 2009).

In many cases, these limited control options are overcome by the weed (Ejeta 2007a). Furthermore, none of these methods were proven to be a single solution that is both economically and practically feasible.

Control is complicated by the life cycle of *Striga* and the fact that it produces 10,000-100,000 seeds per pant, which can be viable for up to 20 years (Yoder et al. 2010). *Striga* seeds stay dormant in the soil until they perceive a signal that assures the presence of a suitable host. This is an important factor for the success of the parasitic plant to complete its life cycle. The failure to find a host plant in the close vicinity within few days after germination means death of the seedlings (Runo et al. 2018). After germination, *Striga* completes most of its life cycle underground. While it is underground, the endosperm can sustain the first 3-7 days without attachment to a host (A. Dawud 2017), but *Striga* needs to attach in that period or it will die. Upon encountering a host root the *Striga* radicle forms a haustorium, an organ that penetrates the host root cortex and endodermis within 6 to 72 hours (Rich et al. 2007).

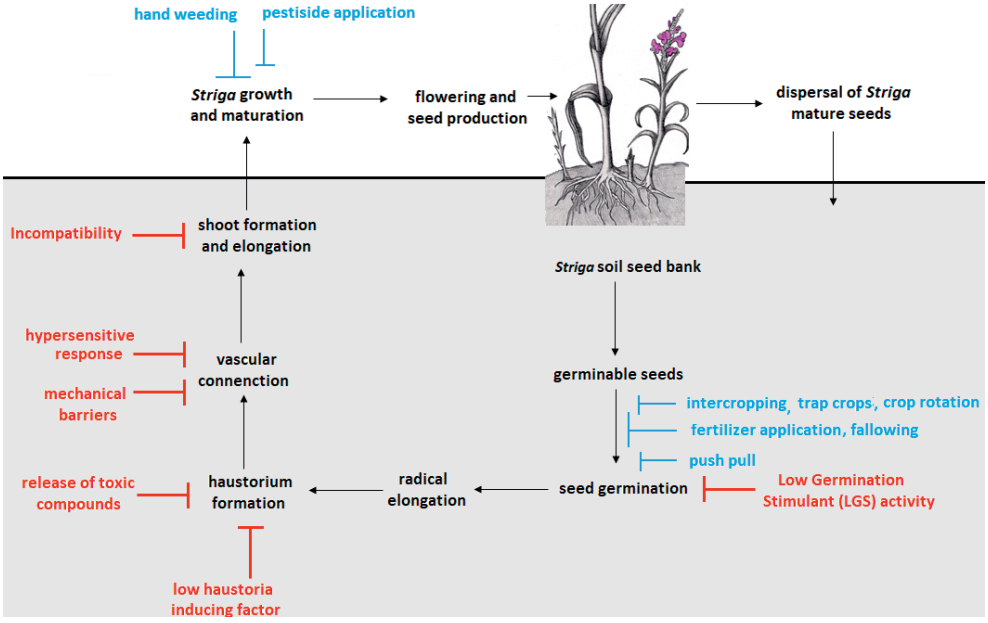


Figure 4. Schematic representation of *Striga* life cycle (black), examples of cultural practices used by farmers (blue) and identified resistance mechanisms by host plants (red) to reduce or eradicate *Striga* infection.

Then, it establishes a xylem-xylem connection and gets access to the nutrients from the host plant. After another 4-7 weeks, the shoots will emerge from the ground. During the underground period, it already causes most of the damage to the host plant. Though early symptoms such as wilting, curling of the leaves and stunting can be used to detect plants that are infected by *Striga*, these symptoms can be confused with drought symptoms that make it difficult for the farmers to decide they have a *Striga* infection. Finally, to complete its life cycle, shoots develop and flower and set and shed seeds thus increasing the number of seeds in the soil seed bank (Figure 4).

Breeding for resistant cultivars has for the past decades been the most important strategy to reduce the *Striga* problem. For the success of this approach, understanding the mechanisms behind resistance is important. Both from wild and domesticated genotypes of sorghum, maize and rice, resistant genotypes have been identified with both pre-and/or post-attachment mechanisms to cope with the presence of the weed or to avoid being infected by the weed. So far, several resistance/tolerance mechanisms have been identified. For instance, a hypersensitive response and incompatibility where the host plants commits local suicide by killing its own cells at the point of attachment are examples of post-attachment resistance mechanisms (Ejeta 2007a; Mohamed et al. 2003). Other post-attachment resistance mechanisms such as cell-wall thickening around the attachment area or avoiding establishment of a xylem-xylem connection with the parasite have been identified in sorghum landrace N13 and rice cultivar Nipponbare (Cissoko et al. 2011; Maiti et al. 1984). Interestingly, some sorghum cultivars such as SRN39, Tetron and 555 can stop or dramatically reduce the germination of *Striga* seeds as a pre-attachment resistance mechanism (Satish et al. 2012). These genotypes exhibit a low level of germination stimulant activity (Gobena et al. 2017).

Several breeding methods have been employed to develop *Striga* resistance. In 2007, low germination stimulant activity (LGS) was first shown to exist in sorghum using an agar assay that determined the distance between the sorghum roots and the germinating *Striga* seeds to select resistant lines (Ejeta 2007a). Later, using 354 recombinant inbred lines derived from SRN39 and Shanqui-Red - a low and high *Striga* seed germination stimulant sorghum line, respectively - the LGS locus was mapped on chromosome 5. The region was fine-mapped using the sorghum genome sequence and comparative analysis of the rice genome to 400 kb encompassing about 30 genes (Satish et al. 2012).

Identification of the sorghum LGS gene that could confer *Striga* resistance can be used as a key tool to fight *Striga* infestation. In order to understand the mechanism underlying LGS, identifying the gene that regulates the change in germination stimuli will be the first step. Then, using genome editing techniques, the gene itself can be used to manipulate sorghum to produce a desired combination of stimuli to improve its performance against *Striga*. In addition, these findings can possibly also be translated to be used in other crops to create both low and high germination stimulant producing genotypes. For example, we can design non-host plants that produce certain strigolactone combinations that trigger high *Striga* germination, which would result in suicide germination (*Striga* germinates but cannot attach). Furthermore, understanding the mechanism behind the regulation of strigolactones by the LGS gene could lead to other tools that can be applied as *Striga* control measure without editing the genome of the host plant. One possibility is exogenous application of inhibitors of certain parts of the strigolactone biosynthetic pathway.

However, any manipulation of these hormones is very challenging due to their multiple roles as plant hormone and as semiochemicals. In addition, they are a group of molecules that are produced and exuded to the rhizosphere in different composition, both in amount and type. A change in its composition could also compromise its role in one of the biological processes

controlled by strigolactones. The general aim of this project was to unravel the puzzle of strigolactone biosynthesis and its diversification, which may assist the effort of controlling *Striga* infection. In addition, understanding the mechanism behind *Striga* resistance and identification of the genes underlying this resistance could be used for breeding programs for multiple crops that are threatened by this weed. In my thesis I focus on identifying and characterizing the gene(s) responsible for germination stimulant activity, read strigolactone biosynthesis, in sorghum.

Scope of the thesis

In **Chapter 1**, I review the literature and describe the severity of the losses caused by *Striga* infestation in several crops. I discuss the fact that the weed is dependent on strigolactones to find its host which can possibly be used as a lead to design strategies to fight the *Striga* problem.

In **Chapter 2,** I study the correlation between strigolactone profile and *Striga* seed germination activity using different sorghum lines. Genome sequencing of several sorghum lines was then used as a validation set to identify and confirm the candidate gene that regulates the low germination stimulant activity in sorghum. In this work, I show that low germination stimulant activity in sorghum is regulated by a gene annotated as a sulfotransferase. Furthermore. I show its effect on strigolactone biosynthesis in a stereospecific manner.

In **Chapter 3,** I further study this sulfotransferase to understand the mechanism by which it controls (stereo-specificity of) strigolactone biosynthesis in sorghum. I use protein modeling and substrate docking as an approach to predict the activity of the sulfotransferase. I develop a model that shows how sulfation may change the strigolactone profile of sorghum, hence affecting its susceptibility towards *Striga*.

Identifying the gene regulating sorgomol production in sorghum is my focus in **Chapter 4**. In this part of the study, I use mapping, biochemistry, RNAseq and genome sequencing as an approach to investigate my hypothesis that a P450 in sorghum is responsible for the conversion of 5-deoxystrigol to sorgomol.

In **Chapter 5**, I contribute to unraveling biosynthesis of strigolactones in sorghum by characterizing the four putative sorghum MAX1 homologs. I use both targeted and untargeted metabolomics to explore the substrates and products of the four sorghum MAX1 homologs using transient expression in *Nicotiana benthamiana***.**

In **Chapter 6**, I discuss the main highlights of the thesis, the challenges and future perspectives. Based on the fact that the success of *Striga* infestation is dependent on the type of strigolactones exuded by sorghum plants, I propose possible tools that can be used to eradicate *Striga*. I also address the concept of integrated *Striga* management.

References

A. Dawud, Maryam. 2017. *'Striga* Resistance in Cereal Crops: Recent Progress and Future Prospects. A Review', *Global Journal of Science Frontier Research; Vol 17, No 3-D (2017): The Global Journal of Science Frontier Research*.

Abe, Satoko, Aika Sado, Kai Tanaka, Takaya Kisugi, Kei Asami, Saeko Ota, Hyun Il Kim, et al. 2014. 'Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18084-89.

Agusti, J., S. Herold, M. Schwarz, P. Sanchez, K. Ljung, E. A. Dun, P. B. Brewer, et al. 2012. 'Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants (Proceedings of the National Academy of Sciences United States of America (2011) 108, (20242-20247) DOI: 10.1073/pnas.1111902108)', *Proceedings of the National Academy of Sciences of the United States of America*, 109.

Akiyama, K., S. Ogasawara, S. Ito, and H. Hayashi. 2010. 'Structural requirements of strigolactones for hyphal branching in AM fungi', *Plant Cell Physiol*, 51: 1104-17.

Al-Babili, Salim, and Harro J. Bouwmeester. 2015. 'Strigolactones, a Novel Carotenoid-Derived Plant Hormone', *Annual Review of Plant Biology*, 66: 161-86.

Alder, Adrian, Muhammad Jamil, Mattia Marzorati, Mark Bruno, Martina Vermathen, Peter Bigler, Sandro Ghisla, et al. 2012. *The Path from -Carotene to Carlactone, a Strigolactone-Like Plant Hormone*.

Ananvoranich, S., L. Varin, P. Gulick, and R. Ibrahim. 1994. 'Cloning and regulation of flavonol 3-sulfotransferase in cell-suspension cultures of *Flaveria bidentis*', *Plant Physiol*, 106: 485-91.

Awad, Ayman, Daisuke Sato, Dai Kusumoto, Hiroaki Kamioka, Yasutomo Takeuchi, and Koichi Yoneyama. 2006. 'Characterization of Strigolactones, Germination Stimulants for the Root Parasitic Plants *Striga* and *Orobanche*, Produced by Maize, Millet and Sorghum', *Plant Growth Regulation*, 48: 221-27.

Baek, Dongwon, Prasanth Pathange, Jung-Sung Chung, Jiafu Jiang, Liqiong Gao, Akira Oikawa, Masami Yokota Hirai, et al. 2010. 'A stress-inducible sulphotransferase sulphonates salicylic acid and confers pathogen resistance in Arabidopsis', *Plant, Cell & Environment*, 33: 1383-92.

Berner, D. K., and O. A. Williams. 1998. 'Germination Stimulation of Striga gesnerioides Seeds by Hosts and Nonhosts', *Plant Disease*, 82: 1242-47.

Bouwmeester, H. J., Chr Roux, J. A. Lopez Raez, and G. Bécard. 2007. "Rhizosphere communication of plants, parasitic plants and AM fungi." In *Trends in Plant Science 12 (2007) 5*. 2007.

Brewer, P. B., K. Yoneyama, F. Filardo, E. Meyers, A. Scaffidi, T. Frickey, K. Akiyama, et al. 2016. 'LATERAL BRANCHING OXIDOREDUCTASE acts in the final stages of strigolactone biosynthesis in Arabidopsis', *Proc Natl Acad Sci U S A*, 113: 6301-6.

Bruno, Mark, Manuel Hofmann, Martina Vermathen, Adrian Alder, Peter Beyer, and Salim Al-Babili. 2014. 'On the substrate-and stereospecificity of the plant carotenoid cleavage dioxygenase 7', *FEBS Letters*, 588: 1802-07.

Bücking, Heike, Liepold, and Ambilwade. 2012. 'The Role of the Mycorrhizal Symbiosis in Nutrient Uptake of Plants and the Regulatory Mechanisms Underlying These Transport Processes.' *Plant Sci*, 4: 108-32.

Cao, Xufeng, Yongrong Liao, Songhao Rong, Changqiong Hu, Xuewei Zhang, Rongjun Chen, Zhengjun Xu, et al. 2016. 'Identification and characterization of a novel abiotic stress responsive sulphotransferase gene (OsSOT9) from rice', *Biotechnology & Biotechnological Equipment*, 30: 227-35.

Cardoso, C., Y. Zhang, M. Jamil, J. Hepworth, T. Charnikhova, S. O. Dimkpa, C. Meharg, et al. 2014. 'Natural variation of rice strigolactone biosynthesis is associated with the deletion of two MAX1 orthologs', *Proc Natl Acad Sci U S A*, 111: 2379-84.

Charnikhova, T. V., K. Gaus, A. Lumbroso, M. Sanders, J. P. Vincken, A. De Mesmaeker, C. P. Ruyter-Spira, et al. 2017. 'Zealactones. Novel natural strigolactones from maize', *Phytochemistry*, 137: 123-31.

Cissoko, M., A. Boisnard, J. Rodenburg, M. C. Press, and J. D. Scholes. 2011. 'New Rice for Africa (NERICA) cultivars exhibit different levels of post-attachment resistance against the parasitic weeds Striga hermonthica and Striga asiatica', *New Phytol*, 192: 952-63.

Coughtrie, M. W. 1996. 'Sulphation catalysed by the human cytosolic sulphotransferases--chemical defence or molecular terrorism?', *Hum Exp Toxicol*, 15: 547-55.

de Saint Germain, Alexandre, Yasmine Ligerot, Elizabeth A. Dun, Jean-Paul Pillot, John J. Ross, Christine A. Beveridge, and Catherine Rameau. 2013. 'Strigolactones Stimulate Internode Elongation Independently of Gibberellins', *Plant Physiology*, 163: 1012.

Dempsey, D'Maris Amick, A. Corina Vlot, Mary C. Wildermuth, and Daniel F. Klessig. 2011. 'Salicylic Acid biosynthesis and metabolism', *The arabidopsis book*, 9: e0156-e56.

Ejeta, Gebisa. 2007a. 'Breeding for Striga Resistance in Sorghum: Exploitation of an Intricate Host-Parasite Biology', *Crop Science*, 47: S216-S27.

Ejeta, Gebisa. 2007b. 'The Striga scourge in Africa: a growing pandemic.' in, *Integrating new technologies for Striga control: towards ending the witch-hunt* (World Scientific).

FAOSTAT. 2012. 'UN. Food and Agriculture Organization (FAO). Trade data base, production data base'.

Gobena, D., M. Shimels, P. J. Rich, C. Ruyter-Spira, H. Bouwmeester, S. Kanuganti, T. Mengiste, et al. 2017. 'Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes Striga resistance', *Proc Natl Acad Sci U S A*, 114: 4471-76.

Gomez-Roldan, M. V., S. Fermas, P. B. Brewer, V. Puech-Pages, E. A. Dun, J. P. Pillot, F. Letisse, et al. 2008. "Strigolactone inhibition of shoot branching." In *Nature 455 (2008)*. 2008.

Haussmann, Bettina I. G., Dale E. Hess, H. Günter Welz, and Hartwig H. Geiger. 2000. 'Improved methodologies for breeding striga-resistant sorghums', *Field Crops Research*, 66: 195-211.

Hayward, Alice, Petra Stirnberg, Christine Beveridge, and Ottoline Leyser. 2009. 'Interactions between Auxin and Strigolactone in Shoot Branching Control', *Plant Physiology*, 151: 400.

Jamil, M., T. Charnikhova, C. Cardoso, T. Jamil, K. Ueno, F. Verstappen, T. Asami, et al. 2011. 'Quantification of the relationship between strigolactones and Striga hermonthica infection in rice under varying levels of nitrogen and phosphorus', *Weed Research*, 51: 373-85.

Kester, M. H., E. Kaptein, T. J. Roest, C. H. van Dijk, D. Tibboel, W. Meinl, H. Glatt, et al. 1999. 'Characterization of human iodothyronine sulfotransferases', *J Clin Endocrinol Metab*, 84: 1357-64.

Kim, Hyun Il, Xiaonan Xie, Han Sung Kim, Jae Chul Chun, Kaori Yoneyama, Takahito Nomura, Yasutomo Takeuchi, et al. 2010. 'Structure–activity relationship of naturally occurring strigolactones in *Orobanche minor* seed germination stimulation', *Journal of Pesticide Science*, 35: 344-47.

Kohlen, W., T. Charnikhova, L. Qing, R. M. E. H. Bours, M. Domagalska, S. Beguerie, F. W. A. Verstappen, et al. 2011. "Strigolactones are transported through the xylem and play a key role in shoot architectural response to phosphate deficiency in nonarbuscular mycorrhizal host Arabidopsis." In *Plant Physiology 155 (2011) 2*. 2011.

Lopez-Obando, Mauricio, Yasmine Ligerot, Sandrine Bonhomme, François-Didier Boyer, and Catherine Rameau. 2015. 'Strigolactone biosynthesis and signaling in plant development', *Development*, 142: 3615.

Maiti, R. K., K. V. Ramaiah, S. S. Bisen, and V. L. Chidley. 1984. 'A Comparative Study of the Haustorial Development of *Striga asiatica* (L.) Kuntze on Sorghum Cultivars', *Annals of botany*, 54: 447-57.

Mohamed, A., A. Ellicott, T. L. Housley, and G. Ejeta. 2003. 'Hypersensitive Response to *Striga* Infection in Sorghum', *Crop Science*, 43: 1320-24.

Mohamed, Kamal I., Papes Monica, Richard Williams, W. Benz Brett, and A. Townsend Peterson. 2006. 'Global Invasive Potential of 10 Parasitic Witchweeds and Related Orobanchaceae', *Ambio*, 35: 281-88.

Mohemed, N., T. Charnikhova, E. J. Bakker, A. van Ast, A. G. Babiker, and H. J. Bouwmeester. 2016. 'Evaluation of field resistance to *Striga hermonthica* (Del.) Benth. in Sorghum bicolor (L.) Moench. The relationship with strigolactones', *Pest Manag Sci*, 72: 2082-90.

Mohemed, N., T. Charnikhova, E. F. Fradin, J. Rienstra, A. G. T. Babiker, and H. J. Bouwmeester. 2018. 'Genetic variation in Sorghum bicolor strigolactones and their role in resistance against Striga hermonthica', *J Exp Bot*, 69: 2415-30.

Motonami, Noriko, Kotomi Ueno, Hitomi Nakashima, Saki Nomura, Masaharu Mizutani, Hirosato Takikawa, and Yukihiro Sugimoto. 2013. 'The bioconversion of 5-deoxystrigol to sorgomol by the sorghum, Sorghum bicolor (L.) Moench', *Phytochemistry*, 93: 41-48.

Nomura, S., H. Nakashima, M. Mizutani, H. Takikawa, and Y. Sugimoto. 2013. 'Structural requirements of strigolactones for germination induction and inhibition of Striga gesnerioides seeds', *Plant Cell Rep*, 32: 829-38.

Nomura, Takahito, Xiaonan Xie, Koichi Yoneyama, Kaori Yoneyama, Takaya Kisugi, Kohki Akiyama, Yoshifumi Nakatani, et al. 2018. 'Which are the major players, canonical or non-canonical strigolactones?', *Journal of Experimental Botany*, 69: 2231-39.

Nouri, Eva, Florence Breuillin-Sessoms, Urs Feller, and Didier Reinhardt. 2014. 'Phosphorus and Nitrogen Regulate Arbuscular Mycorrhizal Symbiosis in Petunia hybrida', *PLOS ONE*, 9: e90841.

Nzioki, H. S., F. Oyosi, C. E. Morris, E. Kaya, A. L. Pilgeram, C. S. Baker, and D. C. Sands. 2016. 'Striga Biocontrol on a Toothpick: A Readily Deployable and Inexpensive Method for Smallholder Farmers', *Front Plant Sci*, 7: 1121.

Parker, Chris, and Charlie R. Riches. 1993. *Parasitic weeds of the world: biology and control* (CAB international).

Parker, C. 2009. 'Observations on the current status of Orobanche and Striga problems worldwide', *Pest Manag Sci*, 65: 453-9.

Piotrowska, A., and A. Bajguz. 2011. 'Conjugates of abscisic acid, brassinosteroids, ethylene, gibberellins, and jasmonates', *Phytochemistry*, 72: 2097-112.

Rich, Patrick J., and Gebisa Ejeta. 2007. 'BIOLOGY OF HOST-PARASITE INTERACTIONS IN *STRIGA* SPECIES.' in, *Integrating New Technologies for Striga Control* (WORLD SCIENTIFIC).

Rodenburg, Jonne, Mamadou Cissoko, Nicholas Kayongo, Ibnou Dieng, Jenipher Bisikwa, Runyambo Irakiza, Isaac Masoka, et al. 2017. 'Genetic variation and host-parasite specificity of *Striga* resistance and tolerance in rice: the need for predictive breeding', *New Phytol*, 214: 1267-80.

Rodenburg, Jonne, Matty Demont, Sander J. Zwart, and Lammert Bastiaans. 2016. 'Parasitic weed incidence and related economic losses in rice in Africa', *Agriculture, Ecosystems & Environment*, 235: 306-17.

Runo, Steven, and Eric K. Kuria. 2018. 'Habits of a highly successful cereal killer, *Striga*', *PLoS pathogens*, 14: e1006731-e31.

Ruyter-Spira, Carolien, Wouter Kohlen, Tatsiana Charnikhova, Arjan van Zeijl, Laura van Bezouwen, Norbert de Ruijter, Catarina Cardoso, et al. 2011. 'Physiological Effects of the Synthetic Strigolactone Analog GR24 on Root System Architecture in Arabidopsis: Another Belowground Role for Strigolactones?', *Plant Physiology*, 155: 721.

Satish, K., Z. Gutema, C. Grenier, P. J. Rich, and G. Ejeta. 2012. 'Molecular tagging and validation of microsatellite markers linked to the low germination stimulant gene (lgs) for Striga resistance in sorghum [Sorghum bicolor (L.) Moench]', *Theor Appl Genet*, 124: 989-1003.

Schwartz, S. H., X. Qin, and M. C. Loewen. 2004. 'The biochemical characterization of two carotenoid cleavage enzymes from Arabidopsis indicates that a carotenoid-derived compound inhibits lateral branching', *J Biol Chem*, 279: 46940-5.

Seto, Yoshiya, Aika Sado, Kei Asami, Atsushi Hanada, Mikihisa Umehara, Kohki Akiyama, and Shinjiro Yamaguchi. 2014. 'Carlactone is an endogenous biosynthetic precursor for strigolactones', *Proceedings of the National Academy of Sciences*, 111: 1640.

Sibhatu, Berhane. 2016. 'Review on Striga Weed Management', *Int. J. Life. Sci. Scienti. Res*, 2: 110-20.

Sorefan, K., J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, E. Foo, S. Chatfield, et al. 2003. 'MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea', *Genes Dev*, 17: 1469-74.

Spallek, Thomas, Musembi Mutuku, and Ken Shirasu. 2013. 'The genus Striga: a witch profile', *Molecular Plant Pathology*, 14: 861-69.

Spíchal, Lukáš, Natalia Yu Rakova, Michael Riefler, Takeshi Mizuno, Georgy A. Romanov, Miroslav Strnad, and Thomas Schmülling. 2004. 'Two Cytokinin Receptors of Arabidopsis thaliana, CRE1/AHK4 and AHK3, Differ in their Ligand Specificity in a Bacterial Assay', *Plant and Cell Physiology*, 45: 1299-305.

Strullu-Derrien, Christine, Marc-André Selosse, Paul Kenrick, and Francis M. Martin. 2018. 'The origin and evolution of mycorrhizal symbioses: from palaeomycology to phylogenomics', *New Phytologist*, 220: 1012-30.

Sun, Huwei, Jinyuan Tao, Pengyuan Gu, Guohua Xu, and Yali Zhang. 2016. 'The role of strigolactones in root development', *Plant Signaling & Behavior*, 11.

Ueda, H., and M. Kusaba. 2015. 'Strigolactone Regulates Leaf Senescence in Concert with Ethylene in Arabidopsis', *Plant Physiol*, 169: 138-47.

Ueno, K., S. Nomura, S. Muranaka, M. Mizutani, H. Takikawa, and Y. Sugimoto. 2011. 'Ent-2'-epi-Orobanchol and its acetate, as germination stimulants for Striga gesnerioides seeds isolated from cowpea and red clover', *J Agric Food Chem*, 59: 10485-90.

Umehara, M., A. Hanada, S. Yoshida, K. Akiyama, T. Arite, N. Takeda-Kamiya, H. Magome, et al. 2008. 'Inhibition of shoot branching by new terpenoid plant hormones', *Nature*, 455: 195-200.

Umehara, Mikihisa, Atsushi Hanada, Hiroshi Magome, Noriko Takeda-Kamiya, and Shinjiro Yamaguchi. 2010. 'Contribution of Strigolactones to the Inhibition of Tiller Bud Outgrowth under Phosphate Deficiency in Rice', *Plant and Cell Physiology*, 51: 1118-26.

Varbanova, M., S. Yamaguchi, Y. Yang, K. McKelvey, A. Hanada, R. Borochov, F. Yu, et al. 2007. 'Methylation of gibberellins by Arabidopsis GAMT1 and GAMT2', *Plant Cell*, 19: 32-45.

Varin, Luc, and Ragai K. Ibrahim. 1989. 'Partial Purification and Characterization of Three Flavonol-Specific Sulfotransferases from Flaveria chloraefolia', *Plant Physiology*, 90: 977.

Visser, T. J., E. Kaptein, H. Glatt, I. Bartsch, M. Hagen, and M. W. Coughtrie. 1998. 'Characterization of thyroid hormone sulfotransferases', *Chem Biol Interact*, 109: 279-91.

Vurro, Maurizio, Angela Boari, Benjamin Thiombiano, and Harro Bouwmeester. 2019. 'Strigolactones and Parasitic Plants.' in.

Wang, Yanting, and Harro J. Bouwmeester. 2018. 'Structural diversity in the strigolactones', *Journal of Experimental Botany*, 69: 2219-30.

Wu, Yuanli, Evgenia Dor, and Joseph Hershenhorn. 2017. 'Strigolactones affect tomato hormone profile and somatic embryogenesis', *Planta*, 245: 583-94.

Xie, Xiaonan. 2017. 'Structural diversity of strigolactones and their distribution in the plant kingdom', *Japanese Journal of Pesticide Science*, 42: 10-16.

Xie, Xiaonan, Kaori Yoneyama, and Koichi Yoneyama. 2010. 'The Strigolactone Story', *Annual Review of Phytopathology*, 48: 93-117.

Yoder, John I., and Julie D. Scholes. 2010. 'Host plant resistance to parasitic weeds; recent progress and bottlenecks', *Curr Opin Plant Biol*, 13: 478-84.

Yoneyama, K., K. Yoneyama, Y. Takeuchi, and H. Sekimoto. 2007. 'Phosphorus deficiency in red clover promotes exudation of orobanchol, the signal for mycorrhizal symbionts and germination stimulant for root parasites', *Planta*, 225: 1031-8.

Yoneyama, Kaori, Ryota Arakawa, Keiko Ishimoto, Hyun Il Kim, Takaya Kisugi, Xiaonan Xie, Takahito Nomura, et al. 2015. 'Difference in Striga-susceptibility is reflected in strigolactone secretion profile, but not in compatibility and host preference in arbuscular mycorrhizal symbiosis in two maize cultivars', *New Phytologist*, 206: 983-89.

Yoneyama, Kaori, Xiaonan Xie, Takaya Kisugi, Takahito Nomura, Hitoshi Sekimoto, Takao Yokota, and Koichi Yoneyama. 2011. 'Characterization of strigolactones exuded by Asteraceae plants', *Plant Growth Regulation*, 65: 495-504.

Yoneyama, Kaori, Xiaonan Xie, Hitoshi Sekimoto, Yasutomo Takeuchi, Shin Ogasawara, Kohki Akiyama, Hideo Hayashi, et al. 2008. 'Strigolactones, host recognition signals for root parasitic plants and arbuscular mycorrhizal fungi, from Fabaceae plants', *New Phytologist*, 179: 484-94.

Zhan, Y., Y. Qu, L. Zhu, C. Shen, X. Feng, and C. Yu. 2018. 'Transcriptome analysis of tomato (Solanum lycopersicum L.) shoots reveals a crosstalk between auxin and strigolactone', *PLOS ONE*, 13: e0201124.

Zhang, Y., A. D. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, et al. 2014. 'Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis', *Nat Chem Biol*, 10: 1028-33.

Chapter 2

Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes *Striga* resistance

Daniel Gobena[1], Mahdere Z. Shimels[2], Patrick J. Rich[1], Carolien Ruyter-Spira[2], Harro Bouwmeester[2,3], Satish Kanuganti[1], Tesfaye Mengiste[4], and Gebisa Ejeta[1]

[1]Department of Agronomy, Purdue University, West Lafayette, IN 47907
[2]Laboratory of Plant Physiology, Wageningen University, 6708 PB Wageningen, The Netherlands
[4]Present address: Plant Hormone Biology, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands.
[4]Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907

Published in PNAS (2017) 114: 4471-7

Abstract

Striga is a major biotic constraint to sorghum production in semiarid tropical Africa and Asia. Genetic resistance to this parasitic weed is the most economically feasible control measure. Mutant alleles at the *LGS1* (*LOW GERMINATION STIMULANT 1*) locus drastically reduce *Striga* germination stimulant activity. We provide evidence that the responsible gene at *LGS1* codes for an enzyme annotated as a sulfotransferase and show that functional loss of this gene results in a change of the dominant strigolactone in root exudates from 5-deoxystrigol, a highly active *Striga* germination stimulant, to orobanchol, a strigolactone with opposite stereochemistry. Orobanchol, although not previously reported in sorghum, functions in the multiple strigolactone roles required for normal growth and environmental responsiveness but does not stimulate germination of *Striga*. This work describes the identification of a gene regulating *Striga* resistance and the underlying protective chemistry resulting from mutation.

Introduction

Infestation by the parasitic weed *Striga* (*Striga asiatica* and *Striga hermonthica*) is a serious constraint to the production of sorghum (*Sorghum bicolor*), a staple cereal crop grown widely across sub-Saharan Africa and the Indian subcontinent. Global estimates of *Striga*'s human toll are lacking. An earlier report by the Food and Agriculture Organization of the United Nations focused on West Africa estimated that the livelihoods of 300 million people were negatively affected by the pest (Mboob 1989). Conservative extrapolation from a recent report on losses to *Striga* in rice (Rodenburg et al. 2016) puts the economic impact on cereal production in sub-Saharan Africa at $1.2 billion annually with losses increasing by $177 million per year. Most of these losses are borne by subsistence farmers (Ejeta 2007a). Genetic resistance to this pest through low *Striga* germination stimulant activity provides control and permits economic production of this crop (Pérez-Vich et al. 2013). Because it is an obligate root parasite, *Striga* seed will not germinate unless it receives a chemical cue from a potential host plant (Rich et al. 2007). Among chemicals identified in sorghum root exudates with *Striga* germination stimulant activity, the most potent are the strigolactones, a class of related compounds used by most terrestrial plants as hormones to regulate shoot and root branching (Gomez-Roldan et al. 2008; Rasmussen et al. 2012). Their presence in root exudates is critical to symbiotic colonization by arbuscular mycorrhizal (AM) fungi (Akiyama et al. 2005). Associations with AM fungi greatly improve the performance of sorghum under nutrient and water deficits (Sun et al. 2013). *Striga* seems to have taken advantage of this signaling to detect its proximity to sorghum roots, germinating at the proper time and place to increase its chances of completing its life cycle on this preferred host. Sorghum produces several strigolactones and exudes them from its roots, particularly under conditions of limited phosphate and nitrogen, probably in an attempt to promote mycorrhizal association (Jamil et al. 2013). Among the strigolactones reported to be present in sorghum root exudates are sorgolactone, strigol, 5-deoxystrigol and sorgomol (Figure 1) (Jamil et al. 2013; Siame et al. 1993; Motonami et al. 2013). These compounds differ from each other by various substitutions on the A- and B-rings but share a common stereochemistry with respect to the β-orientation of their C-rings (Xie et al. 2013). *Striga* is quite sensitive to these strigolactones, able to germinate at concentrations as low as 10^{-11}M depending on the particular strigolactone (Cook et al. 1966).

To facilitate the identification and characterization of resistance to *Striga*, our laboratory developed bioassays that allow observations of the parasitic association at its earliest stages, normally hidden below ground. Among these is the agar gel assay wherein the *Striga* germination stimulant activity of sorghum accessions can be quantified based on the distance between the sorghum root and germinating *Striga* seed in agar (Hess et al. 1992). This useful assay has resulted in the development and release of several *Striga*-resistant sorghum varieties with low germination stimulant activity (Ejeta 2007a). Although not all sorghum lines showing field resistance to *Striga* had low *Striga* germination stimulant activity, all low-stimulant sorghums that were field-tested showed *Striga* resistance (Ejeta 2007a). Low *Striga* germination stimulant activity has been an important resistance trait in sorghum

improvement but less so in other crop hosts of *Striga* such as maize, millet and rice (Pérez-Vich et al. 2013). Genetic studies have shown that inheritance of low *Striga* germination stimulant activity in sorghum is through a mutant allele (*lgs*) expressed in homozygous recessive individuals (K. Vogler et al. 1996). The *Striga*-resistant sorghum variety SRN39 carrying this mutation was mated with a Chinese landrace Shanqui Red, with high germination stimulant activity, to generate a genetic mapping population of 600 recombinant inbred lines (RILs). In a previous genotypic and phenotypic evaluation of 328 RILs by the agar gel assay we created a genetic map with 428 markers, placing the *LGS1* (*LOW GERMINATION STIMULANT 1*) locus in a region near the tip of chromosome 5 with fine mapping that delimited it to a 30-gene region (Satish et al. 2012).

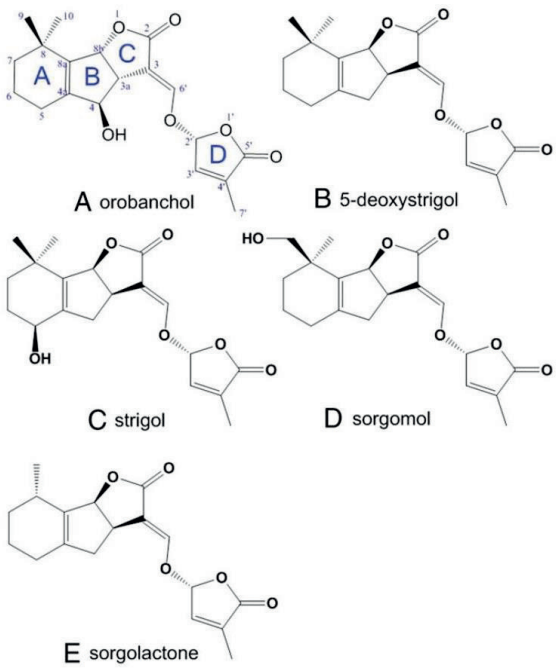


Figure 1. Chemical structures of strigolactones found in sorghum root exudates. Orobanchol (A) has not been previously reported in sorghum. Note the enantiomeric orientation of its C-ring (α-orientation) with respect to the other strigolactones (β-orientation), 5-deoxystrigol (B), strigol (C), sorgomol (D) and sorgolactone (E), previously reported in sorghum root exudates.

Much has been learned over the past decade about biosynthesis of strigolactones, particularly since their roles as growth regulators were discovered. The strigolactones are derived from β-carotene through a series of isomerization, cleavage, oxidation and cyclization steps to form the four distinctive rings of the strigolactones (Alder et al. 2012; Zhang et al. 2014). Four enzymes have been identified to be involved in these steps: DWARF27 (D27), a carotenoid isomerase that converts all-*trans*-β-carotene to 9-*cis*-β-carotene that can be cleaved by CAROTENOID CLEAVAGE DIOXYGENASE 7 (CCD7) to form 9-*cis*-β-apo-10'-carotenal, which is

converted by CAROTENOID CLEAVAGE DIOXYGENASE 8 (CCD8) to carlactone, which contains the A- and D-rings and is, in rice, subsequently oxidized by an ortholog of the Arabidopsis MORE AXILLARY BRANCHES 1 (MAX1), to the first canonical rice strigolactone, 4-deoxyorobanchol (*ent*-2'-*epi*-5-deoxystrigol) (Zhang et al. 2014). Mutant alleles at these loci were identified by plant growth phenotypes that affected shoot branching in model species. Less is known about the later steps of strigolactone biosynthesis, particularly how the additions and/or modifications to functional groups on the member rings occurs. It has been assumed that 5-deoxystrigol is the proto-strigolactone for the strigol-type strigolactones, having a β-oriented C-ring, and 4-deoxyorobanchol for the orobanchol-type strigolactones, with the C-ring in α-orientation (Zhang et al. 2014; Al-Babili et al. 2015). Both groups have the D-ring in R configuration around the chiral center at C-2' (Alder et al. 2012). A major strigolactone in rice root exudates is orobanchol (Figure 1A), and all other strigolactones present in this species share the same stereochemistry with respect to the spatial orientation of the C-ring (Cardoso et al. 2014; Xie et al. 2013). Other plants, including tobacco (*Nicotiana tabaccum*), exude both types of strigolactones (Xie et al. 2013). For sorghum and many other plant species from which strigolactones have been described, the stereochemistries of their strigolactones have not always been determined. However, all strigolactones reported in the root exudates of sorghum (Awad et al. 2006), including 5-deoxystrigol, strigol, sorgomol and sorgolactone, are of the strigol type (Figure 1).

Because mutation at *LGS1* causes a change in *Striga* germination stimulant activity, but without obvious changes to sorghum shoot architecture, we made quantitative and qualitative comparisons of strigolactones in the root exudates of mutant and WT lines.

Results

Striga Germination Stimulation and strigolactones.

Diverse lines were classified for *Striga* resistance, based on the germination distance of *Striga* embedded in agar from the sorghum root as having high maximum germination distance (MGD ≥10 mm; Shanqui-Red, Figure 2A) or low (MGD <10 mm; SRN39, 555, IS7777, SC103 and Tetron) *Striga* germination stimulant activity with four *Striga* sources (Table 1). Low-stimulant genotype SRN39 (Figure 2B), when crossed with high-germination stimulant lines, always result in F1 hybrids with high *Striga* germination stimulant activity, affirming the recessive nature of the *lgs1* mutation. Complementation tests between SRN39 and all of the low germination stimulant lines in this study indicate that they all carry mutations at a common locus because no complementation occurs in their hybrids, that is, all such hybrid plants produce low *Striga* germination stimulant responses (Table 1). The difference in resistance between low- and high- germination stimulant varieties is also apparent when lines are cultivated under *Striga* infestation. SRN39 and its derivatives determined in the agar assay to have low *Striga* germination stimulant activity also support fewer parasites in field plots (Figure 2C).

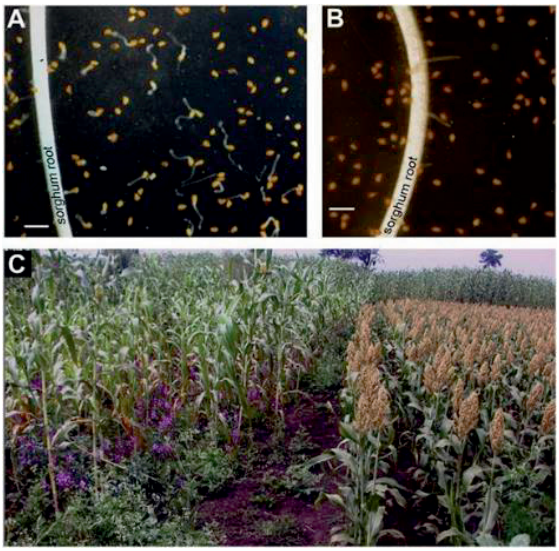


Figure 2. *Striga* resistance phenotypes of *LGS1* variants. Sorghum seedlings with high *Striga* germination stimulant activity (A) will germinate conditioned *S. asiatica* seeds cocultured in agar, a centimetre or more from its root as the germination stimulant, 5-deoxystrigol, diffuses through the medium. Low-germination stimulant sorghum that exudes orobanchol instead of 5-deoxystrigol will not cause *S. asiatica* seeds to germinate in the agar gel assay, even very near its roots (B). (Scale bars, 1 mm.) The photograph (C) shows an *LGS1* WT high- germination stimulant sorghum (left) growing next to a line (right) carrying the *lgs1-1* allele in a field infested *with S. hermonthica* (purple flowers) in Ethiopia.

Table 1. Measures of the *Striga* germination stimulant activity of sorghum lines and hybrids used for genetic mapping of *LGS1*

Sorghum line or hybrid	*S. asiatica* (Derashe, Ethiopia)		*S. asiatica* (North Carolina)		*S. hermonthica* (Samanko, Mali)		*S. hermonthica* (Sinnar, Sudan)	
Shanqui-Red	15.5	± 5.0	19.8	± 3.0	19.6	± 5.2	21.2	± 3.6
SRN39	0	± 0	1.9	± 1.8	0.1	± 0.2	0.1	± 0.2
555	0	± 0	0.5	± 0.5	1.3	± 2.2	0	± 0
IS7777	0	± 0	0	± 0	3.8	± 1.7	2.1	± 2.1
Tetron	1.2	± 2.7	4.3	± 1.5	1.4	± 1.9	2.5	± 3.9
SC103							3.4	± 2.7
(SRN39 × Shanqui-Red)F_1							10.9	± 2.3
(SRN39 × 555)F_1							0.2	± 0.3
(SRN39 × IS7777)F_1							6.6	± 2.8
(SRN39 × Tetron)F_1							7.1	± 2.6
(SRN39 × SC103)F_1							3.1	± 2.5

MGD (millimeters) as measured in the agar gel assay (discussed in the text); MGD >10 mm indicates high *Striga* germination stimulant activity; MGD <10 mm indicates low *Striga* germination stimulant activity. Values are means of measures from three plates ± one SD.

Strigolactone profiles of root exudates from *lgs1* variants consistently display reduced 5-deoxystrigol and enhanced orobanchol levels relative to WT *LGS1* root exudates (Figure 3). Comparison of retention times and mass transitions of dominant strigolactones in sorghum root exudates with standards of known stereochemistry confirmed the β-orientation of the C-ring in 5-deoxystrigol of lines carrying *LGS1* and α-orientation in orobanchol of those with *lgs1* (Figure 4). RILs with low germination stimulant activity have inherited the low 5-deoxystrigol/high orobanchol profile, whereas those with high germination stimulant activity

always contain a threshold level of 5-deoxystrigol and do not accumulate orobanchol, confirming the identity of the gene and the link to this profile (Table S1).

Because strigolactones serve other functions contributing to crop productivity (Gomez-Roldan et al. 2008; Rasmussen et al. 2012; Akiyama et al. 2005; Sun et al. 2013), selecting for mutations that knock out strigolactone production may have undesirable outcomes such as excessive shoot branching or impairment of mycorrhization. Sorghum lines examined in this study carrying *lgs1* alleles all had similar strigolactone exudation rates, typically around 2,000 pmol per plant over the 48 hrs collection period. Although the stereochemistry of the major strigolactone in these exudates profoundly affected *Striga* germination stimulant activity, other strigolactone functions seem to be unchanged by the mutation. Adult SRN39 plants on average have the same number of basal tillers (one) as Shanqui-Red at 0.5 m spacing in a field row. The two lines also do not greatly differ in the degree to which their roots are colonized by three AM fungal species, *Rhizophagus intraradices*, *Rhizophagus clarus* and *Rhizophagus custos*, alone or in combination (Figure S1). Mutation at *LGS1* results in both quantitative and qualitative changes in strigolactone content of root exudates, effectively lowering *Striga* germination stimulant activity without negative productivity side effects.

A search for polymorphisms in PCR products between the parents of the RILs contrasting for *Striga* germination stimulant activity, Shanqui-Red and SRN39, allowed genotyping with eight new markers (Table S2) to refine the position of *LGS1* on the sorghum genetic map. Polymorphisms resulting in PCR product size differences were scored by gel electrophoresis. Most (95%) polymorphic markers in the region cosegregated with the respective trait (RILs with Shanqui-Red alleles had high *Striga* germination stimulant activity, whereas those with SRN39 alleles had low germination stimulant activity). The informative recombinants allowed us to rule out several gene candidates.

For a cluster of candidate genes from position 69,977,147 – 70,011,172 on the sorghum chromosome 5 physical map (Phytozome, *Sorghum bicolor* v3.1, DOE-JGI), a PCR product could not be obtained from SRN39, so, the five genes predicted in this region (*Sobic.005G213500* to *Sobic.005G213832*) could not be scored, except as a presence/absence polymorphism. Whole-genome sequencing of the parents revealed that this five-gene region is deleted in SRN39 (Figure 5 and Table S3). The allele carried by SRN39 is given the designation *lgs1-1*. We also sequenced whole genomes of several unrelated low *Striga* germination stimulant lines in our collection. Examining genomic sequence of this region from these natural variants determined to be allelic to SRN39, we found that the allele in 555, *lgs1-2* also has a large deletion here, but slightly shifted away from the chromosome tip, spanning the position 69,958,403 – 69,986,951, and therefore missing three predicted genes, *Sobic.005G213400*, *Sobic.005G213500* and *Sobic.005G213600*. Its deletion overlapped with that of SRN39 for two genes, *Sobic.005G213500* (*Sb05g026540*), coding for an uncharacterized protein with a functional domain similar to an iron/ascorbate oxidoreductase, and *Sobic.005G213600* (*Sb05g026550*), whose uncharacterized product is predicted to have a sulfotransferase domain. A third allele, *lgs1-3*, with overlapping deletion occurs in IS7777, at

Chr05:69,981,523..70,011,172, resulting in a loss of four genes, *Sobic.005G213600*, *Sobic.005G213700*, *Sobic.005G213766* and *Sobic.005G213832*.

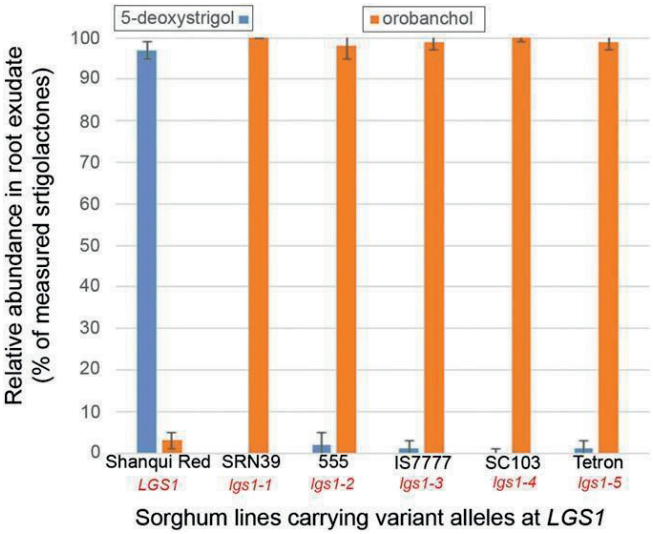


Figure 3. Chemical phenotypes of *LGS1* variants. Strigolactone profiles of root exudates from sorghum Shanqui-Red (*LGS1*) with high *Striga* germination stimulant activity, and of five low-stimulant lines with mutant alleles at SRN39 (*lgs1-1*), 555 (*lgs1-2*), IS7777 (*lgs1-3*), SC103 (*lgs1-4*) and Tetron (*lgs1-5*) are shown. Specific strigolactone quantifications are expressed in relative abundance (percent of total measured strigolactones) in each exudate. Although the absolute amount of the most abundant strigolactone varies from run to run, typical values for 5-deoxystrigol in Shanqui-Red or orobanchol in SRN39 are around 2,000 pmol per plant per 48 hrs. Values are averages of four measures from independent runs ± one SD.

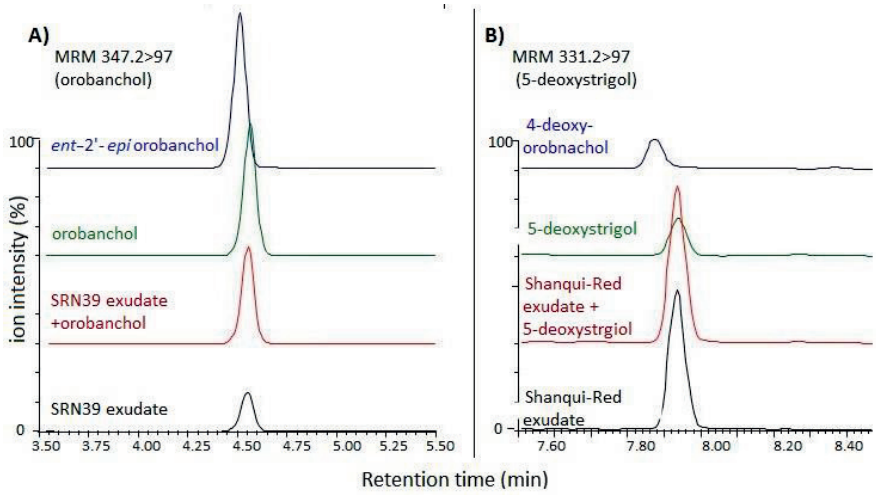


Figure 4. UPLC-MS-MS determination of identity of major strigolactones in SRN39 (low *Striga* germination stimulant activity) and Shanqui-Red (high *Striga* germination stimulant activity). Channels in the chromatograms monitor the mass transitions associated with loss of the D-ring (m/z 97) as strigolactones come off the UPLC column. The major strigolactone in SRN39 root exudate (A) coelutes with authentic orobanchol, not its enantiomer, *ent*-2'-*epi*-orobanchol. The more typical sorghum strigolactone with a β-orientation, 5-deoxystrigol, is the major one in Shanqui-Red root exudate (B). This authentic standard is resolved from the α-oriented enantiomer, 4-deoxyorobanchol.

The common deleted gene for all these alleles is *Sobic.005G213600*, which codes for the sulfotransferase. Sulfotransferases catalyze the transfer of a sulfate group from the universal donor 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to a hydroxyl or amide group of its substrate (Hirschmann et al. 2014). Several PCR primer pairs designed based on the reference genome sequence to amplify portions of *Sobic.005G213600* were used to test a diverse collection of sorghum lines with high and low *Striga* germination stimulant activity. An amplicon was always present for the target locus among accessions with high *Striga* germination stimulant activity but missing in accessions with low *Striga* germination stimulant activity determined to be allelic to SRN39.

An exception to this association was observed in the allelic low-germination stimulant lines, SC103 and Tetron, in which at least some amplicons were obtained from PCR primers targeting this gene. Examining the genomic sequence from *Sobic.005G213600* from these two accessions revealed deletions within the predicted coding region that cause frameshift mutations (Figure 5, Table S3 and Figure S2). The more obvious mutation is in SC103, which contains an allele, *lgs1-4*, having a 421-bp deletion in the second exon. This deletion not only results in a 137-aa residue loss in the predicted protein but also introduces a stop codon 46 residues downstream such that the resulting gene product, if it were translated, would be a protein 244 residues shorter than the WT protein. Tetron contains an allele, *lgs1-5*, with a 10-bp deletion 18 bp upstream of the deleted area of SC103 in the second exon, causing a frameshift that would introduce a stop codon after 39 aberrant residues beyond the deletion. A translated product of this mutant allele would therefore be missing 259 residues relative to the WT gene product (Figure S2). Both of these mutations occur within the annotated sulfotransferase domain of the gene (residues 138 – 439). The one in Tetron destroys the 5' PAPS binding motif (PKSGTTW, Figure S2) highly conserved in all sulfotransferases (Klein et al. 2006). The conserved PAPS binding residues near the end of the protein (FRKGKVGDWKNYMTPDM) would be missing in both mutant peptides (Figure S2). Therefore, all described *lgs1* alleles would lack a functional sulfotransferase product from *Sobic.005G213600*.

Expression of Sobic.005G213600

Publicly available expression profiles of *Sobic.005G231600* based on ESTs from the sorghum reference, BTx623, from the Morokoshi Sorghum Transcriptome Database (http://sorghum.riken.jp/morokoshi/Data/Sobic.005G213600) and in the expression track of Phytozome *Sorghum bicolor* v3.1 (DOE-JGI) indicate that this gene is preferentially expressed in roots and under nitrogen deficiency, two qualities one would expect for genes involved in strigolactone biosynthesis. We monitored the expression of this gene in Shanqui-Red by qRT-PCR and confirmed that expression was significantly greater in roots versus shoots (Figure 6). When seedlings of Shanqui-Red were grown in sand for 1 month irrigated with tap water and compared with seedlings irrigated with nutrient solution (12:2:31) in a potting mix (peat and perlite), expression of this gene was approximately fivefold higher under the nutrient-leached conditions. *LGS1* expression was significantly reduced in Tetron relative to Shanqui-Red in

sand (Figure 6). The qRT-PCR primers targeting the transcript were nested in the 3'-UTR (Table S2). As expected for a completely deleted gene, no expression of this target was observed in SRN39 in either medium. The severe deletion in SC103 also knocked out expression of this gene.

Motifs identified in silico using the PLACE database search tool (Higo et al. 1999) of the presumed promoter region of *Sobic.005G213600* (Figure S3) show some cis-acting regulatory elements (CAREs) that match other genes involved in strigolactone biosynthesis, including root-specificity, drought, phytohormone and nutrient deficiency responsive elements, including a phosphate deficiency response, P1BS. Most CAREs listed in Figure S3 fall within a few hundred base pairs of the transcription start site, in the presumed core promoter.

Discussion

Mutation at *LGS1* does not eliminate strigolactone biosynthesis, but rather changes the type of strigolactones present in the root exudates. In a comparison of all possible stereoisomers of the strigolactones previously reported in sorghum root exudates, strigol, sorgolactone, sorgomol and 5-deoxystrigol, it was shown that *S. hermonthica* germination was much higher when exposed to these strigolactones in their natural (β-oriented C-ring) form than when treated with their α-oriented enantiomers (Nomura et al. 2013). Furthermore, (Yoneyama et al. 2018) predicted that strigolactones containing a hydroxyl group directly on the A- (e.g., strigol) or B-ring (e.g., orobanchol) would be prone to ring-destroying nucleophilic attack and therefore be less persistent in the soil. Together, these results explain why orobanchol-exuding sorghums, like the *lgs1* mutants, would show low *Striga* germination stimulant activity in our laboratory agar assays, as well as when planted in farm fields infested with *Striga*.

We have presented compelling genetic evidence in the form of multiple mutant alleles at this locus that *LGS1* is *Sobic.005G213600*, an uncharacterized gene with a sulfotransferase domain. Unfortunately, the substrates of sulfotransferases other than a few in *Arabidopsis* are largely unknown and cannot be accurately predicted by in silico modeling based on animal enzyme structures (Hirschmann et al. 2014). Plant sulfotransferases resemble their better studied counterparts in animals by the conserved motifs involved in binding PAPS, the universal donor of the sulfate group in the reactions that they catalyze (Klein et al. 2004). They sulfate a variety of substrates and are generally divided into two main classes: membrane associated and cytosolic sulfotransferases. Only three of the formers have been described in plants (*Arabidopsis*), all sulfating tyrosine residues in relatively small secreted peptides with growth regulating activities, one that in turn stabilizes transcription factors (Matsubayashi 2011). The larger class of cytosolic plant sulfotransferases sulfate low molecular weight substrates including flavonoids, coumarins and phytohormones such as brassinosteroids, salicylic acid and jasmonates (Hirschmann et al. 2014).

The *lgs1* mutants preferentially make orobanchol, with an α-oriented C-ring over the common WT strigolactone for sorghum, 5-deoxystrigol, lacking the hydroxyl group at position 4 and having a β-oriented C-ring. The biosynthesis of strigolactones from carotenoids through carlactone continues to be elucidated in model plants such as rice and likely involves hydroxylation of C-18 and carboxylation at C-19 (Zhang et al. 2014; Al-Babili et al. 2015; Abe et al. 2014). The orientation of the C-ring with respect to the B-ring must be determined when these rings form during the cyclization that follows oxidation by the sorghum MAX1 ortholog(s). We, therefore, assume that the sulfotransferase is involved in stereocontrol of ring closure, perhaps by post-translationally modifying proteins at the site where this occurs in such a way that it favors closure to β-orientation. Alternatively, the sulfotransferase may regulate, through sulfated phytohormone intermediates, which MAX1 ortholog or other enzymes metabolize carlactone, influencing the degree to which carlactone is oxidized and the catalytic environment in which its oxidized intermediate cyclizes to a strigolactone. As some sulfotransferases do to other phytohormones, *LGS1* might even sulfate the strigolactone itself.

The hydroxyl of orobanchol, perhaps formed at low levels in sorghum by an alternative pathway, could be sulfated and drive the production of 5-deoxystrigol, whereas accumulation of its unsulfated form, as occurs in *lgs1* mutants, suppresses it. The mutant alleles described at *LGS1* will be useful for further biochemical studies on how stereochemistry of strigolactones is determined or favored and help to establish precursor and product relationships among the various types of strigolactones in sorghum and other plant species.

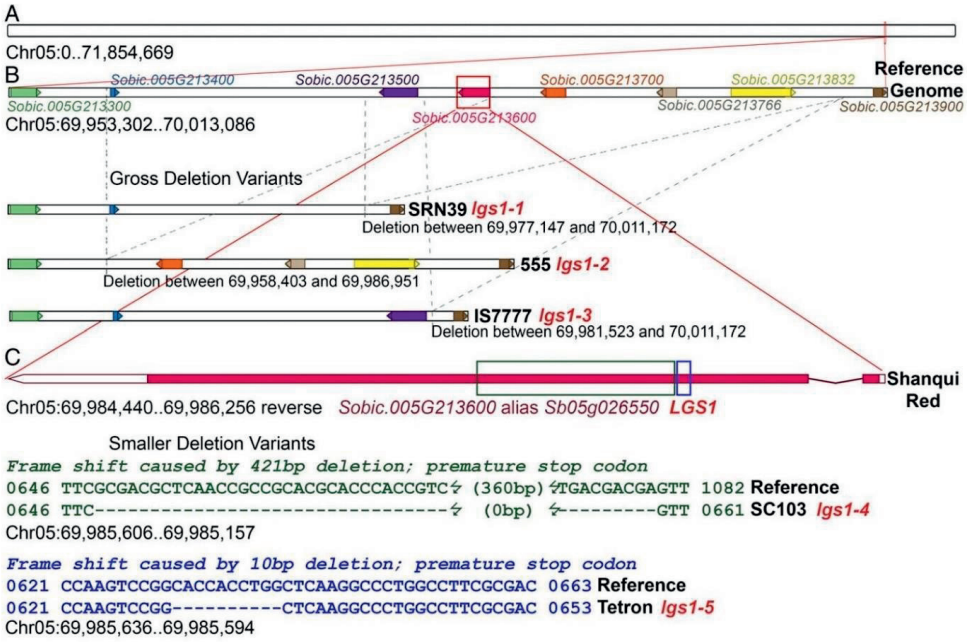


Figure 5. Schematic representation of the *LGS1* locus and its identity based on mutant analysis. (A) Genetic mapping indicated that *LGS1* was near the tip of sorghum chromosome 5 (Satish et al. 2012). (B) Fine mapping based on sequence polymorphisms indicated that the low Striga germination stimulant activity was always associated with a deleted region representing a five-gene loss in the low-stimulant parent of the mapping population, SRN39 (carrying allele *lgs1-1*). Comparing this region to other lines with low *Striga* germination stimulant activity determined to be allelic to SRN39, two other gross deletion variants were discovered with overlapping deletions, 555 (carrying allele *lgs1-2*) and IS7777 (carrying allele *lgs1-3*). The common deletion in these three is *Sobic.005G213600*. Further evidence comes from smaller deletion variants of this gene (C) in SC103 (*lgs1-4*), missing 421 bp in the second exon, and a 10-bp deletion near there in Tetron (*lgs1-5*), both predicted to cause frameshifts and severely truncated peptides without sulfotransferase function.

Striga resistance based on low germination stimulant activity has been long known and successfully exploited in sorghum (Hess et al. 1992; Ejeta 2007a) and improved varieties carrying this trait continue to show resistance to *Striga* populations from both East and West Africa (Bozkurt et al. 2015). Its simple inheritance (K. Vogler et al. 1996), particularly with molecular markers within the *LGS1* locus, make it relatively simple to introgress into existing cultivars. Mutation at *LGS1* does not knock out strigolactones in root exudates; it just changes the relative abundance of certain types, such that the other essential functions of strigolactones (ability for mycorrhizal colonization, favorable tillering and root responsiveness to nutritional deficiencies) remain intact.

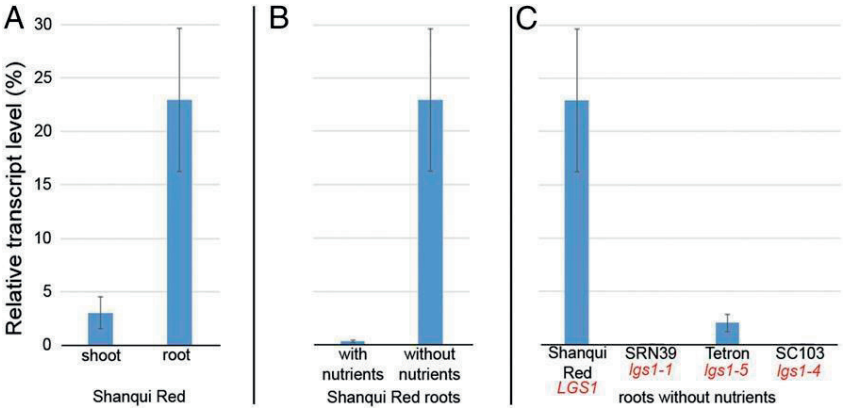


Figure 6. Expression of sorghum *LGS1*. Expression of this gene is at least five-fold higher in roots than in shoots of Shanqui-Red, carrying the WT allele *LGS1* A). RNA was extracted from 4-wk-old seedlings grown in sand without supplemental nutrients. Expression of this gene in roots under these conditions was greatly reduced compared with seedlings of the same age grown in potting mix and irrigated with nutrient solution B). Little or no expression was observed in seedling roots grown in nutrient-leached sand of mutants missing all or part of this gene C). Transcript levels were monitored by qRT-PCR comparing to actin (see Materials and Methods for details). Values are averaged from three technical and three biological replicates ± one SD.

Protection against *Striga* seems to be based on lack of responsiveness to orobanchol of those strains of the weed that parasitize sorghum and the loss of the known chemical cue 5-deoxystrigol. It should be noted, however, that rice, also parasitized by *Striga*, exudes orobanchol-type strigolactones from its roots (Cardoso et al. 2014). The nature of the protection offered by mutation at *LGS1* might be extended to other cereal hosts of *Striga* for which resistance breeding lags behind, such as maize, which exudes a variety of noncanonical strigolactones (Jamil, Kanampiu, et al. 2012), among which several of the strigol type have been reported (Yoneyama et al. 2015).

Materials and Methods

Plant materials

Sorghum lines

The same 328 RILs derived from SRN39 (a line with low *Striga* germination stimulant activity) and Shanqui-Red (high *Striga* germination activity) used to fine-map the *LGS1* locus to a 400-kb region near the tip of an arm of chromosome 5 (Satish et al. 2012) was used to further map its location. In addition to two parental lines, four additional sorghum lines with low *Striga* germination stimulant activity and reported field resistance to *Striga* (555, IS7777, SC103 and Tetron) were used to verify associations between phenotype and genotype outside the mapping population.

Striga sources

Seeds of *S. hermonthica* used to assay germination stimulant activity of sorghum exudates were collected from parasitic weeds in sorghum fields in Feddis, Ethiopia, Sinnar, Sudan and Samanko, Mali. *S. asiatica* seed used was collected from Dereshe, Ethiopia and from North Carolina. *Striga* seed were surface-sterilized in 2.5% sodium hypochlorite and Metricide 28 (2.5% glutaraldehyde; Metrex Research Corp.). After soaking in a fungicide solution containing Benomyl (0.0014% methyl 1-[butylcarbamoyl]-2-benzimidazolecarbamate) for 3 d, the *Striga* seed was embedded in 0.7% agar in 100 mm Petri dishes at a density of ~50 weed seeds per cm^2 for an additional 7 d at 29°C in darkness.

Striga germination stimulant activity

Striga germination stimulant activity of the sorghum lines were determined by the agar gel assay described previously (Hess et al. 1992) with slight modification. Sorghum accessions were surface-sterilized for 30 min in a 50% bleach solution (2.6% sodium hypochlorite) containing 0.2% Tween-20 (polyethylene glycol sorbitan monolaurate; Bio-Rad Corp.) and then imbibed overnight in a 5% aqueous slurry of Captan fungicide (active ingredient: N-trichloromethylthio-4-cyclohexene-1,2-dicarboximide, 48.9%; Arysta LifeScience North America LLC). Four imbibed sorghum seeds were planted per plate into the agar with embedded *Striga* seeds. These were thinned to one sorghum seedling per plate after 2 d in darkness at 29°C. After 2–3 d more of incubation, MGD was determined by averaging the distance, in millimeters, of the three furthest germinated weed seeds from the sorghum root. The tabulated values are the average MGDs from three plates. *Striga* germination stimulant activity was called high if MGD was ≥10 mm and low if the MGD was below 10 mm. All lines except SC103 were measured in the agar gel assay with all four sources of *Striga*. SC103 and the hybrids were measured with the *S. hermonthica* collected from Sinnar, Sudan which showed of all four sources the best germinability (80% with 10−7 M GR24).

Complementation tests among lgs1 mutants

Shanqui-Red, 555, IS7777, SC103 and Tetron were used to pollinate a male sterile version of SRN39. The male sterility used was genic based on recessive alleles at MS3. The Striga germination stimulant activity of the resulting F1 hybrids were determined by the agar gel assay with the Sudanese strain of S. hermonthica. *Striga* germination stimulant activity in the F1 hybrids with SRN39 was considered high if the MGD was ≥10 mm and low if the MGD was less than 10 mm.

Strigolactone profiles of exudates from sorghum with low and high Striga hermonthica stimulant activity

For collection of root exudates, sorghum accessions were grown in a climate room with artificial lighting at 450 µmol·m^{-2}·s^{-1} and controlled conditions [28°C (Ahonsi et al.) 10 h and 25°C (night) 14 hrs at 70% relative humidity] in Wageningen, The Netherlands. Sorghum seeds were surface-sterilized with 2% bleach for 30 min and germinated on moist filter paper in a Petri dish and incubated in darkness for 48 hrs. Germinated sorghum seeds were planted in 14 cm pots filled with sand and watered with half-strength modified Hoagland's nutrient solution containing NH4NO3 (5.6 mM), K2HPO4 (0.4 mM), MgSO4 (0.8 mM), FeSO4 (0.18 mM), CaCl2 (1.6 mM), K2SO4 (0.8 mM), MnCl2 (4.5 µM), CuSO4 (0.3 µM), ZnCl2 (1.5 µM) and Na2MoO4 (0.1 µM). After 1 wk, the seedlings were thinned to three plants per pot. Nutrient solution was applied when needed (500 mL at ~48 hrs intervals). After 4 wk, phosphorus deficiency was created in each pot to increase strigolactone production. Hereto, 1 L phosphorus-deficient nutrient solution (half-strength Hoagland's nutrient solution minus phosphate) was added to each pot and allowed to drain through the holes in the bottom of the pot to remove phosphorus from the sand. The plants were kept under phosphorus deficiency for 1 wk to increase strigolactone production. After 1 wk the pots were drained once more with 1 L of phosphorus-deficient nutrient solution to remove any accumulated strigolactones. Subsequently, 48 hrs later, root exudates were collected in 1 L plastic bottles by passing 1 L of tap water through each pot. The collected root exudates were then run through an SPE C18 column (500 mg) and strigolactones were eluted with 4 mL acetone. The acetone was then evaporated to dryness and the sample re-dissolved in 4 mL hexane and loaded on a preconditioned 200 mg Silicagel Grace Pure SPE column for further purification. The columns were eluted with 2 mL of 10:90 hexane:ethyl acetate. The solvent was subsequently evaporated to dryness and the sample re-dissolved in 200 µl of 25% acetonitrile in water and filtered through a 0.45 µm Minisart SRP filter. Per sample 0.1nmol/mL of D6-5-deoxystrigol was added as internal standard for quantification.

Strigolactones were analyzed using UPLC-MS/MS according to the method described by (Kohlen et al. 2011) with minor modifications. Chromatographic separation was achieved on an Acquity UPLC BEH C18 column (100 × 2.1 mm; 1.7 µm; Waters) by applying a water/acetonitrile gradient (containing 0.1% formic acid) to the column, starting from 5% (vol/vol) acetonitrile for 0.33 min and raised to 27% acetonitrile in 0.34 min, followed by a 4.33

min gradient to 40% acetonitrile, then rising to 65% acetonitrile in 3 min and maintained for 0.67 min, then raised to 90% acetonitrile in 0.2 min, which was maintained for 0.46 min before going back to 5% acetonitrile using a 0.2 min gradient, and maintained for 2.47 min to equilibrate the column before the next run. Operation temperature and flow rate of the column were 50°C and 0.5 mL/min, respectively. The eluent of the column was introduced into the electrospray of the mass spectrometer (Xevo triple quadrupole tandem mass spectrometer; Waters) operating in positive mode. SLs were identified by comparing the MS/MS spectrum and retention time with eight authentic standards [5-deoxystrigol, 4-deoxyorobanchol, orobanchol, *ent*-2'-*epi*-orobanchol, strigol, *epi*-strigol, sorgolactone and sorgomol]. Multiple ion monitoring mode was used to quantify the strigolactones as described previously (Jamil, Kanampiu, et al. 2012). Cone and desolvation gas flows were set to 50 L/h and 1,000 L/h, respectively. The capillary voltage was set at 3.0 kV, the source temperature was 150°C, and the desolvation temperature was 650°C. The cone voltage and collision energy were optimized for each standard compound using the Waters IntelliStart MS Console. Data acquisition and analysis were performed using MassLynx 4.1 (combined with TargetLynx) software (Waters). The proportion of the two major strigolactones per run (orobanchol and 5-deoxystrigol) was calculated from the quantities (picomoles per plant per 48 hrs) determined in four separate runs per sorghum line to construct Figure 3.

Shoot branching observations

Plants of SRN39 and Shanqui-Red were grown under standard agronomic management for sorghum in West Lafayette, IN in two-row field plots with 1 m between rows. Plants were thinned to 0.5 m spacing within rows and the number of basal tillers was counted for each of 35 individuals every 20 d during the growing season.

Mycorrhization observations

Seedlings of Shanqui-Red and SRN39 were planted as described for collecting exudates except they were initially watered with half-strength Hoagland's for the first 10 days then the growing medium (sand) was flushed to remove adsorbed phosphorus. Seven plants (one per pot) arranged in a randomized complete block design for each treatment were then inoculated with 1,500 spores of AM fungi obtained from Mycovitro by mixing the spores with 1 mL of water and injecting the inoculum into the root zone with a syringe. The treatments were R. *intraradices*, *R. clarus* and *R. custos*, a half/half mixture of *R. intraradices* and *R. custos* and an equal parts mixture of all three AM fungal species. Control treatments were inoculated with 1 mL of water containing heat-killed spores. Plants were grown an additional 5 wk, irrigating with modified half-strength Hoagland's containing only 100 µM phosphate. Harvested roots were cut into 2 cm pieces fixed in 70% ethanol and stained according to the protocol of Brundrett et al. (Brundrett et al. 1996). Mycorrhization status was determined on 100 microscopic field views of sections taken at fixed intervals along the sampled roots at 400× magnification and percent colonization calculated based on those views (McGonigle et al. 1990).

Integration of the LGS1 region genetic map onto physical map

A high-resolution map was generated around the *LGS1* locus by mapping all polymorphic markers in the 400 kb region including these new primers and the phenotypes (low or high *Striga* germination stimulant activity) collected from the 354 RILs of the mapping population (Satish et al. 2012). Informative recombinants in the high-resolution map were analyzed to locate the most likely candidate genes determining the *Striga* germination stimulant activity. strigolactone profiles of exudates of these recombinants were also determined to establish their chemical phenotype and association of parental strigolactone profile with *Striga* germination stimulant activity. Putative candidate gene markers, those that always co-segregated with the respective phenotype, were also checked against the set of sorghum lines with known *Striga* germination stimulant activity to see whether the associations held outside of the mapping population. The most likely candidate genes were sequenced in a set of low- and high-stimulating lines to characterize the nature of the mutation.

Genotype Confirmation by Sequencing

The raw reads from whole genome sequence for SC103 were downloaded from the NCBI-SRA (https://www.ncbi.nlm.nih.gov/sra/?term=SC103-14E). High-molecular-weight genomic DNA from Shanqui-Red, SRN39, 555, IS7777 and Tetron was extracted from young seedlings and used to construct indexed paired-end sequencing libraries. DNA samples were run on Illumina HiSeq2500 and generated an average of 100 bp sequence reads in fastq format. Adapter and poor quality (less than Phred-20) were removed from both 5′ and 3′ ends of the fragments then reads below a minimum length of 30 bases were discarded from the sequence dataset. The Bowtie2 sequence alignment software package was used for mapping to the reference genome determine the depth of coverage for each sample. The average depth of coverage for each sorghum line ranged from 10 to 27. The sequence data from all sorghum accessions were also mapped back to the reference genome with CLC Genomics Workbench 8.5.1 version for variant detection. Sequence reads from the *LGS1* region of Shanqui-Red, SRN39, 555, IS7777 and Tetron have been deposited with NCBI-SRA under study accession no. SRP098704.

Expression analysis

Surface-sterilized and imbibed sorghum seeds (as in the germination stimulant activity assays) were germinated at 30°C on moistened sterile filter papers in Petri dishes and germinated seedlings were transferred to 1 dm^3 pots filled with potting mix (milled peat and perlite) or sand. The seedlings in sand were grown for 4 wk and watered with tap water in which no nutrients were provided. Those in the potting mix were irrigated with a 12:2:31 (nitrogen:phosphorus:potassium) solution containing micronutrients (Stewart 1972). At harvest, the roots and shoots were frozen separately in liquid nitrogen and stored at −80°C for RNA extraction. RNA extraction was done using the Spectrum Plant Total RNA kit (Sigma). One hundred micrograms of frozen root or leaf sample was ground to a fine powder in liquid nitrogen. Protocol A of the manufacturer's instructions was followed for cell lysis and

subsequent RNA purification steps. The On-Column DNase I Digestion Set (Sigma) was used according the manufacturer's instructions to eliminate any genomic DNA contamination. One microgram RNA from each sample was reverse-transcribed using the RETROscript kit (Ambion). The template RNA in a 12 µL reaction volume was mixed with 2 µL oligo (dT) primer and heated for 3 min at 85 °C followed by brief spinning. The rest of the reaction components (2 µL 10× RT buffer, 4 µL dNTP, 1 µL RNase inhibitor and 1 µL MMLV-RT reverse transcriptase) was added to the mix and incubated at 44°C for 1 h. The reverse transcription reaction was inactivated by heating at 92°C for 10 min. The resulting reverse transcribed product was diluted with double-deionized water to 100 µL and used as a template for following quantitative PCR. PCR primers targeting the 3' UTR were designed using Primer3 software (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The iTaq Universal SYBR Green Supermix (Bio-Rad) was used for qPCR. Primers at a final concentration of 0.5 µM each, 20 ng template cDNA, 10 µL the 2× reaction mix and double-deionized water to final 20 µL reaction volume were combined for qPCR. The reaction was carried out with three technical and biological replicates with the Mx3005P system. The PCR conditions were 2 min at 95°C followed by 45 cycles of 95°C for 15 s and 60°C for 30 s. The plate values of the target gene and housekeeping gene (actin) (*Sobic.001G112600*) were used to determine the relative transcript level.

Acknowledgments

We thank K. Yoneyama (Weed Science Center, Utsunomiya University, Utsunomiya, Japan) and T. Asami (Department of Applied Biological Chemistry, The University of Tokyo, Tokyo, Japan) for supplying strigolactone standards. This work was supported by Bill and Melinda Gates Foundation Grant OPP1006216 and Netherlands Organization for Scientific Research Vici Grant 865.06.002 and Equipment Grant 834.08.001 (to H.B.).

References

Abe, Satoko, Aika Sado, Kai Tanaka, Takaya Kisugi, Kei Asami, Saeko Ota, Hyun Il Kim, et al. 2014. 'Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18084-89.

Ahonsi, Monday O., Dana K. Berner, Alphonse M. Emechebe, and Segun T. Lagoke. 2002. 'Selection of rhizobacterial strains for suppression of germination of Striga hermonthica (Del.) Benth. seeds', *Biological control*, 24: 143-52.

Akiyama, K., K. Matsuzaki, and H. Hayashi. 2005. 'Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi', *Nature*, 435: 824-7.

Al-Babili, Salim, and Harro J. Bouwmeester. 2015. 'Strigolactones, a Novel Carotenoid-Derived Plant Hormone', *Annual Review of Plant Biology*, 66: 161-86.

Alder, Adrian, Muhammad Jamil, Mattia Marzorati, Mark Bruno, Martina Vermathen, Peter Bigler, Sandro Ghisla, et al. 2012. *The Path from -Carotene to Carlactone, a Strigolactone-Like Plant Hormone*.

Awad, Ayman, Daisuke Sato, Dai Kusumoto, Hiroaki Kamioka, Yasutomo Takeuchi, and Koichi Yoneyama. 2006. 'Characterization of Strigolactones, Germination Stimulants for the Root Parasitic Plants *Striga* and *Orobanche*, Produced by Maize, Millet and Sorghum', *Plant Growth Regulation*, 48: 221-27.

Bozkurt, M. L., P. Muth, H. K. Parzies, and B. I. G. Haussmann. 2015. 'Genetic diversity of East and West African *Striga hermonthica* populations and virulence effects on a contrasting set of sorghum cultivars', *Weed Research*, 55: 71-81.

Brundrett, Mark, Neale Bougher, Bernie Dell, Tim Grove, and Nick Malajczuk. 1996. "Working with mycorrhizas in forestry and agriculture." In.: Australian Centre for International Agricultural Research Canberra.

Cardoso, C., Y. Zhang, M. Jamil, J. Hepworth, T. Charnikhova, S. O. Dimkpa, C. Meharg, et al. 2014. 'Natural variation of rice strigolactone biosynthesis is associated with the deletion of two MAX1 orthologs', *Proc Natl Acad Sci U S A*, 111: 2379-84.

Cook, C. E., Leona P. Whichard, Beverly Turner, Monroe E. Wall, and Grant H. Egley. 1966. 'Germination of Witchweed (Striga lutea Lour.): Isolation and Properties of a Potent Stimulant', *Science*, 154: 1189.

Ejeta, Gebisa. 2007a. 'Breeding for *Striga* Resistance in Sorghum: Exploitation of an Intricate Host-Parasite Biology', *Crop Science*, 47: S216-S27.

Gomez-Roldan, M. V., S. Fermas, P. B. Brewer, V. Puech-Pages, E. A. Dun, J. P. Pillot, F. Letisse, et al. 2008. "Strigolactone inhibition of shoot branching." In *Nature 455 (2008)*. 2008.

Hess, Dale E., Gebisa Ejeta, and Larry G. Butler. 1992. 'Selecting sorghum genotypes expressing a quantitative biosynthetic trait that confers resistance to Striga', *Phytochemistry*, 31: 493-97.

Higo, K., Y. Ugawa, M. Iwamoto, and T. Korenaga. 1999. 'Plant cis-acting regulatory DNA elements (PLACE) database: 1999', *Nucleic Acids Res*, 27: 297-300.

Hirschmann, F., F. Krause, and J. Papenbrock. 2014. 'The multi-protein family of sulfotransferases in plants: composition, occurrence, substrate specificity, and functions', *Front Plant Sci*, 5: 556.

Jamil, M., F. K. Kanampiu, H. Karaya, T. Charnikhova, and H. J. Bouwmeester. 2012. *'Striga hermonthica* parasitism in maize in response to N and P fertilisers', *Field Crops Research*, 134: 1-10.

Jamil, M., T. A. Van Mourik, T. Charnikhova, and H. J. Bouwmeester. 2013. 'Effect of diammonium phosphate application on strigolactone production and Striga hermonthica infection in three sorghum cultivars', *Weed Research*, 53: 121-30.

K. Vogler, R., G. Ejeta, and L. G. Butler. 1996. *Inheritance of Low Production of Striga Germination Stimulant in Sorghum*.

Klein, M., and J. Papenbrock. 2004. 'The multi-protein family of Arabidopsis sulphotransferases and their relatives in other plant species', *J Exp Bot*, 55: 1809-20.

Klein, M., M. Reichelt, J. Gershenzon, and J. Papenbrock. 2006. 'The three desulfoglucosinolate sulfotransferase proteins in Arabidopsis have different substrate specificities and are differentially expressed', *Febs j*, 273: 122-36.

Kohlen, W., T. Charnikhova, L. Qing, R. M. E. H. Bours, M. Domagalska, S. Beguerie, F. W. A. Verstappen, et al. 2011. "Strigolactones are transported through the xylem and play a key role in shoot architectural response to phosphate deficiency in nonarbuscular mycorrhizal host Arabidopsis." In *Plant Physiology 155 (2011) 2*. 2011.

Matsubayashi, Yoshikatsu. 2011. 'Post-translational modifications in secreted peptide hormones in plants', *Plant Cell Physiol*, 52: 5-13.

Mboob, S. S. 1989. 'A regional program for *Striga* control in West and Central Africa', *Striga–Improved management in Africa*: 190-94.

McGonigle, T. P., M. H. Miller, D. G. Evans, G. L. Fairchild, and J. A. Swan. 1990. 'A new method which gives an objective measure of colonization of roots by vesicular—arbuscular mycorrhizal fungi', *New Phytologist*, 115: 495-501.

Motonami, Noriko, Kotomi Ueno, Hitomi Nakashima, Saki Nomura, Masaharu Mizutani, Hirosato Takikawa, and Yukihiro Sugimoto. 2013. 'The bioconversion of 5-deoxystrigol to sorgomol by the sorghum, *Sorghum bicolor* (L.) Moench', *Phytochemistry*, 93: 41-48.

Nomura, S., H. Nakashima, M. Mizutani, H. Takikawa, and Y. Sugimoto. 2013. 'Structural requirements of strigolactones for germination induction and inhibition of *Striga gesnerioides* seeds', *Plant Cell Rep*, 32: 829-38.

Pérez-Vich, Begoña, Leonardo Velasco, Patrick J. Rich, and Gebisa Ejeta. 2013. 'Marker-Assisted and Physiology-Based Breeding for Resistance to Root Parasitic Orobanchaceae.' in Daniel M. Joel, Jonathan Gressel and Lytton J. Musselman (eds.), *Parasitic Orobanchaceae: Parasitic Mechanisms and Control Strategies* (Springer Berlin Heidelberg: Berlin, Heidelberg).

Rasmussen, Amanda, Michael Glenn Mason, Carolien De Cuyper, Philip B. Brewer, Silvia Herold, Javier Agusti, Danny Geelen, et al. 2012. 'Strigolactones Suppress Adventitious Rooting in Arabidopsis and Pea', *Plant Physiology*, 158: 1976.

Rich, Patrick J., and Gebisa Ejeta. 2007. 'BIOLOGY OF HOST-PARASITE INTERACTIONS IN *STRIGA* SPECIES.' in, *Integrating New Technologies for Striga Control* (WORLD SCIENTIFIC).

Rodenburg, Jonne, Matty Demont, Sander J. Zwart, and Lammert Bastiaans. 2016. 'Parasitic weed incidence and related economic losses in rice in Africa', *Agriculture, Ecosystems & Environment*, 235: 306-17.

Satish, K., Z. Gutema, C. Grenier, P. J. Rich, and G. Ejeta. 2012. 'Molecular tagging and validation of microsatellite markers linked to the low germination stimulant gene (lgs) for Striga resistance in sorghum [Sorghum bicolor (L.) Moench]', *Theor Appl Genet*, 124: 989-1003.

Siame, Bupe A., Yohan Weerasuriya, Karl Wood, Gebisa Ejeta, and Larry G. Butler. 1993. 'Isolation of strigol, a germination stimulant for Striga asiatica, from host plants', *J Agric Food Chem*, 41: 1486-91.

Stewart, G. R. 1972. 'The Regulation of Nitrite Reductase Level in Lemna minor L', *Journal of Experimental Botany*, 23: 171-83.

Sun, Xue-Guang, and Ming Tang. 2013. 'Effect of arbuscular mycorrhizal fungi inoculation on root traits and root volatile organic compound emissions of Sorghum bicolor', *South African Journal of Botany*, 88: 373-79.

Xie, X., K. Yoneyama, T. Kisugi, K. Uchida, S. Ito, K. Akiyama, H. Hayashi, et al. 2013. 'Confirming stereochemical structures of strigolactones produced by rice and tobacco', *Mol Plant*, 6: 153-63.

Yoneyama, K., N. Mori, T. Sato, A. Yoda, X. Xie, M. Okamoto, M. Iwanaga, et al. 2018. 'Conversion of carlactone to carlactonoic acid is a conserved function of MAX1 homologs in strigolactone biosynthesis', *New Phytol*, 218: 1522-33.

Yoneyama, Kaori, Ryota Arakawa, Keiko Ishimoto, Hyun Il Kim, Takaya Kisugi, Xiaonan Xie, Takahito Nomura, et al. 2015. 'Difference in Striga-susceptibility is reflected in strigolactone secretion profile, but not in compatibility and host preference in arbuscular mycorrhizal symbiosis in two maize cultivars', *New Phytologist*, 206: 983-89.

Zhang, Y., A. D. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, et al. 2014. 'Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis', *Nat Chem Biol*, 10: 1028-33.

Supplementary information

Supplementary figures

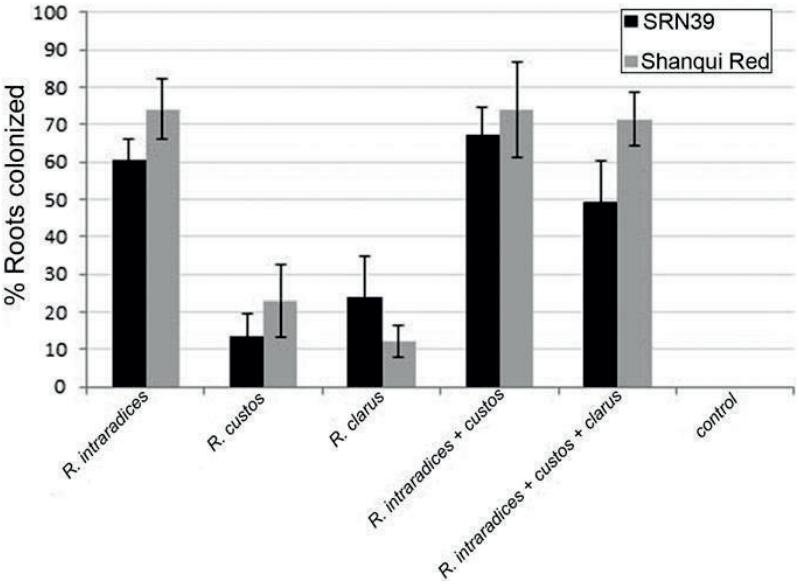


Figure S1. Mycorrhization status (percent roots colonized) of Shanqui-Red (*LGS1*) and SRN39 (*lgs1*) roots by three AM fungi Rhizophagus species 5 wk after inoculation. Plants were grown in sand with suboptimal phosphate. Values are averaged from seven plants ± one SD.

```
5:69,986,256                                                                 MetAsnVal GlnGluArgA rgLysGluLe 010
0001        ATGCTCATAC GGTTATATAA ACCAGTCTAA TAGATCGAGG ACGTACGTTT    ACAGTGCCTC GCACAAAGTA C  ATGAACGTA CAGGAGAGGA GGAAGGAACT 0100
            ATGCTCATAC GGTTATATAA ACCAGTCTAA TAGATCGAGG ACGTACGTTT ACAGTGCCTC GCACAAAGTA      CATGAACGTA CAGGAGAGGA GGAAGGAACT
            ATGCTCATAC GGTTATATAA ACCAGTCTAA TAGATCGAGG ACGTACGTTT ACAGTGCCTC GCACAAAGTA CATGAACGTA CAGGAGAGGA GGAAGGAACT
            ATGCTCATAC GGTTATATAA ACCAGTCTAA TAGATCGAGG ACGTACGTTT         ACAGTGCCTC GCACAAAGTA CATGAACGTA CAGGAGAGGA GGAAGGAACT

     uGluGluArg SerSerThrT  hrLeuGl                                                                            019
0101 CGAAGAGAGA TCATCGACGA CATTGGG              GTG AGTTTTGGCA TGCAAATGCA     AGCAATTCTT CCTTGTTGCC TTCTTTGAAC TGTCATGCAT ACATTGAGGG 0200
            CGAAGAGAGA TCATCGACGA CATTGGGGTG AGTTTTGGCA TGCAAATGCA        AGCAATTCTT CCTTGTTGCC TTCTTTGAAC TGTCATGCAT ACATTGAGGG
            CGAAGAGAGA TCATCGACGA CATTGGGGTG AGTTTTGGCA TGCAAATGCA    AGCAATTCTT     CCTTGTTGCC TTCTTTGAAC TGTCATGCAT ACATTGAGGG
     CGAAGAGAGA TCATCGACGA                    CATTGGGGTG AGTTTTGGCA TGCAAATGCA AGCAATTCTT CCTTGTTGCC TTCTTTGAAC TGTCATGCAT ACATTGAGGG

                                                           yHi sLeuHisThr IleArgAsnT hrProAlaGl ySerSerMet SerThrThrT hrCysTyrSe 040
             0201 ATCCTTAGTT TTGATGCATG CATGTAAACT CTTATAGGCA CCTACATACG    ATACGAAACA CACCAGCAGG CAGCAGCATG AGTACTACTA CTTGCTACTC 0300
            ATCCTTAGTT TTGATGCATG CATGTAAACT CTTATAGGCA CCTACATACG         ATACGAAACA CACCAGCAGG CAGCAGCATG AGTACTACTA CTTGCTACTC
            ATCCTTAGTT TTGATGCATG CATGTAAACT CTTATAGGCA CCTACATACG         ATACGAAACA CACCAGCAGG CAGCAGCATG AGTACTACTA CTTGCTACTC
            ATCCTTAGTT TTGATGCATG CATGTAAACT CTTATAGGCA CCTACATACG         ATACGAAACA CACCAGCAGG CAGCAGCATG AGTACTACTA CTTGCTACTC

     rAlaProAla AlaValValP                roGlyAlaGl yGlyGluVal AlaValValT hrAlaValAl aSerGluAla GlyAlaAlaA laAlaHisAs pGlnSerArg 073
0301        TGCGCCAGCA GCCGTCGTCC CCGGCGCCGG CGGTGAAGTG GCGGTGGTTA         CTGCAGTGGC ATCGGAAGCC GGAGCAGCAG CAGCTCATGA CCAGTCGAGG 0400
            TGCGCCAGCA GCCGTCGTCC CCGGCGCCGG CGGTGAAGTG GCGGTGGTTA         CTGCAGTGGC ATCGGAAGCC GGAGCAGCAG CAGCTCATGA CCAGTCGAGG
            TGCGCCAGCA GCCGTCGTCC CCGGCGCCGG CGGTGAAGTG GCGGTGGTTA         CTGCAGTGGC ATCGGAAGCC GGAGCAGCAG CAGCTCATGA CCAGTCGAGG
            TGCGCCAGCA GCCGTCGTCC CCGGCGCCGG CGGTGAAGTG GCGGTGGTTA         CTGCAGTGGC ATCGGAAGCC GGAGCAGCAG CAGCTCATGA CCAGTCGAGG

     LysLysLysA snAsnHisAr                gSerLeuTyr AlaAsnLeuP roAlaAlaGl uIleIleAsp SerLeuProL euGluThrAr gPheProVal ProHisArgL 107
0401        AAGAAGAAGA ACAACCACCG GAGCCTGTAC GCCAACCTAC CTGCAGCCGA         GATCATCGAC TCGCTGCCCC TGGAGACACG GTTCCCGGTG CCCCATCGCC 0500
            AAGAAGAAGA ACAACCACCG GAGCCTGTAC GCCAACCTAC CTGCAGCCGA GATCATCGAC TCGCTGCCCC TGGAGACACG GTTCCCGGTG CCCCATCGCC
            AAGAAGAAGA ACAACCACCG GAGCCTGTAC GCCAACCTAC CTGCAGCCGA GATCATCGAC TCGCTGCCCC TGGAGACACG GTTCCCGGTG CCCCATCGCC
            AAGAAGAAGA ACAACCACCG GAGCCTGTAC GCCAACCTAC CTGCAGCCGA         GATCATCGAC TCGCTGCCCC TGGAGACACG GTTCCCGGTG CCCCATCGCC

            euTyrGlyGl yPheTrpLys AlaGluPheL euLeuLysGl yMetAlaAla         AlaAlaAlaA rgThrThrSe rCysPheGlu PheGluProA snProSerAs 140
0501        TGTACGGGGG CTTCTGGAAA GCCGAGTTCT TGCTCAAGGG CATGGCAGCC         GCTGCCGCTC GCACCACGTC GTGCTTCGAG TTCGAGCCAA ACCCATCGGA 0600
            TGTACGGGGG CTTCTGGAAA GCCGAGTTCT TGCTCAAGGG CATGGCAGCC         GCTGCCGCTC GCACCACGTC GTGCTTCGAG TTCGAGCCAA ACCCATCGGA
            TGTACGGGGG CTTCTGGAAA GCCGAGTTCT TGCTCAAGGG CATGGCAGCC         GCTGCCGCTC GCACCACGTC GTGCTTCGAG TTCGAGCCAA ACCCATCGGA
            TGTACGGGGG CTTCTGGAAA GCCGAGTTCT TGCTCAAGGG CATGGCAGCC         GCTGCCGCTC GCACCACGTC GTGCTTCGAG TTCGAGCCAA ACCCATCGGA

                                 ┌5'-PAPS binding motif┐
                                                                Phe____________________________________________ 159
                                                    Gl__________ySerArgPro TrpProSerA rgArgSerTr rAlaAlaArg ThrHisArgP roThrProTh 170
     pIlePheLeu AlaSerLeuP    roLysSerGl yThrThrTrp               LeuLysAlaL euAlaPheAl aThrLeuAsn ArgArgThrH isProProSe rAsnAlaAsp 173
0601 CATCTTCCTC GCCAGCTTGC    CCAAGTCCGG CACCACCTGG               CTCAAGGCCC TGGCCTTCGC GACGCTCAAC CGCCGCACGC ACCCACCGTC CAACGCCGAC 0700
            CATCTTCCTC GCCAGCTTGC CCAAGTCCGG CACCACCTGG CTCAAGGCCC TGGCCTTCGC GACGCTCAAC CGCCGCACGC ACCCACCGTC CAACGCCGAC 0700
     CATCTTCCTC GCCAGCTTGC CCAAGTCCGG    ----------               CTCAAGGCCC TGGCCTTCGC GACGCTCAAC CGCCGCACGC ACCCACCGTC CAACGCCGAC 0690
            CATCTTCCTC GCCAGCTTGC CCAAGTCCGG CACCACCTGG CTCAAGGCCC TGGCCTTC--      ---------- ----------      ---------- ---------- 0658

            _________________________________________________________________________________________________ 159
            rAlaSerThr ArgSerAlaT hrAlaThrPr oThrThrVal SerAlaSerT rpSerSer** *                                189
            GlyGlnHisP roPheSerHi sArgAsnPro HisAspCysV alSerPheLe uGluLeuMet MetIleGlnG lyValAspAl aAlaAlaAla AspAspAspA 207
0701        GGCCAGCACC CGTTCAGCCA CCGCAACCCC CACGACTGTG TCAGCTTCCT GGAGCTCATG ATGATCCAGG GCGTCGACGC CGCCGCCGCC GACGACGACG 0800
            GGCCAGCACC CGTTCAGCCA CCGCAACCCC CACGACTGTG TCAGCTTCCT GGAGCTCATG ATGATCCAGG GCGTCGACGC CGCCGCCGCC GACGACGACG 0800
            GGCCAGCACC CGTTCAGCCA CCGCAACCCC CACGACTGTG TCAGCTTCCT GGAGCTCATG ATGATCCAGG GCGTCGACGC CGCCGCCGCC GACGACGACG 0790
     ---------- ----------  ----------  ----------            ---------- ----------  ---------- ----------    ---------- ---------- 0658

            _________________________________________________________________________________________________ 159
            ___________laGlyAla ProArgLeuI leAlaThrHi sLeuProTrp SerTrpLeuP roProAlaIl eThrAlaGly GluGlyGlnG lyGlyGlySe 236
0801        ACGCCGACGA CGCCGGAGCT CCACGGCTGA TCGCCACGCA CTTGCCCTGG TCATGGCTTC CCCCTGCCAT CACGGCGGGG GAAGGACAAG GAGGCGGCTC 0900
                         --CCGGAGCT CCACGGCTGA TCGCCACGCA CTTGCCCTGG TCATGGCTTC CCCCTGCCAT CACGGCGGGG GAAGGACAAG GAGGCGGCTC 0888
            ACGCCGACGA CGCCGGAGCT CCACGGCTGA TCGCCACGCA CTTGCCCTGG TCATGGCTTC CCCCTGCCAT CACGGCGGGG GAAGGACAAG GAGGCGGCTC 0890
     ---------- ----------  ----------  ----------            ---------- ----------  ---------- ----------    ---------- ---------- 0658

            _________________________________________________________________________________________________ 159
            rSerSerArg GlyArgGlyC ysArgIleVa lTyrValCys ArgGluProL ysAspValLe uValSerTyr TrpThrPheS erValLysAl aAlaAlaLys 269
0901        CTCCTCCCGG GGCCGGGGGT GCCGGATCGT GTACGTGTGC CGGGAACCCA AGGACGTGCT GGTCTCCTAC TGGACCTTCA GCGTCAAGGC GGCAGCCAAG 1000
            CTCCTCCCGG GGCCGGGGGT GCCGGATCGT GTACGTGTGC CGGGAACCCA AGGACGTGCT GGTCTCCTAC TGGACCTTCA GCGTCAAGGC GGCAGCCAAG 0988
            CTCCTCCCGG GGCCGGGGGT GCCGGATCGT GTACGTGTGC CGGGAACCCA AGGACGTGCT GGTCTCCTAC TGGACCTTCA GCGTCAAGGC GGCAGCCAAG 0990
     ---------- ----------  ----------  ----------            ---------- ----------  ---------- ----------    ---------- ---------- 0658

            ______________________________________________________________________________V alSerArgAr gLeuSerSer     166
            PheAlaAlaA laAlaAlaAl aGlyGlyAsp AspAspGlyG lyGlyGlyAr gGluSerAla AlaAlaSerL euThrThrSe rPheGluGlu AlaPheGluL 303
1001        TTCGCCGCCG CCGCTGCCGC CGGTGGCGAT GATGATGGCG GCGGCGGCCG GGAGTCTGCA GCAGCAAGCC TGACGACGAG TTTCGAGGAG GCTTTCGAGC 1100
            TTCGCCGCCG CCGCTGCCGC CGGTGGCGAT GATGATGGCG GCGGCGGCCG GGAGTCTGCA GCAGCAAGCC TGACGACGAG TTTCGAGGAG GCTTTCGAGC 1088
            TTCGCCGCCG CCGCTGCCGC CGGTGGCGAT GATGATGGCG GCGGCGGCCG GGAGTCTGCA GCAGCAAGCC TGACGACGAG TTTCGAGGAG GCTTTCGAGC 1090
     ---------- ----------  ----------  ----------            ---------- ----------  ---------- ----------   G TTTCGAGGAG GCTTTCGAGC 0679

     SerSerAlaA rgAspGlyS erProAlaAla HisThrGlyS erThrArgTr pAsnSerGly ValArgAlaA rgGlyGlyPr oThrArgCys CysSerSerG     166
     euPheCysGl uGlyArgPhe ProGlyGlyP roHisTrpLe uHisAlaLeu GluPheTrpA rgGluSerGl nArgArgPro AspGluValL euPheLeuAr     336 protein
1101 TCTTCTGCGA GGGACGGTTC CCCGGCGGCC CACACTGGCT CCACGCGCTG GAATTCTGGC GTGAGAGCCA GAGGAGGCCC GACGAGGTGC TGTTCCTCAG 1200 REF
     TCTTCTGCGA GGGACGGTTC CCCGGCGGCC CACACTGGCT CCACGCGCTG GAATTCTGGC GTGAGAGCCA GAGGAGGCCC GACGAGGTGC TGTTCCTCAG   1188 SQR WT
     TCTTCTGCGA GGGACGGTTC CCCGGCGGCC CACACTGGCT CCACGCGCTG GAATTCTGGC GTGAGAGCCA GAGGAGGCCC GACGAGGTGC TGTTCCTCAG   1190 Tetron 5
     TCTTCTGCGA GGGACGGTTC CCCGGCGGCC CACACTGGCT CCACGCGCTG GAATTCTGGC GTGAGAGCCA GAGGAGGCCC GACGAGGTGC TGTTCCTCAG   0779 SC103    4

     lyThrArgTh rCys***                                                                                          170
     gTyrGluAsp MetLeuArgA spProValGl yAsnLeuArg LysLeuAlaA laPheMetGl yCysProPhe SerAlaGluG luGluThrAl aGlyGlyGly   369   protein
1201 GTACGAGGAC ATGCTGAGAG ATCCGGTAGG GAACCTGAGG AAGCTCGCGG CGTTCATGGG GTGCCCGTTC TCGGCGGAGG AGGAGACGGC CGGCGGCGGC 1300 REF
     GTACGAGGAC ATGCTGAGAG ATCCGGTAGG GAACCTGAGG AAGCTCGCGG CGTTCATGGG GTGCCCGTTC TCGGCGGAGG AGGAGACGGC CGGCGGCGGC 1288 SQR WT
     GTACGAGGAC ATGCTGAGAG ATCCGGTAGG GAACCTGAGG AAGCTCGCGG CGTTCATGGG GTGCCCGTTC TCGGCGGAGG AGGAGACGGC CGGCGGCGGC 1290 Tetron 5
     GTACGAGGAC ATGCTGAGAG ATCCGGTAGG GAACCTGAGG AAGCTCGCGG CGTTCATGGG GTGCCCGTTC TCGGCGGAGG AGGAGACGGC CGGCGGCGGC 0879 SC103   4
```

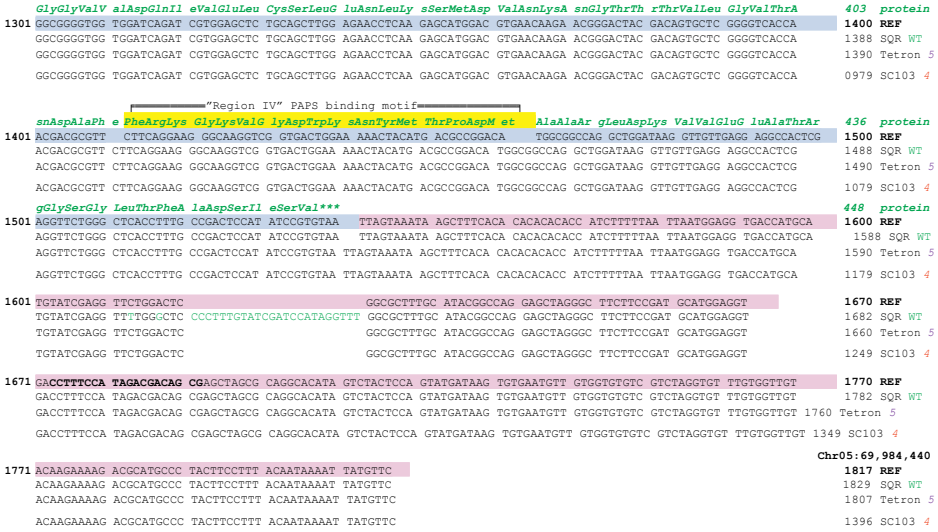


Figure S2. Sequence comparison of WT Shanqui-Red (*LGS1*) and mutant alleles from Tetron (*lgs1-5*) and SC103 (*lgs1-4*) with the sorghum reference genome, BTx623 for the reverse strand of *Sobic.005G213600*. Genomic sequences of Shanqui-Red and the mutants are from consensus of Illumina reads. Variations from the reference DNA sequence are indicated by bases in green, purple and red for Shanqui-Red, Tetron and SC103, respectively. The predicted WT *LGS1* protein is shown in green in three-letter residue code above codons. The predicted variations in mutant peptides are shown in purple and red for Tetron and SC103, respectively. Residues highlighted in yellow are PAPS binding regions that are broadly conserved in sulfotransferases of all organisms. Chromosome positions according to Phytozome *Sorghum bicolor* v3.1, DOE-JGI.

>Upstream region and the 1st 100bp of *Sobic.005G213600* (Chr05: 60,372,676..60,374,475)

```
-1700 +   TGCTCTTAAC GCATATTCGT   ATCTCTTTCA CATACAAAGC ATTCTAATTA ACCATGCTCT GTCTTTTGAT AAGCAGGGGA GATGTAATTA TTCTTTTTGT  -1601
      -   ACGAGAATTG CGTATAAGCA   TAGAGAAAGT GTATGTTTCG TAAGATTAAT TGGTACGAGA CAGAAAACTA TTCGTCCCCT CTACATTAAT AAGAAAAACA  60374376

-1600 + TGTTAAGGTC TAATGGTGTC ATCCATCCTC CCTAGCCTAT ATATCTTATA AAGTTATCTT AATTTCTATT ATCATTCTAT CTTAACATGC AAGCTATCCA  -1501
      - ACAATTCCAG ATTACCACAG TAGGTAGGAG GGATCGGATA TATAGAATAT TTCAATAGAA TTAAAGATAA TAGTAAGATA GAATTGTACG TTCGATAGGT  60374276

-1500 + TGTGTCGCAA CCCATGTCTC ATAAGATGCA TAATGTATCA TCACTGTTTT ATCTCACAAC CGGCAACAAT TAAACTTTAT GTAGTAGCGC ACTAATAAGT  -1401
      - ACACAGCGTT GGGTACAGAG TATTCTACGT ATTACATAGT AGTGACAAAA TAGAGTGTTG GCCGTTGTTA ATTTGAAATA CATCATCGCG TGATTATTCA  60374176

-1400 + GCCCTATCAT TTTTACTAAG TTAAATTAAT ATTTCAGTCA TGTAGGGGCT CTTTGGCCCC TCCAATTCTG TCCAAAAAAA TAATATTTTT AATTTAACCT  -1301
      - CGGGATAGTA AAAATGATTC AATTTAATTA TAAAGTCAGT ACATCCCCGA GAAACCGGGG AGGTTAAGAC AGGTTTTTTT ATTATAAAAA TTAAATTGGA  60374076

-1300 + ATCATGTGTC ACACCTATAC TAAATGCCAA AGGAGGTGGT TTTGTGGGTT ATGGTGTGCA ACATATGTGG ACTCATAAGT CCAGGTTGAC CAGGCTGCCC  -1201
      - TAGTACACAG TGTGGATATG ATTTACGGTT TCCTCCACCA AAACACCCAA TACCACACGT TGTATACACC TGAGTATTCA GGTCCAACTG GTCCGACGGG 60373976

-1200 + TATTTTGATG ACCTTCTTTT GCCGCATAAC ATTGATGTAA TGCACACTGA AAAGAATATT GCCGAGGCCC TTTGGGGCAC TCTTATGAAC ACAGAGAAAT  -1101
      - ATAAAACTAC TGGAAGAAAA CGGCGTATTG TAACTACATT ACGTGTGACT TTTCTTATAA CGGCTCCGGG AAACCCCGTG AGAATACTTG TGTCTCTTTA 60373876

-1100 + CAAAGGACAA TGTTAAGGCT AGAGTGGACT TGTCGACGTT ATGCGATAGA AATAAACAAG AGATGCAACC TCCTAGTGGT CAAAACAAGA AGTGGAAAAA  -1001
      - GTTTCCTGTT ACAATTCCGA TCTCACCTGA ACAGCTGCAA TACGCTATCT TTATTTGTTC TCTACGTTGG AGGATCACCA GTTTTGTTCT TCACCTTTTT  60373776

-1000 + GCCTAAGGTG GATTTCGTCT TAAAAATAGA TGCGAGAAGG GAAGTGCTTG AATGGATCAA GAACTTGATG TTTCCAGATG GCTATGCCGC TAATTTGAGC  -0901
      - CGGATTCCAC CTAAAGCAGA ATTTTTATCT ACGCTCTTCC CTTCACGAAC TTACCTAGTT CTTGAACTAC AAAGGTCTAC CGATACGGCG ATTAAACTCG  60373676

-0900 + AGAGGAGTGA ACTTAACCAC TTTGCGAGTC AACGGGATGA AGAGTCATGA CTACCATGTA TGGATTGAGC GGCTTCTCCC TACAATGGTC CGTGGGTATG  -0801
      - TCTCCTCACT TGAATTGGTG AAACGCTCAG TTGCCCTACT TCTCAGTACT GATGGTACAT ACCTAACTCG CCGAAGAGGG ATGTTACCAG GCACCCATAC  60373576

-0800 + TCCCTGAGCA TGTCTGGAAA GTGTTGGCTG AGTTGAGCTA TTTCTTCCGC CTGTTATGAT GCGCATTTAT GTCATAGATG TTGATACTCT TTTGTATAAA  -0701
      - AGGGACTCGT ACAGACCTTT CACAACCGAC TCAACTCGAT AAAGAAGGCG GACAATACTA CGCGTAAATA CAGTATCTAC AACTATGAGA AAACATATTT
      60373476

-0700 + GTTAGTCAAA CTTTAAAAAA TTTGACTGGC ACGGATTCTA GAAATTAGTT CTTTCGTGGA CGGAGGGAGT GCAAAACATA TTCACGCCAT AATACTTGCT  -0601

      - CAATCAGTTT GAAATTTTTT AAACTGACCG TGCCTAAGAT CTTTAATCAA GAAAGCACCT GCCTCCCTCA CGTTTTGTAT AAGTGCGGTA TTATGAACGA 60373376
 -0600 + TGGCCATCAC ATCATCCTAT GTTTGATCTA ATATCGGATT ACCTTTTTTC ATTATTCCCT CCCTTTTAAA TTATAAGTCA TTTTAGCTTT TCTAGATATA  -0501

      - ACCGGTAGTG TAGTAGGATA CAAACTAGAT TATAGCCTAA TGGAAAAAAG TAATAAGGGA GGGAAAATTT AATATTCAGT AAAATCGAAA AGATCTATAT 60373276

 -0500 + TTATTTTTAT TGTGTATTTA GACATAGTTT ATATCTAAGT GCATATCAAA AGCTACCCGT ATCTAGAAAA GCCAAAACGC CTTAGAATTT GCAATG GAGT -0401

      - AATAAAAATA ACACATAAAT CTGTATCAAA TATAGATTCA CGTATAGTTT TCGATGGGCA TAGATCTTTT CGGTTTTGCG GAATCTTAAA CGTTACCTCA  60373176

 -0400 +    TAGATTA ATTCTTCCAC TAGCTTGTCC GGTTTTGACA GTCTATATCA TGCAAGCAAT AATACAACAA ATTTATTT AT TTCAAA CAAA GTCCTCCTGC -0301

       - CTCATCTAAT TAAGAAGGTG ATCGAACAGG CCAAAACTGT CAGATATAGT ACGTTCGTTA TTATGTTGTT TAAATAAATA    AAGTTTGTTT CAGGAGGACG 60373076

 -0300 + AATGCAGGCG GGCTAGCTAT TCTTGTA ATA TT TAATTATT GTATCAGTCA AGACCAATGT CAGTAACCAA TAAGACT ACG TG GTAGTCCA AGCTAGTCAT -0201

       - TTACGTCCGC CCGATCGATA AGAACATTAT   AAATTAATAA CATAGTCAGT TCTGGTTACA GTCATTGGTT ATTCTGATGC ACCATCAGGT TCGATCAGTA 60372976

 -0200 + ATCCAGTGCG TGGCATTAGG TGCGTCAT  CA GTTG AAATGA TTGACAAAGC GCTGAGCGTA  CGTATAAATA ATTATAAA GG GCC  GGTTAAG CAAT TTAGTT -0101
       - TAGGTCACGC ACCGTAATCC ACGCAGTAGT CAACTTTACT AACTGTTTCG CGACTCGCAT  GCATATTTAT TAATATTT CC CGG CCAATTC GTTAAATCAA 60372876

 -0100 + GTCATTTAAC TTATAAAACA CTGACCAAGT TGGAGTTATA ATGCTCATAC GGTTA TATAA A CCAGTCTAA TAGATCGAGG    ACGTACGTTT ACAGTGCCTC -0001
       - CAGTAAATTG AATATTTTGT GACTGGTTCA ACCTCAATAT TACGAGTATG CCAATATATT  TGGTCAGATT ATCTAGCTCC   TGCATGCAAA TGTCACGGAG 603727

       ⇓Transcription start for Sobic.005G213600 (Chr05:60,372,775 on reverse strand) 5'UTR  1st exon
 +0001 + GCACAAAGTA CATGAACGTA CAGGAGAGGA GGAAGGAACT CGAAGAGAGA TCATCGACGA CATTGGG GTG   AGTTTTGGCA TGCAAATGCA AGCAATTCTT +0100
       - CGTGTTTCAT  GTACTTGCAT GTCCTCTCCT CCTTCCTTGA GCTTCTCTCT AGTAGCTGCT GTAACCCCAC TCAAAACCGT ACGTTTACGT TCGTTAAGAA  60372676
```

Motif	Position (relative to transcription start)	Strand (relative to gene orientation)	Sequence	Description
OSE2ROOTNODULE	-1697	+	CTCTT	root specificity element; associated with expression in mycorrhizal roots
P1BS	-1690	±	GNATATNC	PHR1 binding site; responsive to phosphorous deficiency
SORLIP5,2	-398, -122	+, ±	GAGTGAG GGGCC	motifs often found in promoters of light-responsive genes
ERELEE4	-322	–	ATTCAAA	ethylene-responsive element
ROOTMOTIFTAPOX1	-273	+	ATATT	element found in many promoters of genes with root specific expression
ABRELATERD1	-223	+	AAACAGA	ABA-responsive element
CGTCA-motif	-178	+	CGTCA	cis-acting MeJA-responsive element
MYBCORE	-172	+	CAGTTG	drought-responsive element
CURECORECR	-143, -18	±	GTAC	responsive to copper deficiency
CAAT-box	-110	+	CAAT	core promoter element near transcription start site
TATA-box	-45	+	TATAAA	core promoter element near transcription start site

Figure S3. Promoter sequence analysis indicating possible CAREs for *Sobic.005G213600* based on in silico analysis using the PLACE database search tool. Chromosome positions according to Phytozome *Sorghum bicolor* v2.1, DOE-JGI.

Supplementary tables

Table S1. Major strigolactones in the root exudates of the 24 RILs derived from the cross of SRN39 and Shanqui-Red used to fine-map the *LGS1* region

		Relative abundance in root exudates, % measured strigolactones	
RIL ID	*Striga* germination stimulant activity	5-deoxystrigol	Orobanchol
SSD#3–046	High	98	2
SSD#3–047	High	98	2
SSD#3–080	High	99	1
SSD#3–113	High	100	0
SSD#3–148	High	97	3
SSD#3–246	High	100	0
SSD#3–282	High	100	0
SSD#3–521	High	98	2
SSD#3–660	High	95	5
SSD#3–816	High	99	1
SSD#3–832	High	90	10
SSD#3–177	High	100	0
SSD#3–767	High	100	0
SSD#3–061	Low	0	100
SSD#3–070	Low	0	100
SSD#3–128	Low	0	100
SSD#3–158	Low	5	95
SSD#3–162	Low	0	100
SSD#3–307	Low	0	100
SSD#3–535	Low	0	100
SSD#3–683	Low	1	99
SSD#3–809	Low	0	100
SSD#3–302	Low	1	99
SSD#3–320	Low	1	99

Values are average of three measures.

Table S2. PCR primers used to identify *LGS1* among the candidate genes on sorghum chromosome 5 and the qRT-PCR primers used to monitor its expression

	Primer sequences (5′ → 3′)				
Primer name	Forward	Reverse	Target position in reference genome*		Amplicon size
G2133	TCAGGGAGTGCAGGAGAATC	CCGCATACTTATGATCAGACCTC	60,340,583	60,340,934	351
G2134	GCTTCACAATCCCAGGTGTT	CTGTACCACACGGGCAATAA	60,346,441	60,346,699	258
PDstrigalgs4c	GACAGGCTCCATCTCATGGT	GGGAACTGAACAAAGGCCTAA	60,366,129	60,366,722	593
PDstrigalgs5b	CAAACCCATCGGACATCTTC	CAGCATGTCCTCGTACCTGA	60,371,620	60,372,248	628
PDstrigalgs6b	CTCTTCGGCGACGGGTACTA	TGCAGCAGTACGTCTCAGAACT	60,376,824	60,377,456	632
G2138	AGATGGATCTCGCTTGCCTA	GGCAGTTGCTTCTGGAACTC	60,393,288	60,393,558	270
G2139	ACTTCACGGAGGAGCCTGTA	AGCATGAGAGGCAAAAGCAT	60,399,297	60,399,663	366
5bF + 2/3E3R	CAAACCCATCGGACATCTTC	CACGGATATGGAGTCGGCAA	60,371,318	60,372,248	930
qRT-PCR	TTCCATAGACGACAGCGAGC	AAAGCACTTCTTTGTGGAATAAAGG	60,371,004	60,371,160	156

*Phytozome, Sorghum bicolor v2.1 DOE-JGI.

Table S3. Summary of deletion variants in the LGS1 region

Sorghum accession	Allele name	Deletion position* Chr05:	*Sobic.005G213300*† No annotated functional domains	*Sobic.005G213400* No annotated functional domains	*Sobic.005G213500* 2OG-Fe(II) oxygenase superfamily; nonhaem dioxygenase in morphine synthesis N terminal	*Sobic.005G213600* Weakly similar to *Os09g0555100* protein (sulfotransferase family)	*Sobic.005G213700* Similar to epoxide hydrolase (alpha/beta hydrolase fold)	*Sobic.005G213766* No annotated functional domains	*Sobic.005G213832* UDP-glucose–glucose-phosphate glucosyltransferase	*Sobic.005G213900* Gamma-thionin family
SRN39	lgs1-1	69,977,147..70,011,172	Intact	Intact	Deleted	Deleted	Deleted	Deleted	Deleted	Intact
555	lgs1-2	69,958,403..69,986,951	Intact	Deleted	Deleted	Deleted	Intact	Intact	Intact	Intact
IS7777	lgs1-3	69,981,523..70,011,172	Intact	Intact	Intact	Deleted	Deleted	Deleted	Deleted	Intact
SC103	lgs1-4	69,985,178..69,985,598	Intact	Intact	Intact	421-bp deletion; premature stop	Intact	Intact	Intact	Intact
Tetron	lgs1-5	69,985,617..69,985,626	Intact	Intact	Intact	10-bp deletion; premature stop	Intact	Intact	Intact	Intact

*Phytozome, *Sorghum bicolor* v3.1 DOE-JGI.
†Gene model and annotation.

Chapter 3

Characterization of a sorghum sulfotransferase and its role in diversification of strigolactones

Mahdere Z. Shimels[1], Stefano Rendine[2], Patrick J. Rich[3], Carolien Ruyter-Spira[1], Gebisa Ejeta[3], Harro Bouwmeester[1,4]

[1]Laboratory of Plant Physiology, Wageningen UR, PO Box 658, 6700 AR Wageningen, the Netherlands
[2]Syngenta Crop Protection AG Crop Protection Research Stein, Switzerland
[3]Department of Agronomy, Purdue University, West Lafayette, IN 47907
[4]Present address: Plant Hormone Biology, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands.

Abstract

Sorghum (*Sorghum bicolor*) is among the many African cereals plagued by the root parasitic weed *Striga* (*Striga hermonthica*). Different conventional agronomic practices have been employed to control this parasite. However, these practices have not been effective due to *Striga*'s complex interaction with the host plant and its ability to produce enormous amounts of seeds, which can remain dormant in the soil for many years. Host plant resistance mechanisms towards this root parasite include physical and chemical barriers to parasite establishment, hypersensitivity and reduced *Striga* seed germination stimulant activity. Some *Striga* resistant sorghum genotypes with the latter mechanism were shown to be the result of a change in the type of strigolactone exuded into the rhizosphere (Gobena et al. 2017). Resistant genotypes have a mutation in a sulfotransferase (SbSOT4A) coding gene (*Sobic.005G213600*; *LGS1*) and exude orobanchol, while 5-deoxystrigol is the major strigolactone exuded by the susceptible cultivars. Aside from a hydroxyl group, these strigolactones differ in their stereochemistry with respect to the chiral carbons shared between the B- and C-rings. In the present study, we show that SbSOT4A is localized in the cytosol, suggesting it sulfates small molecules, including hormones like strigolactones. Furthermore, with protein modeling and substrate docking, we show that SbSOT4A has a high affinity to C18-hydroxycarlactone and – based on that – we propose a new model for the role of SbSOT4A – sulfation of 19-hydroxycarlactone. The C18 of this sulfated 18-hydroxycarlactone is oxidized to form a carboxy group, likely by one of the MAX1 homologs, which favors ring closure with the loss of the sulfate group to the β-orientation resulting in 5-deoxystrigol. Loss of SbSOT4A function, and thereby lack of a sulfated intermediate, favors oxidation of the C18 hydroxyl to a carbonyl and ring closure to an α-orientation resulting in orobanchol, a much less effective *Striga* germination stimulant.

Introduction

Plants are constantly exposed to different biotic and abiotic stresses such as drought, extreme temperatures, low nutrient availability, salinity, diseases and pests. Being immobile, they developed different intricate mechanisms to survive these harsh environments. Plant hormones produced under these conditions play a key role in these stress responses through complex interactions and mechanisms. This phenomenon is nicely illustrated by the strigolactones, a group of plant hormones and signaling molecules. The strigolactones are produced in roots and play a key role in regulating plant architecture in response to environmental constraints. Their production is boosted during nutrient limiting conditions - especially phosphate shortage - and they are also exuded into the root environment to serve as signaling molecules. Phosphate is a macronutrient that is essential for the growth and development of plants. Under phosphate shortage, plants respond to maintain several metabolic processes and regulatory pathways such as photosynthesis, respiration and phosphorylation reactions (Czarnecki et al. 2013; Plaxton et al. 2011; Marzec et al. 2013). The secretion by plants of strigolactones into the rhizosphere stimulates the mutualistic interaction with the plant by beneficial microorganisms such as arbuscular mycorrhizal (AM) fungi. In this interaction, the AM fungi solubilize phosphate and facilitate its uptake as well as that of other minerals by the plant (Bouwmeester et al. 2007; Lopez-Raez et al. 2008). Plants with AM fungi can uptake two-fold more phosphate than plants without symbiosis with AM fungi (Sorefan et al. 2003). In return, the host plant provides carbon from photosynthesis to the AM fungi.

The composition of strigolactones in roots and root exudates differs between plant species and sometimes also between different cultivars of the same species, both quantitatively and qualitatively (Gobena et al. 2017; Xie et al. 2010; Kapulnik et al. 2014; Wang et al. 2018). All the known strigolactones are believed to be derived from all-*trans*-β-carotene. Structurally, strigolactones are divided into two groups: the canonical and the non-canonical strigolactones, with the former having a common tricyclic ABC-ring structure linked to the D-ring through an enol ether bridge. The non-canonical strigolactones lack the conventional ABC ring structure but do have the D-ring. Zealactone, avenaol, heliolactone, carlactone, carlactonoic acid and methyl carlactonate are examples of non-canonical strigolactones (Figure 1. 1-6) (Abe et al. 2014; Nomura et al. 2018; Seto et al. 2014; Wang et al. 2018). So far, 23 canonical strigolactones have been identified from plant root exudates (Figure 1.7-21) (Wang et al. 2018; Nomura et al. 2018). The stereochemistry of the canonical strigolactones differs based on the orientation of the C-ring which has either a β- or α-orientation, which classifies them as strigol- and orobanchol-type strigolactones, respectively (Xie et al. 2013). The biosynthesis of both canonical and non-canonical strigolactones has been shown to start from the common strigolactone precursor, carlactone. In many of the plant species investigated carlactone is first converted to carlactonoic acid which has been detected in root exudates of several plant species (Abe et al. 2014; Seto et al. 2014; Yoneyama et al. 2018). Subsequently, in some plant species, it was shown that carlactonoic acid is methylated to form

methyl carlactonoate which is further hydroxylated to an as yet unidentified product by Lateral Branching Oxidoreductase (LBO) (Brewer et al. 2016; Abe et al. 2014). For example, in sunflower carlactonoic acid is likely converted to methyl carlactonoate and then to heliolactone (Takikawa et al. 2017). *In vitro*, it has been shown that carlactonoic acid can also be converted to 18-hydroxy-carlactonoic acid (Yoneyama et al. 2018). In rice, the MAX1 homolog, Os900, was shown to catalyze both the formation of carlactonoic acid and C18 hydroxylation, which was postulated to result in spontaneous ring closure and the formation of 4-deoxyorobanchol (Zhang et al. 2014).

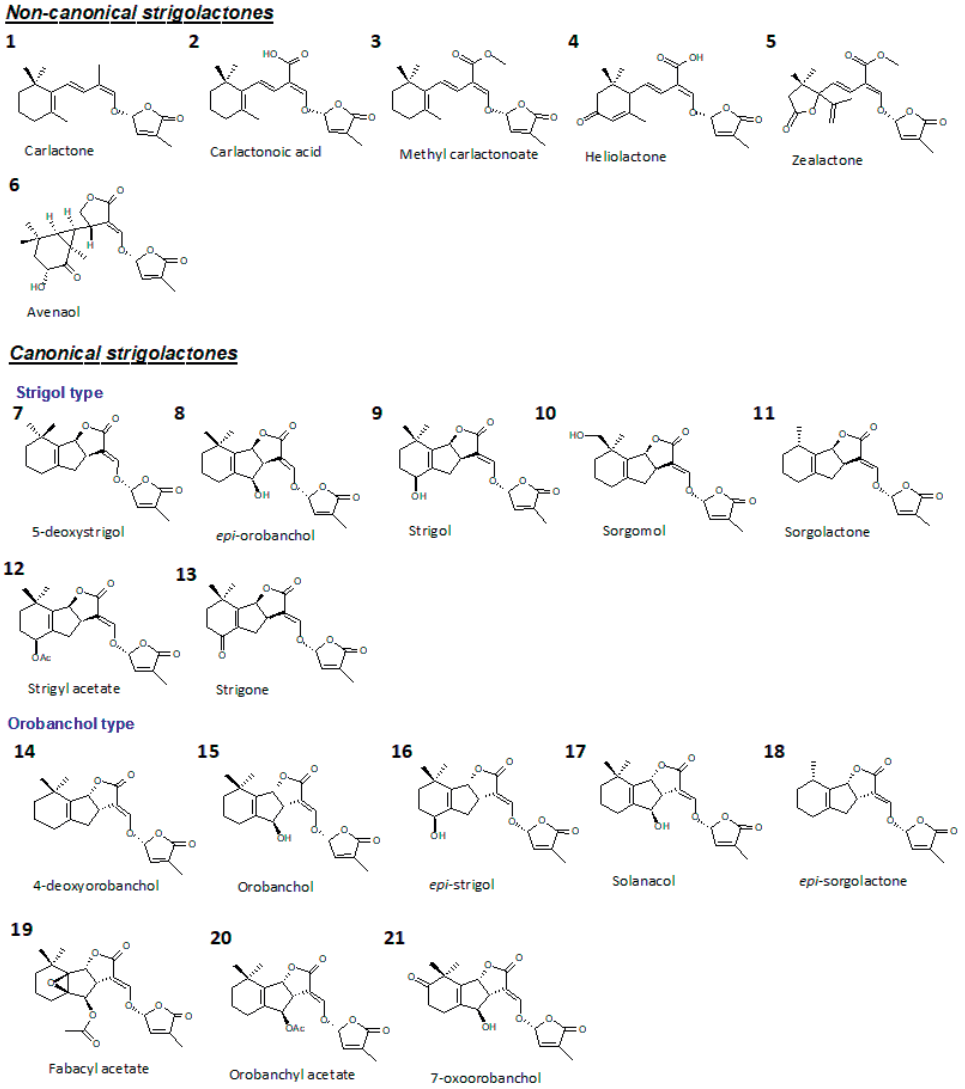


Figure 1. Examples of canonical and non-canonical strigolactones showing the C-ring in either β- (strigol-type) or α- (orobanchol-type) orientation.

The formation of 5-deoxystrigol has not been elucidated; it seems that this is not a MAX1 catalyzed enzymatic reaction. The production of strigol- and orobanchol-type strigolactones from the precursors 5-deoxystrigol and 4-deoxyorobanchol, respectively, proceeds by modifications on the ABC rings to produce different canonical strigolactones (Flematti et al. 2016). One example is the formation of orobanchol from 4-deoxyorobanchol which is catalyzed by the MAX1 homolog, Os1400.

Besides their beneficial roles, strigolactones are used as host detection signaling molecules by root parasitic plants of the Orobanchaceae. Only upon the recognition of a host plant, through the detection of a strigolactone, the seeds of these parasitic plants germinate. The close presence of a host ensures survival of the parasite after germination as it needs to attach to a host within days after germination. Upon contact with the host root, the parasite will initiate a haustorium to be able to penetrate the host root and acquire water, assimilates and nutrients (Cui et al. 2016; Hibberd et al. 2001). To avoid infection, plants have developed pre- and/or post-germination defense mechanisms. For instance, some resistant cultivars develop a hypersensitive response, a programmed cell death at the point of attachment to limit the growth of the parasite (Mohamed et al. 2003). Another form of resistance is the low germination stimulant activity (Gobena et al. 2017; Satish et al. 2012; Rich et al. 2004). Low germination stimulant activity in sorghum was first discovered in vitro in the resistant cultivar IS9830 (Hess et al. 1992). Later, a sorghum mapping population was screened using an agar gel assay, which resulted in the identification of a quantitative trait locus (QTL) for low-germination stimulant activity. The locus resides on chromosome 5 and initially encompassed 30 candidate genes (Satish et al. 2012). In Chapter 2, I describe the identification of a sulfotransferase (SbSOT4A) encoded by *Sobic.005G213600* which we named *LGS1* to be the gene underlying the QTL (Gobena et al. 2017). Several different mutations involving this gene were subsequently discovered among *Striga* resistant sorghum cultivars with low germination stimulant activity. The predominant strigolactone exuded from the roots of these *lgs1* mutants is orobanchol, the C-ring of which is in α-orientation, while high germination stimulant, *Striga* susceptible, sorghum lines exude mainly 5-deoxystrigol, with a β-oriented C-ring.

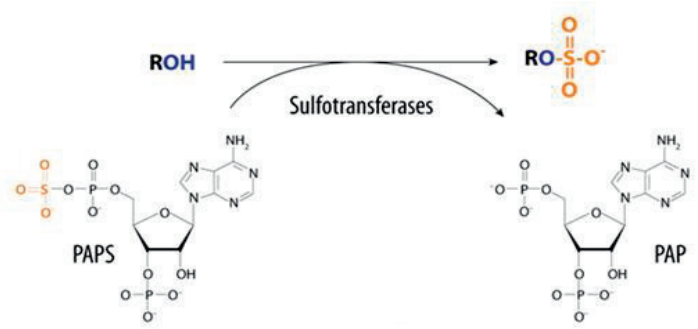


Figure 2. Schematic drawing of sulfation reaction catalyzed by sulfotransferases (Prather et al. 2011).

Sulfotransferases (SOTs) are a group of enzymes present in a wide range of organisms including plants. They catalyze the transfer of a sulfuryl group (-SO3) from a donor, 3`-phosphoadenosine-5`-phosphosulfate (PAPS), to an alcohol or amine acceptor. As a result of this reaction a sulfate conjugate, sulfate ester or sulfamate and 3`-phosphoadenosine-5`-phosphate (PAP) are formed (Figure 2) (Hirschmann et al. 2017).

Sulfotransferases are classified as soluble/cytosolic or membrane-associated (Chapman et al. 2004). The name cytosolic does not necessarily refer to their sub-cellular localization, rather to the ability of the enzyme to be extracted from the cell in a soluble form (Suiko et al. 2017; Hernàndez-Sebastiá et al. 2008). Cytosolic SOTs preferably sulfate small molecules such as hormones while the membrane-associated SOTs are mainly involved in the post-translational modification of macromolecules, such as carbohydrates, proteins, proteoglycans and glycolipids (Chapman et al. 2004). In general, the substrates for the membrane-associated SOTs are highly specific while the cytosolic SOTs can often sulfate a wide range of substrates (Hirschmann et al. 2014). Based on the acceptor group, SOTs can be classified into three groups namely; O, N and S sulfotransferases, of which catalytic activity results in the production of an ester, amide and thioester, respectively. Sulfotransferase family proteins contain four highly conserved regions with critical residues for PAPS binding (Komatsu et al. 1994; Driscoll et al. 1995; Marsolais et al. 1995). Region I, located on the N-terminal of SOTs, contains a 5`-phosphosulfate-binding loop (PSB). Region II contains a 3`-phosphate binding loop (PB) and starts with a highly conserved catalytic histidine. This region is important for proton acceptance during sulfuryl transfer (Varin et al. 1997). Region II and III contain important residues that form a parallel stack with the adenine group of PAPS (Hernàndez-Sebastiá et al. 2008). Region IV, located at the C-terminal of SOTs, contains GxxGxxK and KxxxTVxxxE motifs. KxxxTVxxxE is important for dimerization (Petrotchenko et al. 2001; Komatsu et al. 1994). In plants, several compound classes, such as brassinosteroids, coumarins, flavonoids, gibberellic acids, jasmonates, salicylic acid, glucosinolates, phenolic acids and terpenoids are subjected to sulfation by cytosolic SOTs (Gidda et al. 2003; Hirschmann et al. 2014; Marsolais et al. 2000). These modifications are important in detoxification/inactivation of metabolites but also as a requirement for their biological activity (Rouleau et al. 1999; Matsubayashi et al. 1996). Most sulfated products are water-soluble, which could aid the movement of these molecules (Kester et al. 1999; Visser et al. 1998; Coughtrie 1996). Although several plant SOTs have recently been characterized, little is known about them in terms of their structure, substrate preference and physiological role. From the 21 SOTs of *Arabidopsis*, substrate preference and enzymatic activity have been identified for only ten (Hirschmann et al. 2014). For instance, AtSOT12 and AtSOT13 cluster together in the phylogenetic tree but sulfate brassinosteroids and flavonoids, respectively.

In sorghum, several putative SOTs have been annotated. In Chapter 2, we showed that a gene annotated as a SOT, which we call SbSOT4A, has been linked to stereo-controlled biosynthesis of strigolactones. Furthermore, we showed that SbSOT4A is intact and functional in 5-deoxystrigol exuding LGS1 sorghum with high *Striga* germination stimulant activity (Gobena et al. 2017). The aim of this study was to further characterize this enzyme and propose a

mechanism by which stereo-specific ring closure of its substrate leads to either strigol- or orobanchol-type strigolactones. Here, we determined the subcellular localization of SbSOT4A using both experimental and bioinformatic approaches. Furthermore, using protein modeling and substrate docking, we predicted which substrate the protein may sulfate and what therefore its possible role is in determining *Striga* germination stimulant activity in sorghum.

Results

Striga resistance and strigolactone profile in root exudates of sorghum

Both orobanchol- and strigol-type strigolactones were detected in the root exudate of all sorghum genotypes tested (Figure 3). In Chapter 2, we reported that the major strigolactone produced and exuded by cultivars with low *Striga* germination stimulant activity is orobanchol, an orobanchol-type strigolactone, while susceptible cultivars produced 5-deoxystrigol, a strigol-type strigolactone (Gobena et al. 2017). Five representative sorghum genotypes from both categories were selected and their root exudate strigolactone profiles analyzed. As shown in the heatmap, the low and high *Striga* germination stimulant sorghum lines clustered separately in two major clades (Figure 3). The 1st clade is composed of the high stimulant lines Shanqui-Red, SSD#3-767, P9408, N13 and SSD#3-177 while the second clade is composed of low stimulant lines SRN-39, 555, SSD#3-320, P9401 and SSD#3-302. The 1st clade is characterized by higher levels of 5-deoxystrigol and sorgomol, both strigol-type strigolactones, whereas the 2nd clade is characterized by higher levels of the rest of the strigolactones detected. Within the 2nd clade, orobanchol and strigol together separated from the other strigolactones and as expected for orobanchol, were found to be more abundant in low stimulant sorghum lines than in those with high *Striga* germination stimulant activity. Strigol was also more abundant in root exudates of low stimulant lines than in high stimulant lines. Orobanchol-type strigolactones such as *ent-2'-epi*-strigol and *ent-2'-epi*-sorgolactone clustered together. The resistant genotypes had more of these strigolactones compared to the susceptible genotypes.

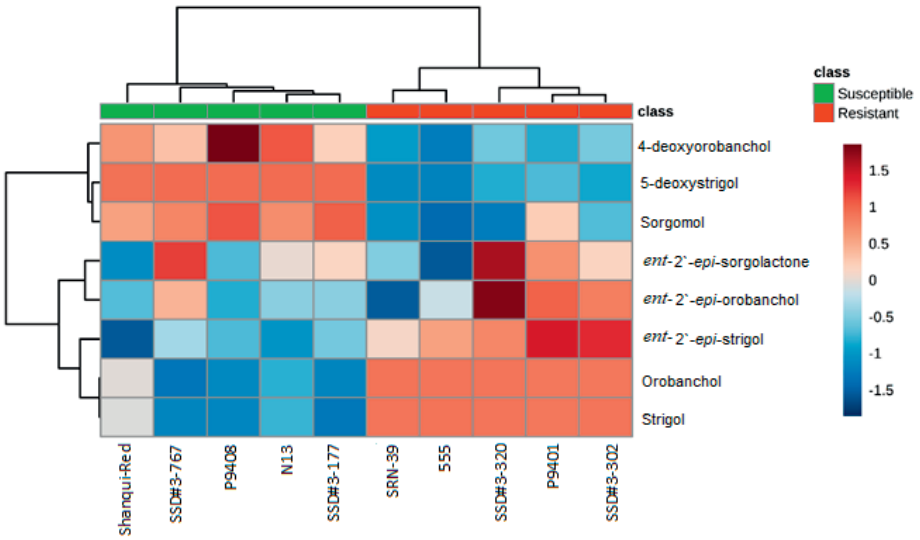


Figure 3. Heat map showing the strigolactone profile in sorghum genotypes with SbSOT4A intact (SSD#3-177, SSD#3-250 and Shanqui-Red) and mutant genotypes (SRN-39, SSD#3-320, SSD#3-302). Data was analyzed using MetaboAnalyst (https://www.metaboanalyst.ca/)with average of three biological replicates.

To confirm the stereochemistry of the orobanchol detected, the samples and the standards of the 4 stereoisomers of orobanchol were injected on an enantio-selective column in LC-MS/MS. The four orobanchol stereo-isomers - *ent*-orobanchol, orobanchol, *ent*-2`-*epi*-orobanchol and 2`-*epi*-orobanchol - eluted at 19.53, 26.34, 20.71 and 21.88 min, respectively (Figure 4 A and B). The chromatograms clearly show that the orobanchol that was exuded by sorghum elutes at a retention time of 26.22 min, which confirms that it is natural orobanchol with C2` R configuration (Figure 4C).

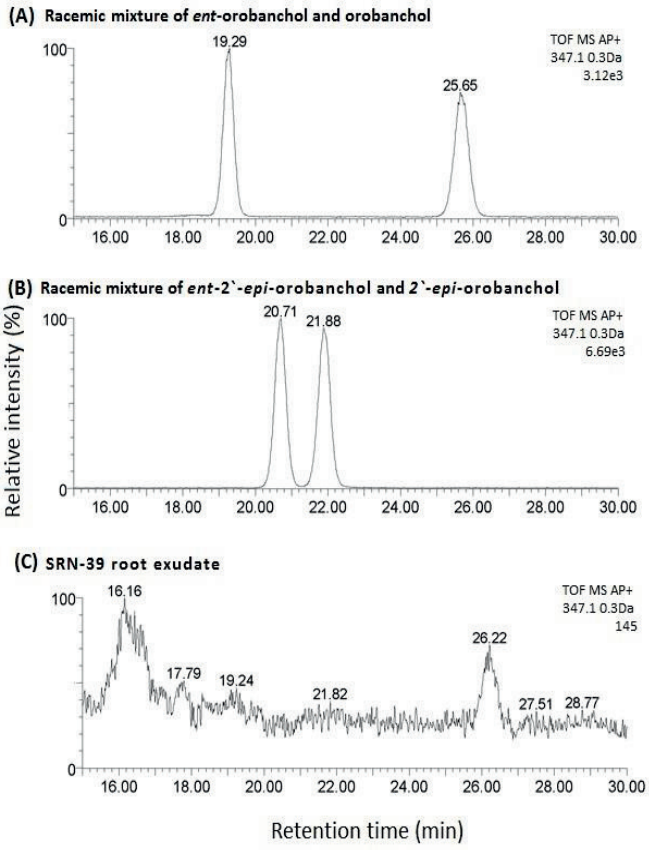


Figure 4. Chiral-LC-MS/MS analysis of orobanchol from root exudates of SRN39, a *Striga* resistant sorghum genotype. (A) Racemic mixture of *ent*-orobanchol (retention time: 19.53 min) and, orobanchol (retention time: 19.53 min). (B) Racemic mixture of *ent*-2`-*epi*-orobanchol (retention time: 20.71 min) and 2`-*epi*-orobanchol (retention time: 21.88 min). (C) SRN39 root exudate. SRN39 was grown for 5 weeks with one week of phosphate starvation. The root exudates were collected and strigolactones were extracted.

Identification of the orobanchol precursor in sorghum

After determining the stereochemistry of orobanchol in the root exudate of SRN39, we carried out a feeding assay to find out what the precursor is of its orobanchol. Sorghum seedlings were grown in hydroponics for two weeks and were fed with equal amounts of three possible precursors (carlactone, 5-deoxystrigol and 4-oxyorobanchol) with and without the presence of fluridone an inhibitor of carotenoid and therefore strigolactone biosynthesis (Matusova et al. 2005). As shown in Figure 5A, compared to the plants grown in water, the application of fluridone significantly reduced the production of orobanchol as anticipated. The feeding of carlactone and 5-deoxystrigol in combination with the fluridone treatment did not restore the production of orobanchol. However, the level of orobanchol was ~4 fold higher when the plants were fed with 4-deoxyorobanchol compared to those treated with only fluridone. This

result suggests that 4-deoxyorobanchol is a precursor for orobanchol in SRN39. Interestingly, the level of orobanchol in root exudates of SRN39 was much lower than we usually find. Therefore, a follow-up experiment was conducted in which plants were grown in hydroponics and sand. As shown in Figure 5B, the level of orobanchol was significantly higher when plants were grown in sand than hydroponics. This makes the feeding results that were obtained with plants grown in hydroponics less reliable.

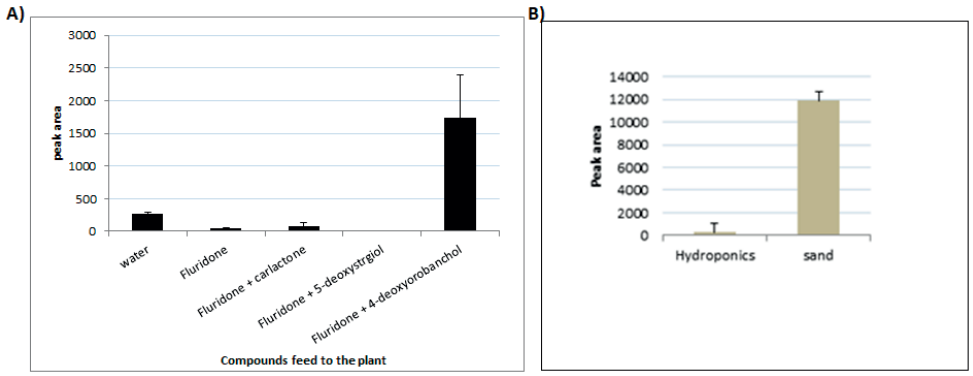


Figure 5. A) Orobanchol profile in feeding assay that were fed with fluridone, fluridone combined with carlactone or 5-deoxystrigol or 4-deoxyorobanchol. B) Orobanchol profile in root exudates of SRN39 when grown in hydroponics and in sand. The bar represents standard error n=3.

Triclosan, a SOT inhibitor, affects strigolactone production in Shanqui-Red

In previous studies, we have shown that SbSOT4A is intact and in sorghum lines with high *Striga* germination stimulant activity (Chapter 2). In the present experiment, SRN39 and Shanqui-Red were grown in sand and treated with Triclosan, a SOT inhibitor and monitored changes in the strigolactone composition. As shown in Figure 6A and C, both orobanchol and 5-deoxystrigol levels in root exudates of Shanqui-Red increased significantly by triclosan treatment. On the other hand, triclosan treated SRN39 seedlings had the same level of both orobanchol and 5-deoxystrigol when compared with the control plants (Figure 6B and D). SRN39 does not have SbSOT4A because of the mutation on the *LGS1* gene. Therefore, the application of the SOT inhibitor is not expected to change the strigolactone profile in SRN39. In Shanqui-Red application of the SOT inhibitor resulted in increased orobanchol production which would agree with the role of SbSOT4A in preventing orobanchol formation. The production of 5-deoxystrigol that we expected to decrease as a result of the increase in orobanchol however was also slightly higher.

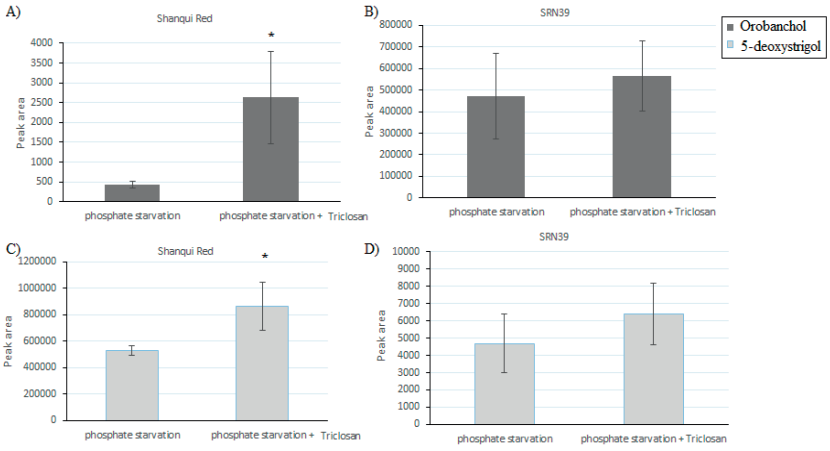


Figure 6. Effect of sulfotransferase inhibitor, triclosan, on strigolactone production by sorghum genotypes Shanqui-Red and SRN39. A) Orobanchol profile of root exudates produced by Shanqui-Red treated with and without triclosan. B) Orobanchol profile of root exudates produced by SRN39 treated with and without triclosan, C) 5-deoxystrigol profile of root exudates produced by Shanqui-Red treated with and without triclosan, and D) 5-deoxystrigol profile of root exudates produced by SRN39 treated with and without triclosan. The bars represent the mean of three replicates +/- the standard error.

Phylogeny of plant SOTs

To be able to say more about the possible function of SbSOT4A, amino acid sequences of all known SOTs from *Arabidopsis* and rice were used to perform BLASTp search in several online databases such as Phytozome and NCBI. 25 Non-redundant sorghum AA sequences with a known SOT domain were retrieved (Table 1). These amino acid sequences were aligned with SOTs from *Arabidopsis*, *Brassica napus*, maize, rice and other species with characterized functions. Only those genes containing one or more of the conserved regions I-IV were used for this analysis. Genes encoding SOTs were used to make a phylogenetic tree using IQ-TREE (http://www.iqtree.org). The *Arabidopsis* SOTs mostly clustered together. For instance, the tyrosylprotein sulfotransferase (AtTPST) clustered together with *Sobic_010G045966*. However, three *Arabidopsis* desulfoglucosinolate sulfotransferases clustered together with flavonol sulfotransferases from *Flaveria chloraefolia* and AtSOT12 were in the same clade with brassinosteroid sulfotransferases from *Brassica napus* (Figure 7). Several clusters that were identified in other studies are also present in our tree. For instance, the four brassinosterioid SOTs from *Brassica napus* and the three flavonol SOTs from *Flaveria chloraefolia* and *Flaveria bidentis* (Figure 7). In Figure 7, the SOTs with identified substrates are indicated, but the SOTs that clustered with SbSOT4A have not functionally been characterized and their substrate is hence unknown.

Table 1. Genes annotated as SOT in *sorghum bicolor*. The data was obtained from phytozome (https://phytozome.jgi.doe.gov/pz/portal.html)

Gene ID	Gene name	PFAM	Annotation	METACYC	Protein length (amino acid)	cDNA length (bp)	Molecular weight (kDa)	No. of introns
Sobic.005G213600	*Sb05g026550 (LGS1)*	PF13469	Sulfotransfer_3	Flavonol 3-sulfotransferase	452	1359	48.50	1
Sobic.005G166700	*Sb05g022710*	PF13469	Sulfotransfer_3	-	420	1263	46.23	5
Sobic.005G217600	*Sb05g026855*	PF13469	Sulfotransfer_3	-	336	1011	37.22	1
Sobic.005G183100	*Sb05g024150*	PF13469	Sulfotransfer_3	Flavonol 3-sulfotransferase	349	1050	38.61	1
Sobic.003G424100		PF13469	Sulfotransfer_3	Flavonol 3-sulfotransferase	341	1026	38.74	1
Sobic.004G046200	*Sb04g003960*	PF13469	Sulfotransfer_3	-	368	1170	40.44	0
Sobic.003G424000		PF13469	Sulfotransfer_3	Flavonol 3-sulfotransferase	322	969	35.83	2
Sobic.004G046100	*Sb04g003950*	PF13469	Sulfotransfer_3	Flavonol 3-sulfotransferase	368	1107	40.10	0
Sobic.002G369400	*Sb02g038840*	PF13469	Sulfotransfer_3	Flavonol 3-sulfotransferase	341	1026	37.43	0
Sobic.002G369500	*Sb02g038850*	PF13469	Sulfotransfer_3		356	1071	38.57	0
Sobic.006G154200	*Sb06g023150*	PF13469	Sulfotransfer_3	CYTOSOLIC SULFOTRANSFERASE 12	347	1044	38.81	
Sobic.004G280000	*Sb04g031660*	PF13469	Sulfotransfer_3	hydroxyjasmonate sulfotransferase	336	1011	37.63	0
Sobic.007G206500		PF00685	Sulfotransfer_1	Flavonol 3-sulfotransferase	425	1278	47.54	2
Sobic.003G437200	*Sb03g046790*	PF00685	Sulfotransfer_1	Desulfoglucosinolate sulfotransferase	335	1008	38.11	0
Sobic.003G437300		PF00685	Sulfotransfer_1		334	1005	37.36	0
Sobic.003G437100	*Sb03g046780*	PF00685	Sulfotransfer_1	Desulfoglucosinolate sulfotransferase	328	987	37.67	0
Sobic.001G136400	*Sb01g011900*	PF13469	Sulfotransfer_3	Desulfoglucosinolate sulfotransferase	345	1038	38.23	0
Sobic.003G437001		PF00685	Sulfotransfer_1	Flavonol 3-sulfotransferase	296	891	33.87	3
Sobic.006G057500		PF00685	Sulfotransfer_1	Flavonol 3-sulfotransferase	185	558	20.74	0
Sobic.001G348000	*Sb01g033350*				364	1095	41.50	5
Sobic.001G362100	*Sb01g034575*	PF00685	Sulfotransfer_1		339	1020	39.36	6
Sobic.002G395900	*Sb02g041100*	PF00685	Sulfotransfer_1		335	1008	38.75	19
Sobic.007G212500	*Sb07g027970*				357	1074	41.01	10
Sobic.010G045833				Tyrosylprotein sulfotransferase	222	669	24.09	2
Sobic.010G045966				Tyrosylprotein sulfotransferase	510	1533	57.94	11

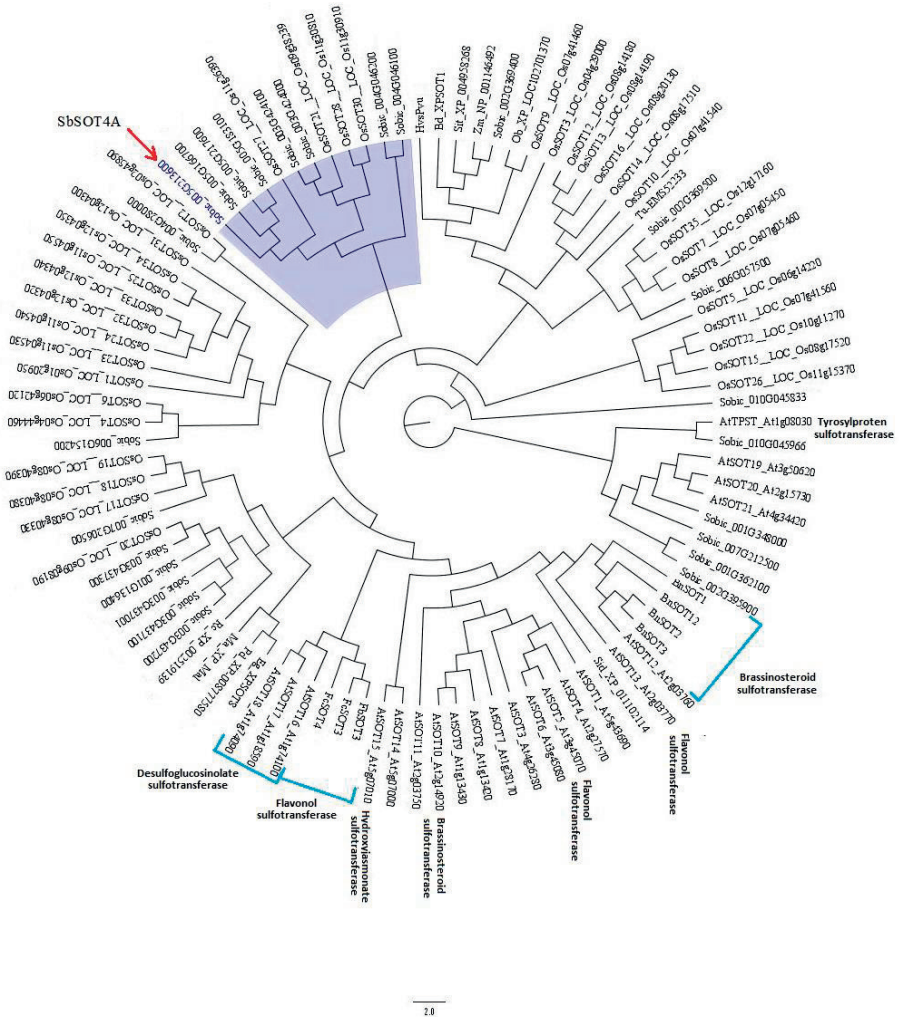


Figure 7. Phylogenetic tree of SOT gene family in plants. Amino acid sequence of gene products annotated as sulfotransferase from *Arabidopsis thaliana*, *Brassica napus*, *Zea maize*, *Oryza sativa*, *Flaveria chloraefolia*, *Flaveria bidentis* and other plant species were obtained from Phytozome and NCBI. Clade where SbSOT4A clustered is shaded in purple. The alignment of these sequences was conducted with MEGA 5 and the phylogenetic tree was constructed using IQ-TREE.

We then had a close look at the sequence homology of SbSOT4A and other SOTs in an attempt to further identify its biological function. SbSOT4A consists of 452 amino acids with sulphotransfer_1 domain (PF00685). It exhibits 48% identity to the gene product of sorghum *Sobic.004G046200* (*Sb04g003960*) which is annotated as brassinosteroid SOT. Close homologs

were also found with gene products from rice (*Os09g0555150*) and *Arabidopsis* (*At3g45070*; annotated as flavonoid SOTs) with 50 and 36% identity, respectively. Looking at the amino acid alignment of SbSOT4A with selected SOTs from *Arabidopsis thaliana*, *Flaveria chlorifolia*, *Flaveria bidentis* and *Brassica napus*, the four highly conserved regions throughout the SOT family could be identified (Figure 8). Furthermore, the GxxGxxK motif in Region IV was present (Figure 8, underlined). However, motifs such as RDP in Region IV, and the KxxxTVxxxE (KTVE) motif were not present in SbSOT4A.

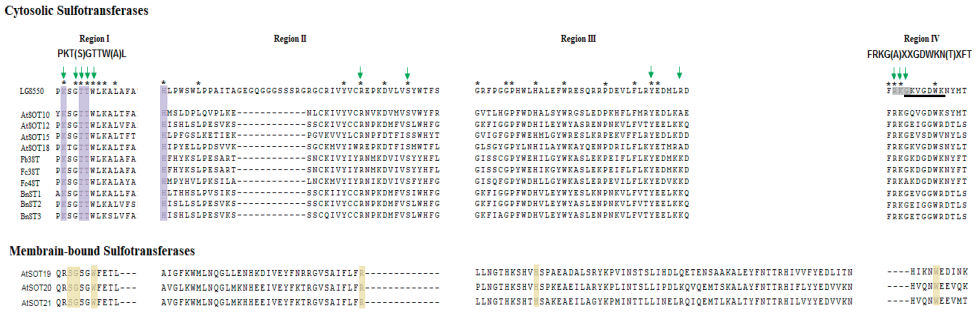


Figure 8. Partial amino acid sequence alignment of SbSOT4A with functionally characterized or annotated plant SOTs. AtSOT10 (product of *At2g14920*, *Arabidopsis thaliana*), AtSOT12 (from *At2g03760*, *Arabidopsis thaliana*), AtSOT15 (*At5g07010*, *Arabidopsis thaliana*), AtSOT18 (*At1g74090*, *Arabidopsis thaliana*), Fb3ST, from *Flaveria bidentis*), Fc3ST and Fc4ST from Flaveria chlorifolia, BnST1, BnST2 and BnST3 (the respective products of *AF000305*, *AF000306* and *AF000307* from *Brassica napus*), AtSOT19, AtSOT20 and AtSOT21 (from *At3g50620*, *At2g15730* and *At4g34420* of *Arabidopsis thaliana*), (The four known conserved regions (I-IV) are indicated on top of the alignment. Residues that are also conserved with membrane-associated SOTs are shaded in gray. The PAPS binding residues are indicated with green arrows, the catalytic residues are indicated with purple highlighting.

Modeling of sulfotransferases

Modeling of SbSOT4A, AtSOT18 and AtSOT19 was conducted using Phyre2 software (Kelley et al. 2015). A single template was selected by the software to be used for the modeling based on the highest confidence score for it to be the closest homolog. For the modeling of SbSOT4A, AtSOT18 from *Arabidopsis thaliana* (c5mekA) was used as a template. For the modeling of AtSOT19, the crystal structure of SOT stf1 from *Mycobacterium tuberculosis* h37rv (c2zq5A) was used as a template. The model quality for all the three models was with 100% confidence for the residues that were covered by the modeling. 322 residues (69% of the input sequence), 322 residues (92% of the input sequence) and 233 residues (69% of the input sequence) were modeled for SbSOT4A, AtSOT19 and AtSOT18 respectively. The result showed that AtSOT18 and SbSOT4A have four central β-strands whereas all the five β-strands of AtSOT19 are centered and surrounded by α-helices (Figure 9A, B and C). These results suggest that SbSOT4A is more similar to AtSOT18, which is a cytosolic SOT, than to AtSOT19, which is a membrane-associated SOT.

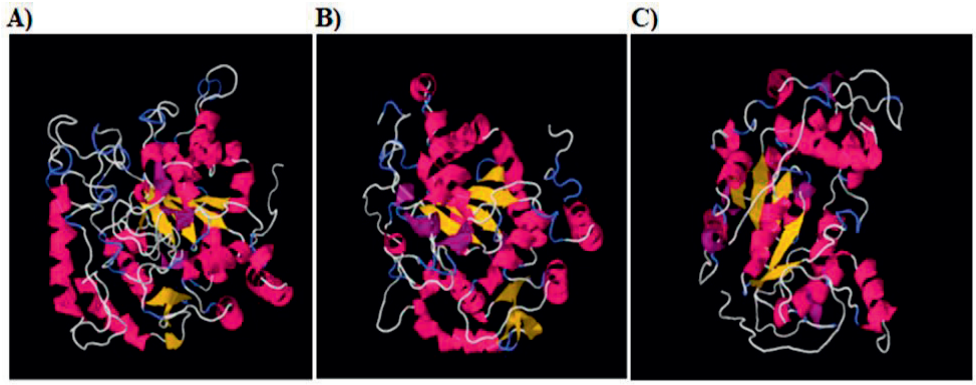


Figure 9 A) Ribbon diagram of LGS550. β-strands and α-helix are indicated in yellow and pink respectively B) Ribbon diagram of AtSOT18, a cytosolic SOT from Arabidopsis thaliana. C) Ribbon diagram of AtSOT19, a predicted membrane-associated SOT from Arabidopsis thaliana. β-strands and α-helix are indicated in yellow and pink color respectively.

SbSOT4A is localized in the cytosol and nucleus

To confirm the predicted subcellular localization of SbSOT4A, we transiently expressed it as a fusion with green fluorescent protein (GFP) at both the N and C terminus in *Nicotiana benthamiana*. Six days after infiltration, *N. benthamiana* leaves were harvested and analyzed by confocal laser scanning microscopy (CLSM). As shown in Figure 11, SbSOT4A-GFP accumulated in the cytoplasm and the nucleus of the epidermis of *N. benthamiana* leaves. The red color indicates auto-fluorescence of the chlorophyll. The green fluorescence in the nucleus was observed throughout all Z-layers. GFP tagging of either the C- or N- terminus did not result in a difference in localization of the protein (Figure 10 A- D).

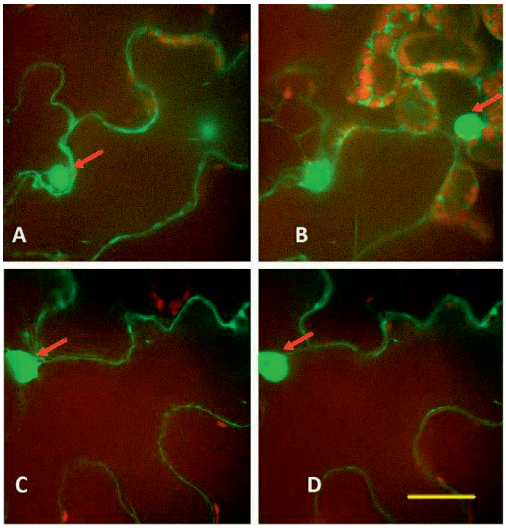


Figure 10. Confocal microscope Z-stack projection showing subcellular localization of SbSOT4A. GFP tagged SbSOT4A were transiently expressed in *N. benthamiana* leaves. The pictures are a merge of the GFP signal (green) and the autofluorescence (red). GFP signal at the N-terminus (A and B) and C-terminus (C and D) of SbSOT4A. Scale bar (yellow line) corresponds to 30 μm.

Structural similarities between At SOT18 and SbSOT4A

So far, AtSOT18 of *Arabidopsis thaliana* is the only plant SOT with a complete structure published. AtSOT18 is a sulfotransferase that catalyzes the sulfation of desulphoglucosinolates. *In vitro*, AtSOT18 was shown to sulfate a broad spectrum of substrates but with preference for 7-methylthioheptyl and 8-methylthiooctyl glucosinolates (Klein et al. 2009). In transient assays, the subcellular localization of AtSOT18 was shown to be in the cytosol. Sulfation plays an important role in glucosinolate biosynthesis since the desulphoglucosinolates are not functional. It has been suggested that the desulphoglucosinolates are a storage or transport form since they are less toxic to the plant (Mithen 2001; Klein et al. 2006). A homology model was built for SbSOT4A, using the X-ray structure of *Arabidopsis thaliana* SOT18 (AtSOT18), co-crystallized with the substrate sinigrin (which was sulfated) and the (desulfated) cofactor PAPS, as template. SbSOT4A displays only 33% sequence identity with AtSOT18 and for an optimal sequence alignment, the longer N- and C-termini of the sorghum SOT were removed. The PAPS binding site showed a high level of identity, with the key residues for cofactor binding and catalytic activity well conserved between AtSOT18 and SbSOT4A (Figure 11a, Table 2). According to the PDB 5MEX, the binding of PAPS to the receptor is stabilized by hydrogen bonds between the 3'-phosphate group and the side chains of Arg177, Ser185 and Arg313, between the 5'-phosphate group and the sidechains of Thr96 and Thr97 and between the adenyl moiety and the sidechains of Cys282 and Tyr243. All these amino acids are conserved in SbSOT4A (Table 2). In AtSOT18, the binding of the cofactor is further stabilized by hydrophobic contacts with Phe284 and the aliphatic chain of Arg247. In SbSOT4A, these residues correspond to two leucines, which are also able to provide stabilizing contacts due to their hydrophobic character (Table 2).

The substrate-binding site, however, is very different between the two SOTs, suggesting a different substrate specificity. For AtSOT18, the analysis of the PDB highlights a network of hydrogen bonds between sulfated sinigrin and the polar side chains of the substrate-binding site (Figure 11b, Table 2). Lys93, Thr96 and His155, which are involved in the coordination of the sulfate moiety, are also conserved in SbSOT4A (Lys103, Thr106 and His171), further confirming the potential sulfation catalytic activity of the enzyme (Table 2). Sinigrin binding to AtSOT18 is further facilitated by hydrogen bonds with the side chains of Arg51, Glu54, Glu193 and Tyr306 (Figure 11b). However, none of these are present in SbSOT4A (Table 2). In particular, Glu193 and Tyr306 are replaced by an alanine and a threonine, respectively, which do not have the correct polar or geometrical features to allow the formation of the same hydrogen bonds. Arg51 and Glu54 belong to a poorly conserved region of the enzyme. It is therefore difficult to build a proper homology model for this part of the protein. Nevertheless, it seems unlikely that SbSOT4 has similar polar amino acids in this region that could be involved in analogous hydrogen bonds with the substrate. Overall, the binding cavity in SbSOT4A thus has a much-increased hydrophobic character compared to AtSOT18, as shown by several polar to apolar changes in the sequence alignment. This indicates a preference of AtSOT18 for polar

substrates (e.g. sinigrin), whereas the SbSOT4A should favor the binding of substrates with an increased hydrophobic character, such as possibly a strigolactone-like molecule.

Table. 2. Residues that are important for PAPS binding site for AtSOT18 and SbSOT4A

PAPS site	92	93	94	95	96	97	98	125	177	185	243	247	282	283	284	287	311	312	313	314	315	316
AtSOT18	P	K	T	G	T	T	W	H	R	S	Y	R	C	S	F	L	Y	F	R	K	G	K
SbSOT4A	P	K	S	G	T	T	W	H	R	S	Y	L	C	S	L	L	F	F	R	K	G	K

Table. 3: Residues that are important for substrate specificity for AtSOT18 and SbSOT4A.

Substrate site	51	54	65	68	92	93	96	125	130	132	133	134	155	186	189	193	215	216	217	299	302	303	304	305	306	311
AtSOT18	R	E	L	L	P	K	T	H	Y	E	I	D	H	M	F	E	S	G	Y	R	R	P	G	V	Y	Y
SbSOT4A	-	-	A	L	P	K	T	H	F	E	L	M	H	Y	F	A	F	P	G	T	V	L	G	V	T	F

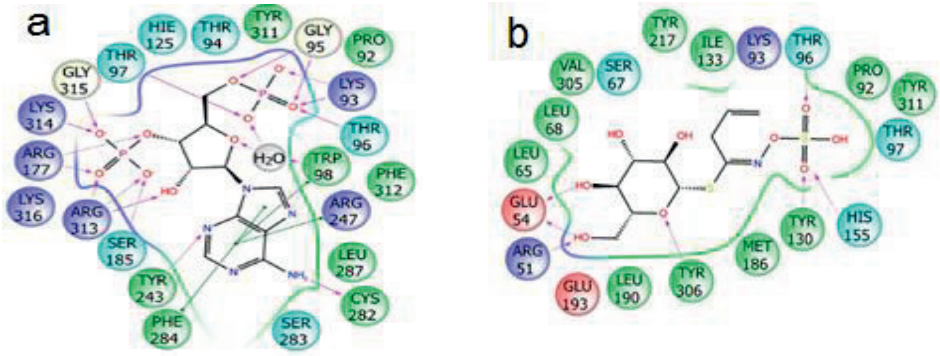


Figure 11. Residues in AtSOT18 that are involved in (a) cofactor (PAPS; desulfated) and (b) substrate (sinigrin; sulfated) binding.

Docking analysis reveals affinity of SbSOT4A for strigolactone-like compounds

In an attempt to explain a possible role for SbSOT4A in the biosynthesis of strigolactones, we hypothesized that sulfation may occur on a hydroxylated carlactone-like or strigolactone-like intermediate. Computer modeling with GOLD (using GoldScore fitness score, flexible amino acid sidechains) was used for docking of several putative substrates into the active site of SbSOT4A (Table 4). This docking technique was also used to explain positional preference in sulfate formation by using substrates that contain a -OH group at different positions. Different compounds were successfully docked into the substrate cavity and displayed a common binding mode, with the sulfate moiety binding deep at the catalytic site, whereas the rest of the molecule extends to the edge of the binding site.

Since SbSOT4 has been shown to play a role in the production of 5-deoxystrigol while a deletion of this enzyme results in the production of orobanchol (Gobena et al. 2017), we checked docking of these strigolactones as well as a number of stereoisomers and further functionalized strigolactones derived from these, such as *ent-2'-epi*-orobanchol, strigol, *ent-2'-epi*-strigol and sorgomol. To investigate the possibility that sulfation of a biosynthetic intermediate plays a role, also a number of carlactone-derived compounds were assessed for

their binding affinity, such as 19-hydroxycarlactone, C18-hydroxycarlactone and 18-hydroxycarlactonoic acid (Table 4). 19-Hydroxycarlactone has been shown to be generated by AtMAX1 as intermediate in the conversion of carlactone to carlactonoic acid (Abe et al. 2014). 18-hydroxycarlactone and 18-hydroxycarlactonoic acid have not been demonstrated to exist in nature but were postulated to be (likely enzyme-bound) intermediates in the biosynthesis of 4-deoxyorobanchol from carlactone by the rice MAX1 homolog Os900 (Zhang et al. 2014). Of all the possible substrates, 18-hydroxycarlactone gave the best docking score (Table 4). Figure 12, therefore, shows the substrate-binding site of the homology model of SbSOT4A with sulfated C18-hydroxycarlactone docked in comparison with the binding site of the protein structure of AtSOT18 with sulfated sinigrin. The sulfate group of the sulfated C18-hydroxycarlactone is coordinated by Lys103, Thr106 and His171, which corresponds to the conserved Lys93, Thr96 and His155 in AtSOT18 that are involved in the same hydrogen bond network with the sulfate group of sinigrin (Figure 12).

In both proteins, the main scaffold of the ligands shows a good shape complimentary with the binding site and is stabilized by Van der Waals contacts with hydrophobic residues in SbSOT4. In AtSOT18, the same region of the binding site is occupied by the glucopyranose ring of sinigrin that is interacting with polar residues. This confirms the hydrophobic character of the active site in SbSOT4, as discussed above based on the sequence alignment. This seems consistent with the binding of the more apolar scaffold of a strigolactone-like molecule. Interestingly, in SbSOT4 an additional hydrogen bond is present between the D-ring carbonyl and the sidechain of Thr330, which in AtSOT18 corresponds to Tyr306 and is instead extending to interact with the oxygen of the glucopyranose ring of sinigrin. A comparable binding mode was also found for all other carlactone analogs with no significant differences in the docking score. The strigolactones (i.e. the compounds with canonical ABC scaffold), instead, showed a less convincing binding mode with much lower docking scores.

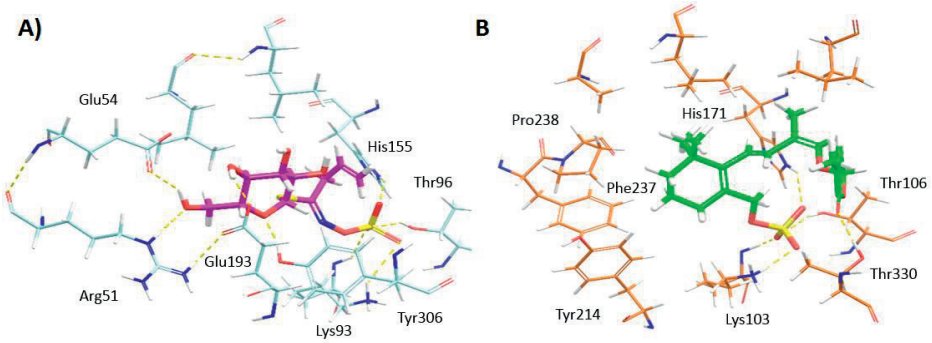
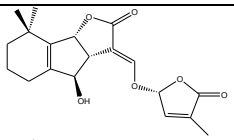
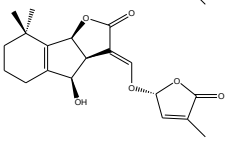
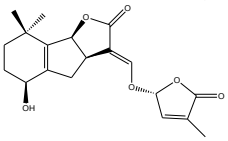
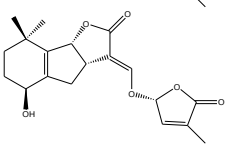
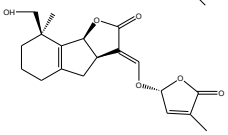
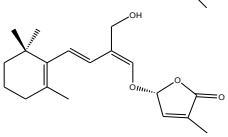
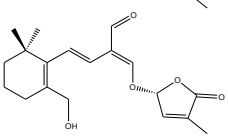
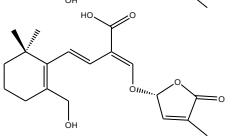
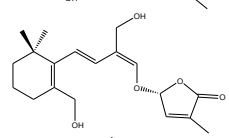
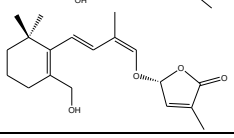


Figure 12. Partial models of the substrate-binding region of AtSOT18 A) with its substrate sinigrin and SbSOT4 with the putative substrate C18-hydroxycarlactone B) which showed the highest docking efficiency of all tested putative substrates. Note the difference in polarity of the amino acids in the binding sites of the two enzymes that might correlate with their specificity for a substrate with strong difference in polarity. Homology models were constructed using the software MODELLER (Sali, 1993 #86) (version 9.18). To build the homology model for SbSOT4A, the X-ray structure of *Arabidopsis* thaliana SOT18 (AtSOT18, Protein Data Bank (PDB) ID 5MEX, Uniprot ID Q9C9C9), co-crystallized with the substrate sinigrin (which was sulfated) and the (desulfated) cofactor PAPS as template was used.

Table 4. Postulated substrates for SbSOT4A and controls used for docking. Computer modeling with GOLD (using Gold Score fitness score, flexible amino acid sidechains) was used for docking these putative substrates into the active site of SbSOT4A

Name of the compound	Chemical Formula	Molecular weight	Structure	Docking score
orobanchol	C19H22O6	346.38		41.759
ent-2`-epi-orobanchol	C19H22O6	346.38		43.717
Strigol	C19H22O6	346.38		55.117
ent-2`-epi-strigol	C19H22O6	346.38		51.041
Sorgomol	C19H22O6	346.38		50.196
C19-hydroxycarlactone	C19H26O4	318.41		58.163
Carlactonoic aldehyde_C18OH	C19H24O5	332.40		63.287
C18-hydroxycarlactonoic acid	C19H24O6	348.40		59.607
C19-hydroxycarlactone _C18OH	C19H26O5	334.41		66.547
C18-hydroxycarlactone	C19H26O4	318.41		66.704

Discussion

Plants can secrete different blends of strigolactones, both in amount and type, into the rhizosphere (Yoneyama et al. 2009). In most cases, plants mainly produce one of the two classes of canonical strigolactones that are divided, based on the C-ring stereochemistry, into strigol- and orobanchol-type strigolactones (Yoneyama et al. 2018; Wang et al. 2018). In some cases, such as in tobacco, tomato and sorghum, both strigol and orobanchol-type strigolactones are produced (Xie et al. 2013; Zhang et al. 2018). These differences in strigolactone composition may potentially affect the fitness of the plant in response to biotic and abiotic stresses. Indeed, in recent work, we showed that the strigolactone profile is a key factor in determining the performance of sorghum genotypes to *Striga* infection (Chapter 3) (Gobena et al. 2017). *Striga* resistant sorghum genotypes have been shown to have low *Striga* germination stimulant activity which is caused by a change in the type of strigolactone exuded into the rhizosphere. Orobanchol producing sorghum genotypes tend to be more *Striga* resistant than those that produce 5-deoxystrigol due to their lower germination stimulant activity. We also showed that this difference is caused by SbSOT4A. Sorghum lines with low *Striga* germination stimulant activity carry mutations in SbSOT4A, also called *LGS1* (*Sobic.005G213600*) resulting in complete deletion of this gene (*lgs1-1*, *lgs1-2* and *lgs1-3* alleles) or severe exon mutations (*lgs1-4* and *lgs1-5*) that prevent SbSOT4A from being properly translated (Gobena et al. 2017).

In this chapter, we further confirmed that lack of SbSOT4A results in differences in the strigolactone profile in the root exudate of sorghum, particularly in a change in the stereochemistry of the strigolactones exuded. A clear correlation between the stereochemistry of strigolactones in root exudates and the presence of the SRN39 allele (*lgs1-1*) was shown in Figure 3. Therefore, we hypothesized that SbSOT4A is involved in the modification of strigolactones or strigolactone-like compounds which somehow results in a different stereochemistry depending on the presence or absence of the SOT. In this chapter, we tried to further characterize this SOT. My main objective was to deepen our understanding of the mechanism by which it can affect the strigolactone composition of sorghum and how the absence or presence of this enzyme results in the formation of strigolactones with different stereochemistry.

Although some plant SOTs have been functionally characterized, sulfation is still a poorly understood phenomenon. The SOTs form a multi-protein family of enzymes that are present in many organisms including plants. The enzymes are only active when the sulfate donor, PAPS, is available (Hirschmann et al. 2017). Sulfation in plants has been shown to be an important enzymatic activity that plays a role in the modulation of signaling in response to stress. For instance, the first characterized SOTs from *Flaveria chloraefolia* shed light on their possible role in regulation of plant growth. All three FcSOTs were shown to catalyze the sulfation of the flavonoid quercetin. The sulfated quercetin stimulates auxin transport from the apical tissues and promotes plant growth (Ananvoranich et al. 1994; Varin et al. 1989). AtSOT12 (the product of *At2g03760*) that has SOT activity towards brassinosteroids was also

shown to use salicylic acid and flavanone as substrates (Baek et al. 2010; Hashiguchi et al. 2014; Marsolais et al. 2007). Interestingly, it was also shown to have a broader spectrum of substrates and play a detoxification role by sulfating cycloheximide and toxins produced by bacteria(Chen et al. 2015). The expression of three brassinosteroid SOTs from *Brassica napus* and two from *Arabidopsis thaliana* was shown to be induced by salicylic acid (SA) which suggests they may have a role in plant defense against pathogens (Rouleau et al. 1999; Marsolais et al. 2007). In humans and animals, cytosolic SOTs have been shown to inactivate signal molecules such as steroids and thyroids or to play a role in the storage of some molecules which later can be reactivated by sulfatases or be transported by transporters (Coughtrie 2016; Reed et al. 2005). A good example of the latter is the *Arabidopsis* cytosolic SOT, AtSOT18 that sulfates glucosinolates, which can facilitate transport or storage (Klein et al. 2006; Graser et al. 2001). Salicylic acid sulfation is mediated by AtSOT12 which results in a less toxic form and was suggested to happen when the level of SA is high after the plant has been exposed to stress such as pathogen infection (Baek et al. 2010). In other cases, sulfation is a form of activation of metabolites. For instance, several studies on SOTs demonstrated their involvement in the plant response to biotic and abiotic stresses. For instance, mutants in AtSot12, which is known to sulfate brassinosteroids, have been shown to be hypersensitive to salt and display an increased level of ABA and higher susceptibility to pathogens (Baek et al. 2010; Hirschmann et al. 2014). In the case of the response to *Striga* infestation, we can speculate that sulfation in the strigolactone pathway is a resistance mechanism that regulates the production of orobanchol-type strigolactones. Orobanchol has a very low activity in the induction of germination of *Striga*, while it is a highly active hyphal branching inducing factor in arbuscular mycorrhizal fungi (Chapter 2).

Though evidence is not yet available we speculate that the route to the production of orobanchol is different in different plant species or even different varieties within one species. As shown in chapter 2, Shanqui-red - which has an intact SbSOT4A - is also producing a small amount of orobanchol which could suggest the presence of another route to orobanchol, possibly similar to rice. However, none of the sorghum homologs of the rice MAX1 that converts 4-deoxyorobanchol to orobanchol, Os1400, catalyzed the same reaction in our transient assay (Chapter 5). Sulfation in sorghum thus seems to be a mechanism to change the stereochemnistry of the strigolactone produced from strigol- to orobanchol-type. However, we cannot rule out the possibility that strigolactone/strigolactone-like compounds are sulfated for the purpose of facilitating transportation and this needs further study.

So, what compound could be the substrate of SbSOT4A? We first determined if SbSOT4A is a cytosolic or membrane-associated sulfotransferase. This is important since the substrates of membrane-associated sulfotransferase are macromolecules such as peptides and proteins. The few plant membrane-associated sulfotransferases that have been characterized have been shown to sulfate regulatory or signaling proteins (Hernàndez-Sebastiá et al. 2008). The cytosolic SOTs transfer the sulfuryl group of PAPS to small molecules such as hormones and xenobiotics (Gamage et al. 2006; Chapman et al. 2004). In plants, these group of enzymes are either predicted or functionally characterized to have SOT activity towards plant hormones

and secondary metabolites such as brassinosteriods, salicylic acid, hydroxyl-jasmonate, flavonoids and desulphoglucosinolates (Gidda et al. 2003; Marsolais et al. 2007; Hirschmann et al. 2014; Gidda et al. 2006). A close look at the highly conserved regions (I-IV) of both plant and human SOTs reveals that these regions are present in SbSOT4A. In addition, highly conserved motifs such as GxxGxxK in Region IV were present. Region II, a 3'-phosphate-binding (3'-PB) motif, starts with a highly conserved histidine, which is common for the entire SOT family. It is responsible for proton acceptance during surfuryl transfer (Kakuta et al. 1998). At the end of this region, an Arg and a Ser with an eight amino acid gap in between, are also conserved. These residues are responsible for the binding of the 3`-phosphate of PAP (Hirschmann et al. 2014). Also, additional residues such as the RKG residues and GxxGxxK motif (in Region IV; as mentioned above) are highly conserved among SOTs (Hirschmann et al. 2014; Marsolais et al. 1995), but the former was not found in SbSOT4A. The GxxGxxK motif is only known to be conserved in cytosolic SOTs but not in membrane-associated SOTs whereas the reverse is true for the other motif. The GxxGxxK motif is required for the binding of PAPS. A substitution of G for K resulted in a loss of enzyme activity (Komatsu et al. 1994). Other conserved regions such as the KTVE motif which consists of ten amino acid residues (KxxxTVxxxE) near the C-terminus was also shown to be highly conserved in the cytosolic SOTs. This motif generally mediates dimerization of cytosolic SOTs (Negishi et al. 2001; Weitzner et al. 2009; Petrotchenko et al. 2001). However, recent work showed that this motif is only conserved in human SOTs and is not found in any of the 22 *Arabidopsis* SOTs (Hirschmann et al. 2017; Tibbs et al. 2015). This motif was also not found in SbSOT4A and using an online protein modeling tool the protein was predicted to be a monomer and therefore does not need this motif. Furthermore, in terms of substrate recognition, a typical cytosolic SOT recognizes mostly hydrophobic molecules while the membrane-associated ones recognize hydrophilic peptides (Rath et al. 2004). Targeted mutagenesis studies in flavonol 3-SOTs showed that amino acid residues 92-194 determine their substrate specificity (Marsolais et al. 1995). The substrate-binding sites have certain differences and determine the substrate specificity of each SOT. For instance, Arg51, Glu54, Glu193 and Tyr306 are residues that are present in AtSOT18 but not in SbSOT4A and this results in a more hydrophobic pocket (Figure 11 and Table 2 and 3). This hydrophobic character suggests that more hydrophobic molecules such as strigolactone-like compounds serve as substrate rather polar compounds such as sinigrin.

The structure of the cytosolic and membrane-associated SOTs also differs in the substrate-binding site. For the membrane-associated sulfotransferases, a large open cleft with hydrophilic surface serves as the binding site whereas a deep hydrophobic pocket is the binding site for the cytosolic SOTs which determines the substrate specificity (Gamage et al. 2006). In general, SOTs have a single α/β domain with five-strand parallel β-sheets surrounded by α-helices. The cytosolic SOTs have a central core of four of these β-strands that is surrounded by the α-helices (Hirschmann et al. 2017). SbSOT4A has four central β-strands surround by α-helices which is similar to an *Arabidopsis* cytosolic SOT, AtSOT18. On the other

hand, all five β-strands of AtSOT19 are centered and surrounded by α-helices (Figure 9A, B and C).

All this suggests that SbSOT4A is a cytosolic SOT, which was further supported by the subcellular localization using GFP labeled SbSOT4A that was clearly localized in the cytosol. We also found GFP signals in all Z-layers of the nucleus. The positively charged amino acids, lysine and arginine, have low abundance, which suggests that SbSOT4A is negatively charged. If this is so, the GFP signal observed in the nucleus might be the result of an active diffusion of this 72.9 kDa protein (SbSOT4A with GFP) into the nuclear pore complexes. The nuclear pores can allow diffusion of a protein up to a size of 110 kDa (Wang et al. 2007). The cytosolic localization of SbSOT4A is consistent with the production of strigolactones in the cytosol (Booker et al. 2005). All this seems to rule out the possible hypothesis of a role for SbSOT4A in post-translational protein modification, which would be possible with a membrane-associated sulfotransferase.

To find possible targets of SbSOT4A we looked at the function of all known plants SOTs. In *Brassica napus*, 71 SOTs have been found while 22 and 35 are present in *Arabidopsis thaliana* and *Oryza sativa*, respectively (Hirschmann et al. 2014; Hirschmann et al. 2015; Chen et al. 2012). 21 out of the 22 *Arabidopsis* SOTs are cytosolic. AtSOT19 is predicted to be a membrane-associated SOT. In sorghum, we found 25 genes annotated as SOTs by blasting the amino acid sequence of SbSOT4A against the sorghum genome in the NCBI database. In order to be able to conclude more about the substrate specificity of SbSOT4A we first analyzed its phylogenetic relationship with other plant SOTs. However, this is not without challenges. In the first place, SbSOT4A did not cluster with any functionally characterized SOT. In the second place, other studies have shown that SOTs that cluster together do not necessarily share the same substrate specificity (Hernàndez-Sebastiá et al. 2008; Klein et al. 2004; Hirschmann et al. 2014). In some cases, SOTs with more than 80% identity did not have the same substrate preference (Klein et al. 2004; Klein et al. 2008). These phylogenetic studies have been conducted using the entire gene sequence which makes it difficult to zoom in on less conserved regions such as the binding pocket, which is highly diverse both in plant and human SOTs (Gamage et al. 2006). In this work, we also showed that the amino acids in the (modeled) substrate-binding site of SbSOT4A have very different properties from those of AtSOT18 (Figure 11, Table 2). Where AtSOT18 has a very polar substrate-binding site, the binding site of SbSOT4A is much more apolar. This suggests that they sulfate very different substrates. Indeed, our docking analysis showed that the more lipophilic substrates, such as strigolactones and carlactone-like compounds, dock much better into SbSOT4A than the polar sinigrin (Table 3). Of the tested potential substrates C18-hydroxycarlactone had the highest docking score, but also other carlactone analogs showed comparable docking scores suggesting they may also be possible substrates for SbSOT4A. However, the canonical strigolactones had substantially lower docking scores, making them unlikely substrates. Also, *Arabidopsis* SOTs have been shown to accept multiple substrates. For instance, flavonoids, brassinosteroids and salicylic acid were all sulfated by AtSOT12 (Baek et al. 2010; Marsolais et al. 2007; Lacomme et al. 1996; Hashiguchi et al. 2013). In some cases, it has been shown that SOTs sulfate in a

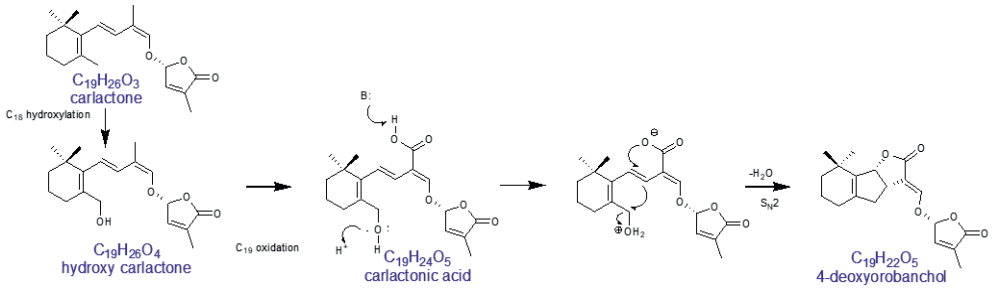
stereospecific manner. For instance, AtSOT12 was shown to be stereospecific for 24-*epi*-brassinosteroid (Marsolais et al. 2007).

Different SOTs showed specificity for certain positions of the hydroxyl group. The preference between *Arabidopsis thaliana* and *Flaberia bidentis* flavonol SOTs are good examples of this. The first prefer the hydroxyl position 3 or 7, the latter shows more affinity with quercetin. More interestingly, a test on the substrate specificity of five different hydroxyflavones that contain a hydroxyl group at different positions showed that AtSULT202B7, a flavonol SOT from *Arabidopsis*, sulfates region specifically (Hashiguchi et al. 2014). The specificity of these enzymes was found to be different in different accessions of *Arabidopsis*. For instance, for flavonoids, four positions have been identified as targets for sulfation.

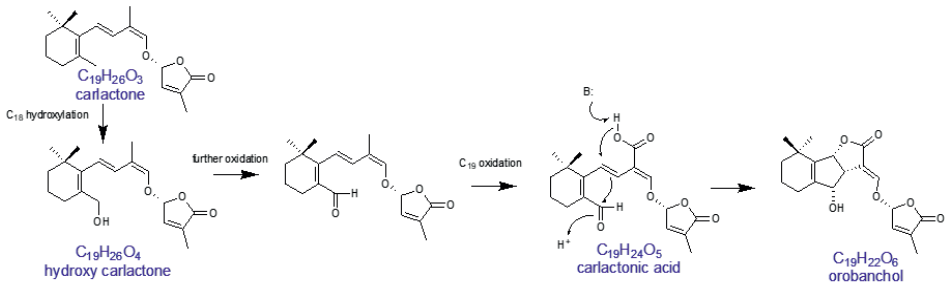
The unexpected outcome that C18-hydroxy-carlactone is the substrate with the highest affinity for the SbSOT4A binding site prompted us to revisit our ideas about strigolactone biosynthesis and how this may be affected by this sorghum sulfotransferase (Figure 13). In the case of rice, biosynthesis of strigolactones from carlactone is mediated by a P450 that hydroxylates C18 and oxidizes C19 which results in the production of 4-deoxyorobanchol. Possibly, in sorghum genotypes with a non-functional SbSOT4A, C18 oxidation proceeds to the aldehyde. If this is followed by oxidation at C19 – possibly by one of the MAX1 homologs - to form C18-oxo-carlactonoic acid, ring closure can result in direct orobanchol formation. We cannot completely exclude that some sorghum lines can produce orobanchol as in rice (in our feeding assay 4-deoxyorobanchol seemed to be a precursor of orobanchol, Figure 5A). Nevertheless, in the presence of a functional SbSOT4A, the amount produced by these lines is by far not as high as in rice. In Chapter 5, none of the four sorghum MAX1 homologs that were cloned from SRN39 were able to produce 4-deoxyorobanchol, which might suggest that in sorghum 4-deoxyorobanchol is not a natural precursor for orobanchol but rather that it is produced through cyclization of the 18-oxo-carlactonoic acid. We postulate that if a functional SOT is present, C18-hydroxycarlactone is sulfated before the second oxidation step that forms the aldehyde occurs. This sulfated C18-hydroxycarlactone can then be further converted to sulfated C18-hydroxycarlactonoic acid by oxidation at the C19 position, again possibly by one of the MAX1 homologs of sorghum. The removal of the sulfate group could be followed by ring closure and could result in the production of 5-deoxystrigol. Taken together, the results suggest that SbSOT4a may control strigolactone stereochemistry in sorghum by sulfation or not (in *lgs1* mutants) of the biosynthetic intermediate C18-hydroxycarlactone, which results in the production of 5-deoxystrigol or of orobanchol, if sulfation does not occur, and high or low *Striga* germination stimulant activity, respectively. Transgenic sorghum in which SbSOT4A is knocked down or out can hopefully prove this proposed mechanism. To be able to translate these findings to other *Striga* susceptible crops such as rice, we need to further study the presence and function of SbSOT4A homologs. In our phylogenetic analysis, OsSOT27 *LOC_OS11g26390* was found in the same branch as SbSOT4A. This gene is annotated as putative flavonol 3-sulfotransferase and is predicted to be cytosolic (Chen et al. 2012). Four more rice genes were found in the same clade which might suggest they have a similar function (Figure 7). However, none of these genes has been functionally characterized so far.

It would be interesting to study if these genes play any role in strigolactone biosynthesis in rice.

In Rice



<u>In sorghum:</u> where there is no functional SOT



<u>In sorghum:</u> when SOT is intact

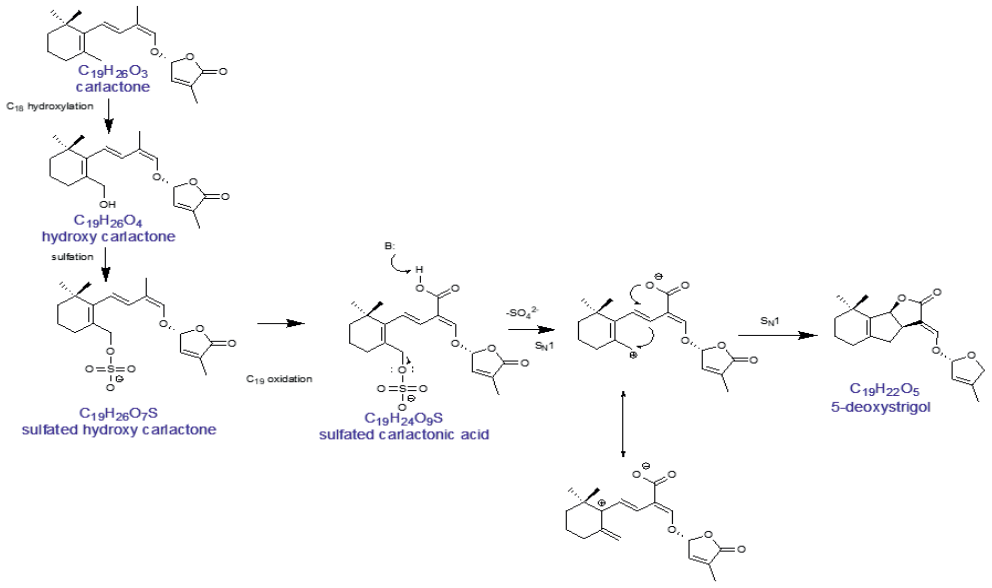


Figure 13. Overview of the (putative) biosynthesis pathway of some strigolactones in rice and sorghum and the possible role of a sulfotransferase, SbSOT4A, in this.

Materials and Methods

Strigolactone extraction and detection from root exudates and extracts

Sorghum seeds were surface sterilized and germinated at 25^{o}C for 48 hrs in darkness on moistened filter paper. The germinated seeds were transferred to 14 cm pots (three plants per pot) filled with river sand. The seedlings were grown for two weeks in a climate chamber with temperature setting of 28^{o}C and 16 hrs light. The seedlings were watered with half-strength Hoagland's nutrient solution. After two weeks, the pots were flushed with 1 L of phosphate deficient half strength Hoagland's to remove the phosphate from the pots. The plants were then grown for an additional week and watered with half strength Hoagland's without phosphate to induce strigolactone production. After that week, a second flushing was used to rinse out any metabolites, including strigolactones, accumulated in the pot. Another 48 hrs later, 1 L of root exudate was collected by passing 1 L of tap water through each pot. Subsequently, the roots were collected, washed, frozen in liquid nitrogen and stored at -80^{o}C for further analysis.

The collected root exudates were passed through a C18 column (500 mg, Grace pure) that was preconditioned with 3 ml acetone and 3 ml of demineralized water. The strigolactones were eluted with 4 ml acetone. Two ml of the eluent were used for further analysis. 200 µl of internal standard (D6-*epi*-5-deoxystrigol) to a final concentration 0.1 nmol/ml was added, and the acetone was evaporated to dryness. The samples were then re-dissolved in 50 µl of ethyl acetate and 4 ml hexane and passed through a silica column (200 mg) (preconditioned with 2 ml ethyl acetate and 4 ml hexane). The columns were eluted with 2 ml of 10:90 hexane: ethyl acetate (v/v). The eluent was evaporated to dryness and re-dissolved in 200 µl of 25% acetonitrile in water. The samples were filtered through a 0.45 µm Minisart SRP filter (Sartorius, Germany). Strigolactones were then analyzed by comparing the retention times and mass transition with the internal standards using a Waters Xevo tandem quadrupole mass spectrometer. An Acquity UPLC BEH C18 column was used for chromatographic separation by applying a water: acetonitrile gradient with 0.1% formic acid to the column. The starting gradient was 5% acetonitrile for 0.33 min. Then, it was raised to 27% (v/v) acetonitrile in 0.34 min, followed by a 4.33 min gradient to 40% (v/v) acetonitrile, then rising to 65% acetonitrile in 3 min and maintained for 0.67 min. A 90% acetonitrile gradient replaced that in 0.2 min was used and was maintained for 0.46 min. The gradient was then set back to 5% acetonitrile in 0.2 min and maintained for 2.47 min to equilibrate the column prior to the next run. The column temperature was set at 50^{o}C with flow rate of 0.5 mL/min. 10 – 20 µl of samples were injected. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L/hr, respectively. The source temperature was set at 150^{o}C, the capillary voltage was set at 3.0 kV and the desolvation temperature at 650^{o}C. For each standard, the cone voltage was optimized using Waters IntelliStart MS Console. MRM was used for identification of strigolactones. Data were analyzed using MassLynx 4.1 (combined with TargetLynx) software.

The stereochemistry of orobanchol was determined using a Waters Alliance e2695 HPLC connected to a Waters LCT Premier XE time-of-flight (TOF) mass spectrometer with an atmospheric chemical ionization source (APCI). Separation was performed with an Astec® Cellulose DMP chiral HPLC column (250 x 4.6 mm id, 5 µm, Supelco, Bellafonte, PA, USA) using a flow rate of 0.5 mL/min. Racemic mixture of orobanchol standard was injected and the separation of the four orobanchol stereo-isomers (*ent*-orobanchol, orobanchol, *ent*-2`-epi-orobanchol and 2`-*epi*-orobanchol) was achieved with a 6:1:1 hexane/methanol/methyl-tert-butyl ether. The column was held at 25°C and sample injection volume was 5 to 20 µL. The mass spectrometer was operated in positive APCI mode. Cone and desolvation gas flows were set to 250 and 650 $L \cdot h^{-1}$, respectively. The conrona voltage was set at 3.0 kV, the source temperature at 100°C, and the desolvation temperature at 350°C.

To extract strigolactones from roots, 2 ml ethyl acetate containing 0.1 µM internal standard GR24 was added to 500 mg frozen root powder. The samples were sonicated for 10-15 min in a Branson 3510 ultrasonic bath, centrifuged for 10 min at 2000 rpm and the supernatant transferred to a clean vial. An additional 2 ml of ethyl acetate were used to re-extract the pellet which was then centrifuged again for 10 min. The combined supernatants were evaporated to dryness and the residue re-dissolved in 50 µl ethyl acetate and 4 ml hexane. The samples were then loaded on a silica column (200 mg) (preconditioned with 2 ml ethyl acetate and 4 ml hexane). The column was eluted in fractions: fraction 1 (eluted with 4 ml 35:80 ethyl acetate: hexane) and fraction 2 (4 ml 90:20 EtOAc: hexane). The fractions were combined, and the solvent was evaporated to dryness after which the residue was re-dissolved in 200 µl of 25% acetonitrile in water. Strigolactones were then analyzed as described above.

Feeding Assay

Sorghum seedlings were grown for two weeks in hydroponics consisting of modified half-strength Hoagland's nutrient solution. The seedlings were transferred to 50 ml glass bottles containing 40 ml of tap water with and without fluridone and grown for 3 days. Then the water was replaced with clean water and in each pot, 0.4mg of carlactone, 5-deoxystrigol or 4-deoxyorobanchol were added. After 24 hrs, the root exudates were collected and strigolactones were extracted and analyzed as described above.

SOT Inhibition assay

Seeds were grown in river sand as described above. After a week of phosphate starvation, the pots were flushed with 1 L of tap water to remove accumulated strigolactones. The plants were subsequently watered with 100 ml water with and without 10 µM of the SOT inhibitor triclosan. After 48 hrs, the pots were flushed with 1 L of water to collect root exudates and strigolactones were extracted and analyzed as described above.

Generation of expression constructs and microscopy

LGS1 was amplified by PCR from cDNA synthesised from sorghum genotype Shanqui-Red using gene specific primers. Subsequently, the cDNA was fused with Green Fluorescent Protein (GFP) gene in the two expression vectors, pK7FWG2 (for N-terminal GFP fusion) and pK7WGF2 (For C-terminal GFP fusion). The constructs were transformed to *Agrobacterium tumefaciens.* Infiltration of the *Agrobacterium* into *Nicotiana benthamiana* through the abaxial leaf surface was done after adjusting the OD_{600} of the suspension to 0.5. Cellular location of the labeled SbSOT4A protein was monitored six days after infiltration by examining leaves for the GFP signal under a confocal laser scanning microscopy (CLSM) at 60x magnification. Pictures were taken at several heights at fixed places (Z-layers) on slices of leaf tissue. GFP and red fluorescent were excited using 488 nm and the 561 nm wavelengths respectively. Representative images from the abaxial side of the *N. benthamiana* leaf discus were taken and processed using ImageJ (FIJI).

Protein modeling and docking

The protein modeling of SbSOT4A, AtSOT18 and AtSOT19 was conducted by feeding the amino acid sequence of the genes to the Phyre2 software (Kelley et al. 2015). A single template was selected by the software to be used for the modeling based on the highest confidence score for it to be homologous to the input sequence(s). For the modeling of SbSOT4A Sulfotransferase-18 from *Arabidopsis thaliana* (AtSOT18; c5mekA) was used as a template. For AtSOT19 the crystal structure of sulfotransferase stf1 from *Mycobacterium tuberculosis* h37rv (c2zq5A) was used as template.

For the docking studies, a homology model of SbSOT4A was based on the X-ray structure of AtSOT18 (Protein Data Bank (PDB) ID 5MEX, Uniprot ID Q9C9C9), co-crystallized with the substrate sinigrin (which was sulfated) and the (desulfated) cofactor PAPS. The software MODELLER(Sali et al. 1993). (version 9.18) was used to generate the homology model. SbSOT4A displays only 33% sequence identity with AtSOT18 and for an optimal sequence alignment, the longer N- and C-termini of the sorghum SOT were removed.

Docking of the strigolactone analogs onto the SbSOT4A homology model was performed with GOLD software (CCDC, version 5.2) (Jones et al. 1997), using GoldScore as fitness score and introducing flexible amino acid sidechains. Docking was done with the sulfated form of the strigolactone analogs, in order to compare with the binding mode of the reference sulfated sinigrin in the PDB 5MEX.

Data analysis

A heat map was created for the strigolactones detected in root exudates of different sorghum lines. The sorghum lines were selected from the LGS validation set used in Chapter two. Shanqui-Red, SSD#3-767, SSD#3-177, P9408 and N13 are sorghum lines with SbSOT4A intact and SRN-39, SSD#3-320, SSD#3-302, P9401 and 555 have SbSOT4A mutant genotypes. The average peak area of each strigolactone was analyzed using MetaboAnalyst, an online tool

(https://www.metaboanalyst.ca/). The default setting (with distance measure of Euclidean with Ward clustering algorithm) to normalize data. ANOVA was used for statistical analysis.

Acknowledgments

We thank Gavin R Flematti from School of Chemistry and Biochemistry, The University of Western Australia, Crawley, Australia for determining the stereochemistry of orobanchol, and K. Yoneyama (Weed Science Center, Utsunomiya University, Utsunomiya, Japan) and T. Asami (Department of Applied Biological Chemistry, The University of Tokyo, Tokyo, Japan) for supplying SL standards. This work was supported by Bill and Melinda Gates Foundation Grant OPP1006216 and Netherlands Organization for Scientific Research Vici Grant 865.06.002 and Equipment Grant 834.08.001 (to H.B.).

References

Abe, Satoko, Aika Sado, Kai Tanaka, Takaya Kisugi, Kei Asami, Saeko Ota, Hyun Il Kim, et al. 2014. 'Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18084-89.

Ananvoranich, S., L. Varin, P. Gulick, and R. Ibrahim. 1994. 'Cloning and regulation of flavonol 3-sulfotransferase in cell-suspension cultures of Flaveria bidentis', *Plant Physiol*, 106: 485-91.

Baek, Dongwon, Prasanth Pathange, Jung-Sung Chung, Jiafu Jiang, Liqiong Gao, Akira Oikawa, Masami Yokota Hirai, et al. 2010. 'A stress-inducible sulphotransferase sulphonates salicylic acid and confers pathogen resistance in Arabidopsis', *Plant, Cell & Environment*, 33: 1383-92.

Booker, J., T. Sieberer, W. Wright, L. Williamson, B. Willett, P. Stirnberg, C. Turnbull, et al. 2005. 'MAX1 encodes a cytochrome P450 family member that acts downstream of MAX3/4 to produce a carotenoid-derived branch-inhibiting hormone', *Dev Cell*, 8: 443-9.

Bouwmeester, H. J., Chr Roux, J. A. Lopez Raez, and G. Bécard. 2007. "Rhizosphere communication of plants, parasitic plants and AM fungi." In *Trends in Plant Science 12 (2007) 5*. 2007.

Brewer, P. B., K. Yoneyama, F. Filardo, E. Meyers, A. Scaffidi, T. Frickey, K. Akiyama, et al. 2016. 'LATERAL BRANCHING OXIDOREDUCTASE acts in the final stages of strigolactone biosynthesis in Arabidopsis', *Proc Natl Acad Sci U S A*, 113: 6301-6.

Chapman, E., M. D. Best, S. R. Hanson, and C. H. Wong. 2004. 'Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility', *Angew Chem Int Ed Engl*, 43: 3526-48.

Chen, J., L. Gao, D. Baek, C. Liu, Y. Ruan, and H. Shi. 2015. 'Detoxification function of the Arabidopsis sulphotransferase AtSOT12 by sulphonation of xenobiotics', *Plant Cell Environ*, 38: 1673-82.

Chen, Rongjun, Yunyun Jiang, Jiali Dong, Xin Zhang, Hongbo Xiao, Zhengjun Xu, and Xiaoling Gao. 2012. 'Genome-wide analysis and environmental response profiling of SOT family genes in rice (Oryza sativa)', *Genes & Genomics*, 34: 549-60.

Coughtrie, M. W. 1996. 'Sulphation catalysed by the human cytosolic sulphotransferases-chemical defence or molecular terrorism?', *Hum Exp Toxicol*, 15: 547-55.

Coughtrie, M. W. H. 2016. 'Function and organization of the human cytosolic sulfotransferase (SULT) family', *Chem Biol Interact*, 259: 2-7.

Cui, Songkui, Takanori Wakatake, Kei Hashimoto, Simon B. Saucet, Kiminori Toyooka, Satoko Yoshida, and Ken Shirasu. 2016. 'Haustorial Hairs Are Specialized Root Hairs That Support Parasitism in the Facultative Parasitic Plant *Phtheirospermum japonicum*', *Plant Physiology*, 170: 1492.

Czarnecki, Olaf, Jun Yang, David J. Weston, Gerald A. Tuskan, and Jin-Gui Chen. 2013. 'A dual role of strigolactones in phosphate acquisition and utilization in plants', *International journal of molecular sciences*, 14: 7681-701.

Driscoll, W. J., K. Komatsu, and C. A. Strott. 1995. 'Proposed active site domain in estrogen sulfotransferase as determined by mutational analysis', *Proceedings of the National Academy of Sciences*, 92: 12328.

Flematti, G. R., A. Scaffidi, M. T. Waters, and S. M. Smith. 2016. 'Stereospecificity in strigolactone biosynthesis and perception', *Planta*, 243: 1361-73.

Gamage, N., A. Barnett, N. Hempel, R. G. Duggleby, K. F. Windmill, J. L. Martin, and M. E. McManus. 2006. 'Human sulfotransferases and their role in chemical metabolism', *Toxicol Sci*, 90: 5-22.

Gidda, S. K., O. Miersch, A. Levitin, J. Schmidt, C. Wasternack, and L. Varin. 2003. 'Biochemical and molecular characterization of a hydroxyjasmonate sulfotransferase from Arabidopsis thaliana', *J Biol Chem*, 278: 17895-900.

Gidda, S. K., and L. Varin. 2006. 'Biochemical and molecular characterization of flavonoid 7-sulfotransferase from Arabidopsis thaliana', *Plant Physiol Biochem*, 44: 628-36.

Gobena, D., M. Shimels, P. J. Rich, C. Ruyter-Spira, H. Bouwmeester, S. Kanuganti, T. Mengiste, et al. 2017. 'Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes *Striga* resistance', *Proc Natl Acad Sci U S A*, 114: 4471-76.

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

Hashiguchi, T., Y. Sakakibara, Y. Hara, T. Shimohira, K. Kurogi, R. Akashi, M. C. Liu, et al. 2013. 'Identification and characterization of a novel kaempferol sulfotransferase from Arabidopsis thaliana', *Biochem Biophys Res Commun*, 434: 829-35.

Hashiguchi, Takuyu, Yoichi Sakakibara, Takehiko Shimohira, Katsuhisa Kurogi, Masao Yamasaki, Kazuo Nishiyama, Ryo Akashi, et al. 2014. 'Identification of a novel flavonoid glycoside sulfotransferase in Arabidopsis thaliana', *Journal of biochemistry*, 155: 91-97.

Hernàndez-Sebastiá, Cinta, Luc Varin, and Frédéric Marsolais. 2008. 'Sulfotransferases from Plants, Algae and Phototrophic Bacteria.' in Rüdiger Hell, Christiane Dahl, David Knaff and Thomas Leustek (eds.), *Sulfur Metabolism in Phototrophic Organisms* (Springer Netherlands: Dordrecht).

Hess, Dale E., Gebisa Ejeta, and Larry G. Butler. 1992. 'Selecting sorghum genotypes expressing a quantitative biosynthetic trait that confers resistance to Striga', *Phytochemistry*, 31: 493-97.

Hibberd, J. M., and W. D. Jeschke. 2001. 'Solute flux into parasitic plants', *J Exp Bot*, 52: 2043-9.

Hirschmann, F., F. Krause, and J. Papenbrock. 2014. 'The multi-protein family of sulfotransferases in plants: composition, occurrence, substrate specificity, and functions', *Front Plant Sci*, 5: 556.

Hirschmann, Felix, Florian Krause, Petra Baruch, Igor Chizhov, Jonathan Wolf Mueller, Dietmar J. Manstein, Jutta Papenbrock, et al. 2017. 'Structural and biochemical studies of sulphotransferase 18 from Arabidopsis thaliana explain its substrate specificity and reaction mechanism', *Scientific Reports*, 7: 4160.

Hirschmann, Felix, and Jutta Papenbrock. 2015. 'The fusion of genomes leads to more options: A comparative investigation on the desulfo-glucosinolate sulfotransferases of Brassica napus and homologous proteins of Arabidopsis thaliana', *Plant Physiology and Biochemistry*, 91: 10-19.

Jones, G., P. Willett, R. C. Glen, A. R. Leach, and R. Taylor. 1997. 'Development and validation of a genetic algorithm for flexible docking', *J Mol Biol*, 267: 727-48.

Kakuta, Yoshimitsu, Lee G. Pedersen, Lars C. Pedersen, and Masahiko Negishi. 1998. 'Conserved structural motifs in the sulfotransferase family', *Trends in Biochemical Sciences*, 23: 129-30.

Kapulnik, Y., and H. Koltai. 2014. 'Strigolactone involvement in root development, response to abiotic stress, and interactions with the biotic soil environment', *Plant Physiol*, 166: 560-9.

Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass, and M. J. Sternberg. 2015. 'The Phyre2 web portal for protein modeling, prediction and analysis', *Nat Protoc*, 10: 845-58.

Kester, M. H., E. Kaptein, T. J. Roest, C. H. van Dijk, D. Tibboel, W. Meinl, H. Glatt, et al. 1999. 'Characterization of human iodothyronine sulfotransferases', *J Clin Endocrinol Metab*, 84: 1357-64.

Klein, M., and J. Papenbrock. 2004. 'The multi-protein family of Arabidopsis sulphotransferases and their relatives in other plant species', *J Exp Bot*, 55: 1809-20.

Klein, M., and J. Papenbrock. 2009. 'Kinetics and substrate specificities of desulfo-glucosinolate sulfotransferases in Arabidopsis thaliana', *Physiol Plant*, 135: 140-9.

Klein, M., M. Reichelt, J. Gershenzon, and J. Papenbrock. 2006. 'The three desulfoglucosinolate sulfotransferase proteins in Arabidopsis have different substrate specificities and are differentially expressed', *Febs j*, 273: 122-36.

Klein, Marion, and Jutta Papenbrock. 2008. 'Sulfotransferases and Their Role in Glucosinolate Biosynthesis.' in Nafees A. Khan, Sarvajeet Singh and Shahid Umar (eds.), *Sulfur Assimilation and Abiotic Stress in Plants* (Springer Berlin Heidelberg: Berlin, Heidelberg).

Komatsu, K., W. J. Driscoll, Y. C. Koh, and C. A. Strott. 1994. 'A P-Loop-Related Motif (GxxGxxK) Highly Conserved in Sulfotransferases Is Required for Binding the Activated Sulfate Donor', *Biochemical and Biophysical Research Communications*, 204: 1178-85.

Lacomme, C., and D. Roby. 1996. 'Molecular cloning of a sulfotransferase in Arabidopsis thaliana and regulation during development and in response to infection with pathogenic bacteria', *Plant Mol Biol*, 30: 995-1008.

Lopez-Raez, J. A., T. Charnikhova, V. Gomez-Roldan, R. Matusova, W. Kohlen, R. De Vos, F. Verstappen, et al. 2008. 'Tomato strigolactones are derived from carotenoids and their biosynthesis is promoted by phosphate starvation', *New Phytol*, 178: 863-74.

Marsolais, F., J. Boyd, Y. Paredes, A. M. Schinas, M. Garcia, S. Elzein, and L. Varin. 2007. 'Molecular and biochemical characterization of two brassinosteroid sulfotransferases from Arabidopsis, AtST4a (At2g14920) and AtST1 (At2g03760)', *Planta*, 225: 1233-44.

Marsolais, Frédéric, Satinder Gidda, Jason Boyd, and Luc Varin. 2000. *Chapter Fourteen Plant soluble sulfotransferases: Structural and functional similarity with mammalian enzymes*.

Marsolais, Frédéric, and Luc Varin. 1995. 'Identification of Amino Acid Residues Critical for Catalysis and Cosubstrate Binding in the Flavonol 3-Sulfotransferase', *Journal of Biological Chemistry*, 270: 30458-63.

Marzec, Marek, Aleksandra Muszynska, and Damian Gruszka. 2013. 'The Role of Strigolactones in Nutrient-Stress Responses in Plants', *International journal of molecular sciences*, 14.

Matsubayashi, Y., and Y. Sakagami. 1996. 'Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of Asparagus officinalis L', *Proc Natl Acad Sci U S A*, 93: 7623-7.

Matusova, Radoslava, Kumkum Rani, Francel W. A. Verstappen, Maurice C. R. Franssen, Michael H. Beale, and Harro J. Bouwmeester. 2005. 'The Strigolactone Germination Stimulants of the Plant-Parasitic *Striga* and *Orobanche spp.* Are Derived from the Carotenoid Pathway', *Plant Physiology*, 139: 920.

Mithen, Richard. 2001. 'Glucosinolates – biochemistry, genetics and biological activity', *Plant Growth Regulation*, 34: 91-103.

Mohamed, A., A. Ellicott, T. L. Housley, and G. Ejeta. 2003. 'Hypersensitive Response to *Striga* Infection in Sorghum' Purdue Agric. Res. Programs, Journal Paper no. 16 745', *Crop Science*, 43: 1320-24.

Negishi, M., L. G. Pedersen, E. Petrotchenko, S. Shevtsov, A. Gorokhov, Y. Kakuta, and L. C. Pedersen. 2001. 'Structure and function of sulfotransferases', *Arch Biochem Biophys*, 390: 149-57.

Nomura, Takahito, Xiaonan Xie, Koichi Yoneyama, Kaori Yoneyama, Takaya Kisugi, Kohki Akiyama, Yoshifumi Nakatani, et al. 2018. 'Which are the major players, canonical or non-canonical strigolactones?', *Journal of Experimental Botany*, 69: 2231-39.

Petrotchenko, Evgeniy V., Lars C. Pedersen, Christoph H. Borchers, Kenneth B. Tomer, and Masahiko Negishi. 2001. 'The dimerization motif of cytosolic sulfotransferases', *FEBS Letters*, 490: 39-43.

Plaxton, William C., and Hue T. Tran. 2011. 'Metabolic Adaptations of Phosphate-Starved Plants', *Plant Physiology*, 156: 1006.

Prather, Brittany, Cheryl Ethen, Miranda Machacek, and Leon wu. 2011. *Phosphatase-coupled Sulfotransferase Assay*.

Rath, V. L., D. Verdugo, and S. Hemmerich. 2004. 'Sulfotransferase structural biology and inhibitor discovery', *Drug Discov Today*, 9: 1003-11.

Reed, M. J., A. Purohit, L. W. Woo, S. P. Newman, and B. V. Potter. 2005. 'Steroid sulfatase: molecular biology, regulation, and inhibition', *Endocr Rev*, 26: 171-202.

Rich, Patrick J., Cécile Grenier, and Gebisa Ejeta. 2004. 'Striga Resistance in the Wild Relatives of Sorghum This research was supported by the Rockefeller Foundation Grant No. 2000FS024. Purdue Agricultural Res. Programs, Journal Article No. 17064', *Crop Science*, 44: 2221-29.

Rouleau, Michèle, Frédéric Marsolais, Martine Richard, Ludovic Nicolle, Brunhilde Voigt, Günter Adam, and Luc Varin. 1999. 'Inactivation of Brassinosteroid Biological Activity by a Salicylate-inducible Steroid Sulfotransferase from Brassica napus', *Journal of Biological Chemistry*, 274: 20925-30.

Sali, A., and T. L. Blundell. 1993. 'Comparative protein modelling by satisfaction of spatial restraints', *J Mol Biol*, 234: 779-815.

Satish, K., Z. Gutema, C. Grenier, P. J. Rich, and G. Ejeta. 2012. 'Molecular tagging and validation of microsatellite markers linked to the low germination stimulant gene (lgs) for Striga resistance in sorghum [Sorghum bicolor (L.) Moench]', *Theor Appl Genet*, 124: 989-1003.

Seto, Yoshiya, Aika Sado, Kei Asami, Atsushi Hanada, Mikihisa Umehara, Kohki Akiyama, and Shinjiro Yamaguchi. 2014. 'Carlactone is an endogenous biosynthetic precursor for strigolactones', *Proceedings of the National Academy of Sciences*, 111: 1640.

Sorefan, K., J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, E. Foo, S. Chatfield, et al. 2003. 'MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea', *Genes Dev*, 17: 1469-74.

Suiko, M., K. Kurogi, T. Hashiguchi, Y. Sakakibara, and M. C. Liu. 2017. 'Updated perspectives on the cytosolic sulfotransferases (SULTs) and SULT-mediated sulfation', *Biosci Biotechnol Biochem*, 81: 63-72.

Takikawa, Hirosato, Kasumi Shida, Kazuma Kuwabara, Masaharu Mizutani, Moe Iseki, Takatoshi Wakabayashi, and Yukihiro Sugimoto. 2017. 'Evidence for species-dependent biosynthetic pathways for converting carlactone to strigolactones in plants', *Journal of Experimental Botany*, 69: 2305-18.

Tibbs, Zachary E., Katie Jo Rohn-Glowacki, Frank Crittenden, Amber L. Guidry, and Charles N. Falany. 2015. 'Structural plasticity in the human cytosolic sulfotransferase dimer and its role in substrate selectivity and catalysis', *Drug Metabolism and Pharmacokinetics*, 30: 3-20.

Varin, L., F. Marsolais, M. Richard, and M. Rouleau. 1997. 'Sulfation and sulfotransferases 6: Biochemistry and molecular biology of plant sulfotransferases', *Faseb j*, 11: 517-25.

Varin, Luc, and Ragai K. Ibrahim. 1989. 'Partial Purification and Characterization of Three Flavonol-Specific Sulfotransferases from *Flaveria chloraefolia*', *Plant Physiology*, 90: 977.

Visser, T. J., E. Kaptein, H. Glatt, I. Bartsch, M. Hagen, and M. W. Coughtrie. 1998. 'Characterization of thyroid hormone sulfotransferases', *Chem Biol Interact*, 109: 279-91.

Wang, R., and M. G. Brattain. 2007. 'The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa', *FEBS Lett*, 581: 3164-70.

Wang, Yanting, and Harro J. Bouwmeester. 2018. 'Structural diversity in the strigolactones', *Journal of Experimental Botany*, 69: 2219-30.

Weitzner, Brian, Thomas Meehan, Qifang Xu, and Roland L. Dunbrack Jr. 2009. 'An unusually small dimer interface is observed in all available crystal structures of cytosolic sulfotransferases', *Proteins: Structure, Function, and Bioinformatics*, 75: 289-95.

Xie, X., K. Yoneyama, T. Kisugi, K. Uchida, S. Ito, K. Akiyama, H. Hayashi, et al. 2013. 'Confirming stereochemical structures of strigolactones produced by rice and tobacco', *Mol Plant*, 6: 153-63.

Xie, Xiaonan, Kaori Yoneyama, and Koichi Yoneyama. 2010. 'The Strigolactone Story', *Annual Review of Phytopathology*, 48: 93-117.

Yoneyama, K., N. Mori, T. Sato, A. Yoda, X. Xie, M. Okamoto, M. Iwanaga, et al. 2018. 'Conversion of carlactone to carlactonoic acid is a conserved function of MAX1 homologs in strigolactone biosynthesis', *New Phytol*, 218: 1522-33.

Yoneyama, K., X. Xie, K. Yoneyama, and Y. Takeuchi. 2009. 'Strigolactones: structures and biological activities', *Pest Manag Sci*, 65: 467-70.

Zhang, Y., X. Cheng, Y. Wang, C. Diez-Simon, K. Flokova, A. Bimbo, H. J. Bouwmeester, et al. 2018. 'The tomato MAX1 homolog, SlMAX1, is involved in the biosynthesis of tomato strigolactones from carlactone', *New Phytol*, 219: 297-309.

Zhang, Y., A. D. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, et al. 2014. 'Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis', *Nat Chem Biol*, 10: 1028-33.

Chapter 4

Towards identification of the gene regulating sorgomol production in sorghum

Mahdere Z. Shimels[1], Patrick J. Rich[2], Carolien Ruyter-Spira[1], Daniel Gobena[2], Adedayo O. Adeyanju[2], Gebisa Ejeta[2], Harro Bouwmeester[1,3]

[1]Laboratory of Plant Physiology, Wageningen UR, PO Box 658, 6700 AR Wageningen, the Netherlands
[2]Department of Agronomy, Purdue University, West Lafayette, IN 47907
[3]Present address: Plant Hormone Biology group, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands.

Abstract

Strigolactones are carotenoid derived plant hormones that are also exuded into the rhizosphere where they serve as signaling molecules that activate the beneficial arbuscular mycorrhizal fungi but also stimulate germination of parasitic weeds such as *Striga hermonthica*. So far, more than 20 strigolactones have been identified and sorgomol is one that is typically found in sorghum. In this study, we used a multidisciplinary approach to elucidate sorgomol biosynthesis in sorghum. Hereto, we used Recombinant Inbred Lines (RILs) derived from parents contrasting for the presence of sorgomol. Bulks of selected high and low sorgomol producing lines were used for RNA-sequencing, which demonstrated there were several differentially expressed genes including a cytochrome P450, *Sobic.001G319900* - an enzyme that could conceivably catalyze the proposed biosynthetic step from 5-deoxystrigol to sorgomol - that displayed higher expression in the high sorgomol lines. However, genetic mapping of sorgomol production using the RIL population demonstrated a QTL on chromosome 8 and not on chromosome 1, where *Sobic.001G319900* is located. Bulk segregant analysis of re-sequenced genomic pools of RILs contrasting for sorgomol exudation narrowed the locus to a region containing *Sobic.008G106200* predicted to also encode a cytochrome P450. The likelihood of the involvement of these two candidate genes in sorgomol production is discussed.

Introduction

The strigolactones form a relatively new family of plant hormones that are produced in small amounts by plants. Inside the plant, they regulate plant architecture such as shoot branching, stem thickness and, growth of primary and lateral roots (Gomez-Roldan et al. 2008; Sun et al. 2016; Agusti et al. 2012). The amount of strigolactones produced by a plant is significantly increased when the plant is exposed to stress, especially phosphate starvation. When exuded into the rhizosphere, strigolactones serve as signaling molecules that trigger hyphal branching in arbuscular mycorrhizal (AM) fungi (Akiyama et al. 2005; Besserer et al. 2006). In this mutualistic relationship, the AM fungi enhance acquisition of nutrients, including phosphate, by the plant in return for carbon from photosynthesis (Smith et al. 2011; Harrison 2005). However, other pathogenic organisms such as the parasitic plant, *Striga hermontica*, also use strigolactones as a signal for host plant presence. Strigolactones trigger seed germination of *Striga* after which it attaches to the roots of the host plant and siphons off water, nutrients and assimilates to sustain its life cycle (Yoneyama et al. 2010; Bouwmeester et al. 2007).

More than 20 strigolactones have been identified so far (Figure 1A). Each plant species produces its own unique blend of type and amount of strigolactones. All strigolactones are derived from carotenoids. Several enzymatic steps have been postulated and/or proven to be involved in the production of strigolactones. The first steps of the biosynthesis are catalyzed by iron-binding protein DWARF27 and Carotenoid Cleavage Dioxygenases 7 and 8 (CCD7 and CCD8) producing carlactone, the central intermediate in strigolactone biosynthesis (Lin et al. 2009; Arite et al. 2007; Zou et al. 2006; Booker et al. 2004; Sorefan et al. 2003; Alder et al. 2012). From carlactone, two classes of strigolactones are derived, canonical and non-canonical strigolactones where the latter ones do not have the classical ABC ring. Examples of non-canonical strigolactones are in carlactonoic acid, methyl carlactonoate and heliolactone (Nomura et al. 2018; Al-Babili et al. 2015; Xie et al. 2013; Wang et al. 2018). However, all strigolactones have the D-ring that is linked by an enol ether bridge to the ABC-ring or the structure that replaces the ABC-ring in the non-canonical strigolactones (Umehara et al. 2015). Canonical strigolactones are grouped into two types, with α- or β-stereochemistry at the B-C ring junction, called orobanchol- and strigol-type, respectively, with 4-deoxyorobanchol and 5-deoxystrigol as the ideotype precursor of both types, respectively (Flematti et al. 2016; Xie et al. 2013; Al-Babili et al. 2015). The rice enzyme that catalyzes the conversion of carlactone to 4-deoxyorobanchol is a member of the CYP711A class of the superfamily of cytochrome P450s, Os900, a MAX1 homolog (Zhang et al. 2014). Further steps for the diversification of strigolactones can be achieved by hydroxylation, acetylation, demethylation, esterification, decarboxylation, epoxidation and oxidation of the AB-rings of 5-deoxystrigol and 4-deoxyorobanchol (Motonami et al. 2013; Al-Babili et al. 2015). The evolutionary reasons for the large structural diversity in the strigolactones are not yet fully understood. However, there is evidence that this diversity is of biological relevance. In germination assays, for example, it has been shown that not all strigolactones induce germination of all parasitic plant species with the same efficiency. For instance, sorgomol, a strigol type of strigolactone, is significantly

more active on *S. hermonthica* than on *Orobanche minor* (Xie et al. 2009; Mohemed et al. 2018).

Sorgomol was identified in the root exudate of certain varieties of sorghum, which gave it its name, as a germination stimulant of *Striga* (Xie et al. 2008). Just as the seemingly more universal strigolactone exuded by sorghum roots, 5-deoxystrigol, sorgomol has high *Striga* germination stimulant activity (Nomura et al. 2013) at nanomolar and even picomolar concentrations (Toh et al. 2015). Unlike the hydroxylated strigolactone orobanchol, sorgomol shares the β-oriented B-C-ring (strigol-type) stereochemistry of the active *Striga* germination stimulants and the hydroxyl substitution at C-9 does not render the rings more vulnerable to nucleophilic attack which might compromise its stability in the rhizosphere as expected for orobanchol and strigol where the hydroxyl is on the B- or A-ring, respectively (Yoneyama et al. 2009). As shown in Figure 1 the biosynthesis of sorgomol is postulated to proceed through 5-deoxystrigol. This was demonstrated using a feeding assay where 5-deoxystrigol was shown to be converted by sorghum to sorgomol likely by hydroxylation at C-9 (Motonami et al. 2013). Further oxidation of the hydroxyl group of sorgomol to a carboxylic acid and subsequent decarboxylation would produce sorgolactone (Motonami et al. 2013). Based on the hypothesis of sorgomol biosynthesis through hydroxylation of its postulated precursor 5-deoxystrigol at C9, the four sorghum MAX1s were our fist candidate genes for this hydroxylation reaction (Motonami et al. 2013). As shown in Chapter 3, our transient assay where we co-infiltrated the four MAX1s with the rice carlactone pathway showed a clear reduction of carlactone. However, we did not detect sorgomol or its precursor 5-deoxystrigol in any of the assays. This result suggests that the four sorghum MAX1 homologs play no role in sorgomol biosynthesis. Therefore, we used an untargeted genomics and transcriptomics approach to identify a shortlist of genes leading to two *a priori* candidate genes controlling the production of sorgomol in sorghum.

Hereto, we took advantage of a recombinant inbred population derived from a cross between two sorghum lines that contrasted for the presence of sorgomol in their root exudates. Although both parental lines produce root exudates with high *Striga* germination stimulant activity, they are distinguished by the strigolactones present in the exudates. The major strigolactone of K1597 is 5-deoxystrigol, while the root exudate of the other parent, KP33-2 is dominated by sorgomol. The recombinant inbred lines (RILs) were found to produce mainly one or the other strigolactone. By chemically phenotyping the strigolactones in the root exudates of these parental lines and their recombinant inbred derivatives, RNAseq and genotyping through genome sequencing, we attempted to find the gene controlling the supposed conversion of 5-deoxystrigol to sorgomol.

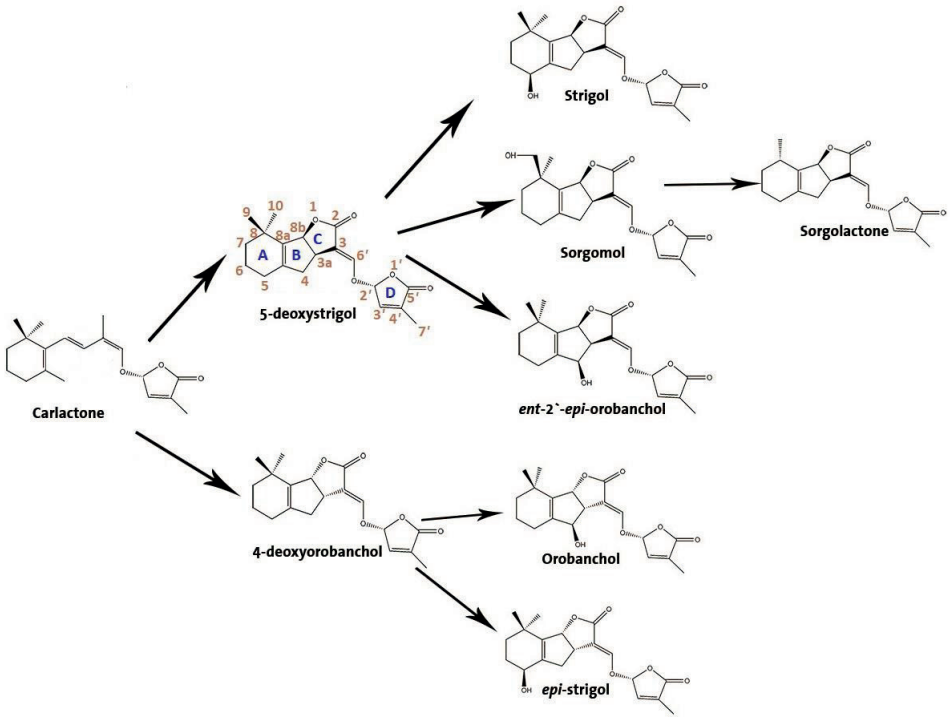


Figure 1. Schematic drawing of strigolactone biosynthesis.

The search for a candidate gene for sorgomol biosynthesis involves several assumptions. Firstly, the ability to make sorgomol is controlled in this population by one locus. Both parents and their resulting progeny have functional alleles in genes coding for the other strigolactone biosynthetic enzymes leading to 5-deoxystrigol. The allele at the "sorgomol locus" is assumed to be dominant, that is, its gene product is capable of converting 5-deoxystrigol to sorgomol and this conversion proceeds in all lines where it is present such that the RILs have either one or the other parental phenotype with respect to the dominant strigolactone in their root exudate. The protein product of the candidate gene resulting from the recessive allele cannot convert 5-deoxystrigol to sorgomol. We also assumed that the candidate gene, like other strigolactone biosynthetic genes, is expressed in roots. Furthermore, we expect that the expression of the candidate gene is higher in lines able to produce sorgomol. From earlier unpublished analyses, we found that BTx623, the sorghum genotype that was used to create the sorghum reference genome, produces small amount of strigolactones composed of orobanchol, 5-deoxystrigol, strigol and sorgomol in a ratio of 5:75:5:15%, respectively (Benjamin Thiombiano, pers. communication). Therefore, the allele at the candidate gene could be the same as in the low sorgomol RILs. We assume that some recognizable homolog of the candidate gene exists in the sorghum reference genome and that some variation with predicted functional variation in the protein for which the alternate allele codes is evident

between the resequenced parents of the population when aligned to the reference sequence. Further, those variations must correlate with all members of the derived RILs that share the chemical phenotype of their respective parents. In other words, all RILs that do not make sorgomol should share, without exception, the allele at this candidate gene with K1597, the parental line whose root exudate contains mainly 5-deoxystrigol, and all sorgomol producing RILs should have the KP33-2 allele, the functional and dominant allele. And, of course, we must be confident that the phenotype we measure, the strigolactone profile of the root exudate, is clearly and consistently expressed under the conditions of our experiments such to reliably distinguish the parental types among the RILs.

Result

Strigolactone profile of parental line Lines

We started our untargeted approach to look for the gene that converts 5-deoxystrigol to sorgomol by checking the strigolactone profile of two sorghum lines that differ strongly in sorgomol production, KP33-2 and K1597. Figure 2 shows the results of the analysis of root exudates and extracts of these two genotypes in comparison with Shanqui-Red and SRN39 that we used in Chapters 2, 3 and 4. CK32-7 is also included in this experiment because it has been shown to exhibit *Striga* resistance through a hypersensitive response, similar to KP33-2 (Mohamed et al. 2010). KP33-2 had a relatively high level of sorgomol both in root extracts and exudates when compared to the other sorghum varieties, while K1597 produced little or no sorgomol both in root extract and exudate. Both lines exuded similar amounts of 5-deoxystrigol. On the other hand, in root extracts, the level of 5-deoxystrigol was significantly higher in K1597 compared to KP33-2. Other strigolactones including orobanchol, the major strigolactone in SRN39, were below the detection level both in root exudates and extracts (Figure 2).

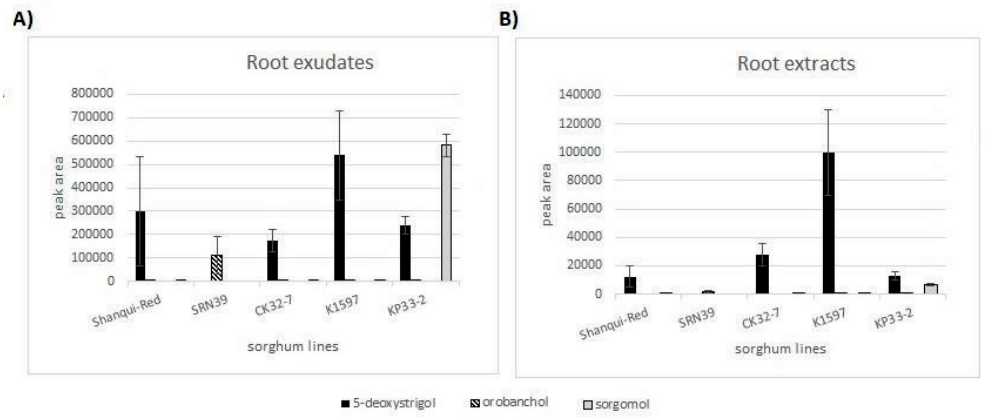


Figure 2. Strigolactone profile in root exudates and extracts of sorghum lines. One plant per pot was grown for four weeks under normal conditions (watered with ½-strength Hoagland nutrient solution) and was then phosphate starved (Watered with ½-strength Hoagland nutrient solution without phosphate)

for one week after which the root exudate was collected and analyzed on MRM-LC-MS-MS. The bar represents the standard error (n=3).

Strigolactone profile of RIL lines to determine sorgomol locus

Based on the variation in sorgomol production of the parents, the strigolactone profile of 96 RILs was checked. A large variation in sorgomol level was observed between the RILs (Figure 3). The variation between replicates was rather large, likely because the replicates were grown in different batches.

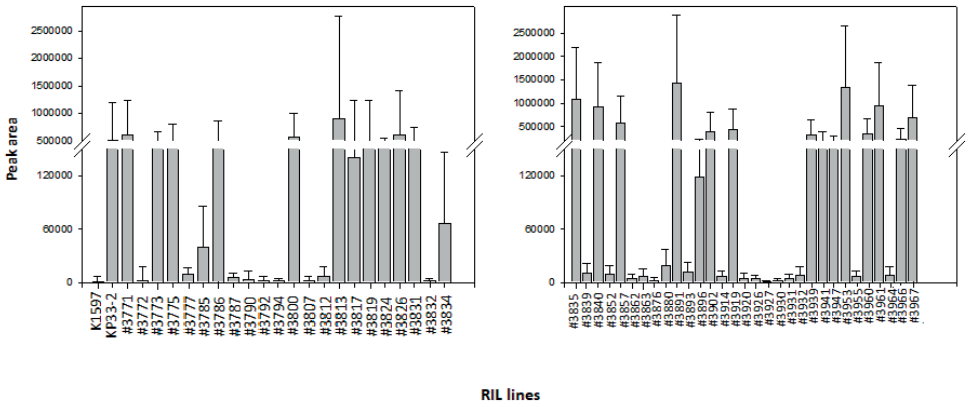


Figure 3. Sorgomol profile in root exudates of selected RIL lines and parental line. Lines used for genome sequencing are indicated with red arrows. Three plants per pot were grown for two weeks under normal conditions (watered with ½ Hoagland nutrient solutions) and were grown under phosphate starvation (Watered with ½ Hoagland nutrient solution without phosphate) for 1 week. The bar represents standard error n=3.

In addition, several other strigolactones were detected such as 5-deoxystrigol and strigol and minor amounts of *ent*-2'-*epi*-orobanchol in most samples (Table 2). All strigolactones were of the strigol-type with β-oriented C-ring. The concentration range of these strigolactones varied from 1-1148 pmol/plant for sorgomol, 6-191 pmol/plant 5-deoxystrigol and 4-544 pmol/plant strigol. Generally, those lines in which 5-deoxystrigol was the major strigolactone of the root exudate had only trace amounts of sorgomol, and in those for which sorgomol was the dominant strigolactone, 5-deoxystrigol amounts were relatively low.

Given that we also observed other hydroxylated strigolactones in the parents and RILs besides sorgomol, we had to consider alternative biosynthetic fates of 5-deoxystrigol, the presumed precursor of sorgomol, strigol and *ent*-2'-*epi*-orobanchol. The strigolactones detected by MRM in the root exudates are provided in Table 2 and are arranged in descending order based on the absolute sorgomol amount measured in the root exudates. The table is divided into two, separating the RILs based on the supposed parental types using sorgomol levels, showing the 54 HIGH and 46 LOW in the left (shaded in grey) and right side of the table, respectively, fitting the expected 1:1 ratio for parental types (by the chi-squared test) in an F2 derived recombinant inbred population in which sorgomol levels are determined by two contrasting alleles at a single locus.

In relative abundance sorgomol ranges from 17-95% of the total strigolactones. In the LOW group, sorgomol is only present at 1 to 20 pmol/plant, or 1-15% of the total strigolactones. Using the ratio of sorgomol:5-deoxystrigol also helped to distinguish the parental types. KP33-2, the sorgomol producing progenitor, has a ratio of sorgomol:5-deoxystrigol of 17 while the low sorgomol parent of the RIL population (K1597) is over 500-fold less at 0.03 for this ratio. The high sorgomol RILs varied for this ratio between 1 and 139, which is at least 100-fold higher than those of the low sorgomol RILs with a ratio between 0.01 and 0.37. The sorgomol:5-deoxystrigol ratio supports the assumption that the K1597 types share a general impairment in the ability to make sorgomol which distinguishes them from KP33-2 types.

Table 2 shows that the amount of *ent*-2'-*epi*-orobanchol in the parents and RILs is relatively low compared to the other measured strigolactones showing that hydroxylation at C-4 is less likely than at C-5 (leading to strigol) or C-9 (leading to sorgomol). Actually, the amount of *ent*-2'-*epi*-orobanchol is negligible compared to those of strigol and sorgomol and will not be considered further.

An interesting relationship between strigol and sorgomol emerges from these strigolactone data. With few exceptions, the two parental strigolactone profiles clearly emerge from the RILs. In addition to their sorgomol levels and sorgomol:5-deoxystrigol ratios, the HIGH sorgomol (KP33-2-types) are distinguished from the LOW sorgomol (K1597-types) by the ratio of sorgomol:strigol. This ratio is generally above one in the KP33-2 types and below one in the K1597 types. The ratio of strigol:5-deoxystrigol, on the other hand, does not clearly distinguish the parental types. This ratio averages 1.87 for the HIGH sorgomol RILs and 2.22 for the LOW RILs. The amount and proportion of strigol along with the sorgomol:strigol ratio in the root exudates of the two parental types, implies that conversion of 5-deoxystrigol to strigol occurs at a similar rate in both groups, though the amount that is converted is generally higher in the LOW sorgomol group.

Table 2. Strigolactone profile in root exudates of high and low sorgomol producing lines.

Line	Absolute abundance of measured strigolactones (pmol/plant/48hrs)					Line	Absolute abundance of measured strigolactones (pmol/plant/48hrs)				
	5-deoxystrigol	*ent*-2'-*epi*-orobanchol	strigol	sorgomol	total		5-deoxystrigol	*ent*-2'-*epi*-orobanchol	strigol	sorgomol	total
3966	144	3	317	1148	1612	**3880**	79	2	183	20	284
3821	51	0	4	1131	1186	**3936**	101	5	449	13	568
3780	85	0	95	1096	1276	**3856**	90	2	95	10	197
3942	47	0	66	987	1100	**3778**	229	4	358	9	600
3967	26	1	86	934	1047	**3828**	88	4	247	8	347
3771	142	5	430	850	1427	**3772**	161	6	544	7	718
3953	40	2	116	818	976	**3964**	96	4	467	7	574
3817	54	1	92	792	939	**3858**	191	4	189	6	390
3813	59	1	89	779	928	**3790**	142	3	314	5	464
3827	51	2	101	742	896	**3803**	21	0	7	5	33
3952	85	1	117	678	881	**3806**	149	6	441	5	601
3786	27	0	40	662	729	**3815**	139	2	143	5	289
3869	5	5	5	640	655	**3852**	93	3	220	5	321
3943	72	1	61	629	763	**3893**	98	3	155	5	261
KP33-2	36	1	6	593	636	**3920**	87	3	255	5	350
3857	24	1	40	580	645	**K1597**	124	3	11	4	142
3957	79	3	188	556	826	**3777**	130	2	245	4	381
3840	65	1	82	541	689	**3804**	114	6	412	4	536
3795	43	1	55	530	629	**3812**	112	3	262	4	381
3835	44	1	63	519	627	**3832**	93	2	157	4	256
3842	17	1	94	500	612	**3839**	54	2	167	4	227
3829	23	0	26	465	514	**3914**	100	5	340	4	449
3960	103	4	207	463	777	**3926**	102	4	265	4	375
3818	28	1	53	437	519	**3931**	61	1	107	4	173
3819	22	1	38	419	480	**3932**	94	3	223	4	324
3961	26	0	31	392	449	**3959**	141	7	470	4	622
3826	86	3	337	391	817	**3787**	80	2	137	3	222
3935	48	1	79	387	515	**3862**	117	3	262	3	385
3800	26	0	39	365	430	**3863**	63	2	131	3	199
3896	8	0	31	363	402	**3933**	42	1	124	3	170
3860	22	0	15	354	391	**3955**	34	1	94	3	132
3903	14	0	45	351	410	**3962**	65	2	137	3	207
3928	27	1	52	329	409	**3792**	62	2	121	2	187
3891	15	0	39	314	368	**3794**	89	3	158	2	252
3773	33	1	43	307	384	**3807**	97	3	216	2	318
3919	11	0	82	282	375	**3845**	72	3	186	2	263
3775	24	1	23	256	304	**3925**	6	0	7	2	15
3913	19	1	32	235	287	**3930**	169	2	213	2	386
3796	42	1	58	225	326	**3782**	132	1	112	1	246
3831	30	0	55	212	297	**3814**	8	0	1	1	10
3866	14	14	14	190	232	**3870**	12	12	12	1	37
3902	10	0	10	178	198	**3876**	15	1	73	1	90
3941	97	3	191	165	456	**3927**	41	1	118	1	161
3939	18	0	27	161	206						
3909	22	0	13	152	187						
3824	51	2	117	120	290						
3785	84	3	266	117	470						
3776	8	0	20	99	127						
3947	62	2	218	91	373						
3904	13	0	13	85	111						
3875	6	6	6	83	101						
3774	39	0	10	65	114						
3834	6	0	7	65	78						
3924	101	1	85	37	224						
3882	6	0	9	31	46						

Differentially expressed genes for sorgomol production in sorghum

Earlier work suggested that the candidate gene for sorgomol production is a cytochrome P450 capable of oxidizing C-9 of 5-deoxystrigol (Motonami et al., 2013). Therefore, we decided to use RNAseq and genomic mapping under the assumption that the expression of the "sorgomol gene" would be noticeably higher among HIGH sorgomol RILs. Six high and six low sorgomol producing lines were selected for RNAseq (Table 1). The RNA was isolated from individual root samples. For each group, two biological replicates were formed by combing three RNA samples. The RNA from each parent was used as a third replicate. The data were aligned against the publicly available *Sorghum bicolor* v3.1 genome sequence for the first part of the analysis.

Normalized-RPKM (reads per kilobase per million) values were used to extract differentially expressed genes. First, to validate the reproducibility of the data between replicates and to decide the cut-off for the RPKM values, RPKM values of candidate strigolactone biosynthesis genes were checked. These genes were not differentially expressed between HIGH and LOW bulks. A putative D27 homolog, *Sobic.009G030800*, had a similar RPKM value ~3 in both HIGH and LOW bulks. *CCD7* (*Sobic.006G170300*) had an average RPKM value of 17.23 and 18.73 in HIGH and LOW bulk respectfully whereas *CCD8d* (*Sobic.002G168800*) had 21.5 and 30.56 RPKM values in HIGH and LOW bulk, respectively. Of the MAX1 homologs, *Sb2210* (*Sobic.003G269500*) had the highest RPKM value of 152.89 and 218.72 in the HIGH and LOW bulk, respectively. The other MAX1 homologs, *Sb7880* (*Sobic.004G095500*), *Sb2310* (Sobic.010G170400) and *Sb2220* (*Sobic.003G269600*), exhibited RPKM values ranging between 1.71 - 9.50 with no significant difference between the HIGH and LOW bulk. *Sb2310* had a low RPKM value around 2. Based on these results, RPKM values ≥2 were selected for further analysis (Supplementary Table 1). RPKM values were reproducible between replicates.

Out of the total 18,130 genes expressed, 69 genes were down-regulated in HIGH RILs and 57 up-regulated (Figure 4). Out of these 57, only 32 were up-regulated ≥ 2-fold in the high sorgomol producing lines with reproducibility between the replicates (Table 3). We focused on genes that are only up-regulated in high sorgomol producing lines based on the hypothesis that the increase in the production of sorgomol likely correlates with higher expression of the candidate gene in these lines. Interestingly, none of the four sorghum *MAX1*s were among these candidate genes since there was no significant difference in expression of these genes between high and low sorgomol producing lines (Supplementary Table 2). We also checked the expression of other putative strigolactone biosynthetic genes and none of these showed differential expression between high and low sorgomol producing lines (Supplementary Table 2).

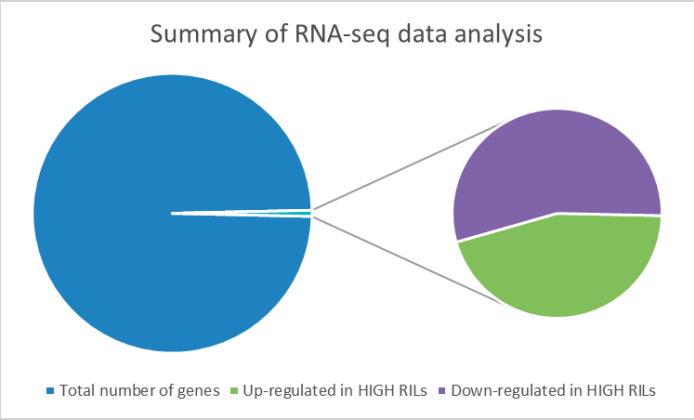


Figure 4. Summary of RNA-seq data analysis. Venn diagram representation of total number of genes expressed (18,130) (in blue), genes upregulated in HIGH RILs (57) (in green) and down-regulated genes in HIGH RILs (69) (in purple). A cut-off value of ≥2 RPKM in one of the replicates was applied. RPKM values with P-value <0.05 was considered significantly different.

To narrow down the list of candidate genes, a targeted approach was used by looking at the functional annotation of each gene (Table 3). Most genes did not have a functional annotation, but there was *Sobic.001G319900*, a putative cytochrome P450-like protein, and this was selected as an *a priori* candidate gene for further analysis. The RPKM value of this gene was 5-fold higher in the HIGH bulk when compared to the LOW bulk including the parental lines (P= 0.000985).

We sequenced the genome of a pool of high and low sorgomol producing lines in an attempt to search for more evidence that supports the involvement of this gene in regulating sorgomol production. Lines used for RNAseq were also included in the list selected for genome sequencing (Table 1). The consensus sequences of the LOW and HIGH bulks and KP33-2 and K1597 were aligned with the BTx623 genome sequence. There was a three-nucleotide addition in the TATA-box of the promotor region of *Sobic.001G319900* in the HIGH bulk consensus sequence and KP33-2 (Figure 5). Since this region is the core region of the promotor which is involved in the regulation of transcription, the addition of nucleotides that changes the distance between the TATA-box and cap site could possibly cause differential expression. A maize mutant with a duplication or deletion of the TATA box displayed reduced expression of *alcohol dehydrogenase-1* (*Adh*) in an organ-specific manner (Kloeckener-Gruissem et al. 1992). However, although it was present in the consensus sequence, this sequence variation was not consistently present in all HIGH genome sequences. There were no other consistent differences in the promoter or coding sequence of the gene that could support its candidacy. To look for additional candidate genes, therefore, we decided to use the sorgomol exudation data obtained for the RILs to map the sorgomol level to the sorghum genome.

Table 3. List of genes that are differentially expressed between low and high sorgomol producing lines. The RPKM value of the genes was higher in high sorgomol producing lines and the fold change is calculated from the average of the three replicates.

No	Feature ID	Description	Fold change	P-value
1	*Sobic.008G109400*	No annotated domains	4.96	0.00098
2	*Sobic.005G208100*	K04733 - interleukin-1 receptor-associated kinase 4 (IRAK4)	4.87	0.00082
3	*Sobic.008G009400*	No annotated domains	3.11	0.0066
4	*Sobic.008G106800*	Similar to Isoflavone reductase, putative	6.23	0.028785
5	*Sobic.008G090400*	No annotated domains	9.05	5.75554E-05
6	*Sobic.001G139900*	PTHR12064//PTHR12064:SF30	2.98	2.90211E-08
7	*Sobic.005G204900*	Similar to Lipase family protein; CGI-141-RELATED/LIPASE CONTAINING PROTEIN	2.80	0.001468035
8	*Sobic.008G094900*	No annotated domains	9.47	3.22644E-06
9	*Sobic.008G090300*	No annotated domains	9.10	8.64608E-15
10	*Sobic.008G065932*	No annotated domains	8.96	2.069E-14
11	*Sobic.005G071100*	K04802 - proliferating cell nuclear antigen (PCNA)	5.88	2.57528E-07
12	*Sobic.002G427000*	No annotated domains		
13	*Sobic.008G065400*	No annotated domains	7.63	5.74627E-08
14	*Sobic.001G165000*	PTHR31920:SF20 - B3 DOMAIN-CONTAINING PROTEIN REM23	21.57	1.6770E-12
15	*Sobic.001G152280*	No annotated domains	8.03	3.81449E-05
16	*Sobic.007G170700*	Similar to Putative fatty acyl coA reductase	4.79	7.89265E-07
17	*Sobic.008G069700*	No annotated domains	7.00	2.69534E-07
18	*Sobic.008G079032*	No annotated domains	6.58	3.96403E-05
20	*Sobic.006G041266*	Glycerol-3-phosphate 1-O-acyltransferase / Glycerol-3-phosphate O acyltransferase	5.95	0.000173011
21	*Sobic.001G152320*	No annotated domains	44.26	5.25114E-08
22	*Sobic.001G319900*	similar to Cytochrome P450-like protein, orthologues AT2G45510, OS10G0525000, OS10G0525200, OS10G0524700	4.96	0.000985093
23	*Sobic.005G123700*	No annotated domains	112.52	3.30822E-08
24	*Sobic.001G288301*	No annotated domains	3.73	0.039404307
25	*Sobic.008G067301*	No annotated domains	8.27	7.54664E-05
26	*Sobic.008G112100*	Similar to 1-acyl-sn-glycerol-3-phosphate acyltransferase PLS1	6.31	0.00085996
27	*Sobic.008G079100*	No annotated domains	7.22	0.00093722
28	*Sobic.008G090100*	No annotated domains	37.38	0.045372398
29	*Sobic.005G065200*	No annotated domains	7.63	0.001468035
30	*Sobic.008G105850*	No annotated domains	8.33	0.001837077
31	*Sobic.006G101200*	No annotated domains	233.51	3.36048E-05
32	*Sobic.006G092000*	No annotated domains	50.56	0.012722075
33	*Sobic.001G163000*	No annotated domains	45.81	0.00440565

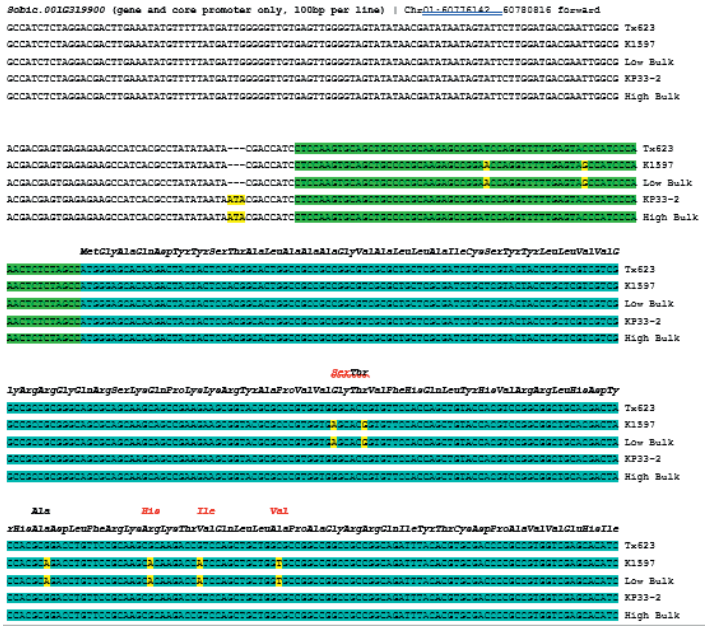


Figure 5. Alignment of genomic sequence to the core promoter, 5'-UTR (highlighted in green) and beginning of first exon (blue highlight) of *Sobic.001G319900*. Deviations from the reference genome are highlighted in yellow. The sequences are ordered from top to bottom as BTx623 (reference), K1597 (low sorgomol producing parent), LOW bulk consensus, KP33-2 (high sorgomol producing parent) and HIGH bulk consensus.

Molecular mapping for identifying the sorgomol locus

In order to map this locus to a specific chromosome, we used a set of short sequence repeat (SSR) markers with known genomic locations to genotype the RILs for which we obtained strigolactone profiles. We first screened K1597 and KP33-2 with SSRs to determine which gave visually scorable size polymorphisms after electrophoresis of PCR products on agarose gel. From this screen, we identified 30 polymorphic SSRs which we then used to genotype the population. With this initial set, we found apparent linkage of the KP33-2 allele at marker *Xtxp354* on Chromosome 8 (position 55, 558, 543 v3.1) with the HIGH sorgomol phenotype and thereafter added 14 additional polymorphic SSRs on Chromosome 8 to confirm this observation. The KP33-2 allele at the additional markers was present in all HIGH sorgomol RILs from Chr08:54, 320, 841-56, 617, 759 (v3.1) but the co-segregation began to break beyond position 57,000,000 and the HIGH sorgomol trait segregated nearly independently from the *Xtxp28* marker located near the beginning of Chromosome 8 (3, 852, 232). Although the physical interval of the QTL is quite large, the evidence gathered by these associations for linkage to Chromosome 8 is compelling. A summary of the SSR genotyping of those RILs eventually used for bulk sequencing is provided in Table 4 and the entire set of genotyped RILs in Supplemental Table 1. In these tables, the phenotype calls and alleles at each marker associated with the two parental types are indicated as 'B' for the HIGH sorgomol parent (KP33-2) or 'A' for the LOW sorgomol parent (K1597).

SSD7 RIL#	2013 row	Genotype calls from SSR markers polymorphic between K1597 (LOW sorgomol parent; A alleles) and KP33-2 (HIGH sorgomol parent, B alleles)																							
		Chr01				Chr02						Chr03		Chr04			Chr05					Chr06		Chr07		
		Xtxp71 (Chr01_6,677,622)	Xtxp11 (Chr01_52,962,862)	Xtxp279 (Chr01_61,146,468)	Xtxp284 (Chr01_65,087,651)	Xtxp197 (Chr02_1,449,033)	Xtxp297 (Chr02_4,229,787)	Xtxp4 (Chr02_13,161,049)	Xtxp298 (Chr02_57,050,889)	Xtxp56 (Chr02_61,503,546)	Xtxp8 (Chr02_75,906,289)	Xtxp285 (Chr03_67,805,717)	Xtxp98 (Chr03_73.959.820)	Xcup05 (Chr04_3,852,180)	Xtxp343 (Chr04_48,757,775)	Xtxp29 (Chr04_57,546,300)	Xtxp303 (Chr05_5,686,170)	Kaf2 (Chr05_57,781,081)	SBKAFGK1 (Chr05_57,781,085)	PDStriga1b (Chr05_60,300,767)	SSR1_UP26580 (Chr05_60,447,508)	Xtxp45 (Chr06_49,269,476)	Xtxp265 (Chr06_51,167,019)	Xtxp312 (Chr07_4,522,365)	Sb6_342 (Chr07_37,531,281)	Xcup57 (Chr07_64,091,798)
KP33-2		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
29	#3821	B	B	A	B	B	B	B	B	B	B	B	B	B	B	A	A	B	B	A	A	A	A	A	A	A
8	#3780	B	B	B	A	A	A	B	B	B	A	A	A	B	A	B	B	B	B	B	B	B	B	A	A	A
79	#3942	A	A	A	B	A	A/B	A	A	B	A	B	B	A	B	A	B	_	B	B	B	A	A	A	B	B
83	#3953	B	A	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	B	A
26	#3817	B	A	B	A	B	A	B	A	B	A	A	B	B	B	B	A	A	A	A	A	B	B	A	B	A
23	#3813	A	B	B	B	B	B	B	B	B	A	B	B	A	B	A	B	B	B	B	B	B	B	A	A	A
32	#3827	B	B	A	B	B	B	B	B	B	_	A	A	A	B	B	B	A	A	A	A	B	B	A	B	B
82	#3952	A	B	B	B	A	A	A	A	B	B	A	A	A	A	B	A	B	B	B	B	A	B	B	A	B
11	#3786	B	B	B	B	B	B	A	A	B	B	A	B	B	B	B	B	A	A	A	A	A	A	A/B	A	A
50	#3869	B	B	B	B	A	A	A	B	A	A	B	B	A	B	A	B	B	B	A	A	A	A	B	B	B
45	#3857	A	B	_	A	B	B	B	B	B	A	B	A	A	B	B	A	B	B	A	B	A	A	A	B	A
85	#3957	A	A	A	A	B	B	B	B	B	A	A	B	A	B	B	A	B	A	A	A	A	A	A	A	A
K1597		A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
73	#3932	B	A	A	A	A	B	B	B	B	B	B	B	A	A	B	B	_	A	A	A	A	A	A	B	A
7	#3777	A	A	A/B	A/B	A	A	A	A	B	B	B	B	B	B	A	A	B	B	B	B	A	A	A	A	A
86	#3959	B	A	A	A	A	A	A	A	A	A	A	A	B	B	A	B	B	B	A	A	A	A	B	A	A
22	#3812	A	B	B	B	A	A	A	A	B	A	A	B	A	A	A	B	A	A	A	A	A	A	B	B	A
19	#3804	A	A	B	B	A	A	A	A	A	A	A	A	A	A	B	B	A	A	B	B	B	B	B	A	B
36	#3832	A	A	B	B	A	A	B	B	B	A	B	B	B	B	A	A	A	A	A	A	A	A	A	A	B
63	#3914	A	B	A	B	B	B	B	B	A	A	A	A	B	B	B	B	B	B	A	A	B	B	B	A	A
68	#3926	B	B	A	A	B	B	A	B	A	B	A	A	A	B	B	A	A	A	A	A	B	B	A	A	A/B
48	#3862	A	B	_	B	B	A	A	B	A	B	A	A	B	B	A	B	B	B	B	B	A	A	B	A	B
21	#3807	B	A	B	B	A	A	B	A	A	A	A	B	A	B	B	B	A	A	A	A	A	A	B	A	A
71	#3930	B	B	B	B	B	B	A	A	A	A	B	B	B	B	B	A	B	B	B	B	A	A	A	B	B
9	#3782	B	A	B	B	B	B	B	B	B	A	A	A	B	B	A	A	A	A	_	A	A	A	B	B	B

SSD7 RIL#	2013 row	Chr08																Chr09	Chr10		Phenotype Call (B=HIGH sorgomol; A=LOW sorgomol)	Sorgomol (pMol/plant/48hrs)	5-deoxystrigol (pMol/plant/48hrs)	Sorgomol : 5-deoxystrigol
		Xtxp28 (Chr08_3,824,703)	SB4510 (Chr08_46,965,912)	SB4511 (Chr08_46,975,858)	SB4513 (Chr08_47,243,683)	Ch08_TG(26) (Chr08_47,318,464)	Ch08_TA(15) (Chr08_47,319,133)	Ch08_AT(59) (Chr08_47,338,690)	SB4516 (Chr08_47,409,902)	Xtxp354 (Chr08_48,122,810)	SB4524 (Chr08_48,283,763)	SB4528 (Chr08_48,961,423)	SB4529 (Chr08_49,026,500)	SB4531 (Chr08_49,265,554)	SB4534 (Chr08_49,403,512)	SB4536 (Chr08_49,448,160)	SB4540 (Chr08_49,794,898)	Xtxp107 (Chr09_57,746,020)	Xcup49 (Chr10_233,351)	Xtxp331 (Chr10_11,673,195)				
KP33-2		B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	593	36	17
29	#3821	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	1131	51	22
8	#3780	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	1096	85	13
79	#3942	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	987	47	21
83	#3953	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	818	40	20
26	#3817	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	792	54	15
23	#3813	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	779	59	13
32	#3827	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	742	51	15
82	#3952	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	678	85	8
11	#3786	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	662	27	25
50	#3869	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	B	A	B	640	5	139
45	#3857	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	580	24	25
85	#3957	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	556	79	7
K1597		A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	4	124	0.03
73	#3932	B	A	A	A	A	A	A	A	A	A	B	B/A	B	B	B	B	A	B	A	A	4	94	0.05
7	#3777	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	A	A	A	4	130	0.03
86	#3959	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	4	141	0.03
22	#3812	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	A	4	112	0.04
19	#3804	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	4	114	0.03
36	#3832	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	4	93	0.04
63	#3914	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	4	100	0.04
68	#3926	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	A	A	4	102	0.03
48	#3862	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	3	117	0.03
21	#3807	B	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	A	2	97	0.02
71	#3930	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	B	A	A	2	169	0.01
9	#3782	_	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	B	B	A	A	1	132	0.01

The linkage to Chromosome 8 was also evident from the bulk segregant analysis on the genomic DNA pools of 12 HIGH and 12 LOW sorgomol RILs already mentioned above. Initial single nucleotide polymorphisms (SNPs) and indels between the HIGH and LOW pools compared by Popoolation2 (Kofler et al. 2011) and QTL-seq (Takagi et al. 2013) software on reads mapped to either the BTx623 reference (Phytozome 12, *Sorghum bicolor* v3.1.1) or SbRIO (Phytozome 12, *Sorghum bicolor* Rio v2.1) indicated multiple significant variations between the pools on Chromosome 8, but these variations occurred over spans too large to search for candidate genes controlling sorgomol production. We used SbRIO because we knew from previous strigolactone analysis of BTx623 root exudates that this line produces just a low level of sorgomol. We reasoned that genes controlling sorgomol production might be completely absent from the BTx623 genome analogous to the *LGS1* story presented in Chapter 2. Having an independently assembled sorghum reference genome would at least diminish that chance, though we have not looked at the strigolactones in Rio.

We struggled with narrowing down candidates. Therefore, we re-analyzed the genomic sequence data by aligning the reads of the HIGH and LOW bulks to an "alternate reference" created by The Fasta Alternate Reference Maker (GATK version 3.8-0) that replaced the reference bases at variation sites between the standard sorghum reference genomes (BTx623 and Rio) with the parental bases in areas where neither KP33-2 nor K1597 matched the standard reference assemblies. The Illumina HiSeq paired-end sequencing reads of the LOW and HIGH sorgomol bulks were then mapped to this alternate reference created for each parent and variant calls were computed against the alternate reference Fasta. This method of mapping the SNPs between KP33-2 and K1597 on the sorghum genome assembly reduced the contrasts between the HIGH and LOW bulks to a few significant loci (Figure 6). The QTLs defined by the bulk segregant analysis with the alternate reference based on the RIL parents BTx623 illustrated in Figure 6 are most significant on Chromosomes 1, 3 and 8. The plot of Δ(SNP- index) on Chromosomes 1 and 3 cross the 95% confidence interval threshold while that on Chromosome 8 crosses the 99% confidence interval.

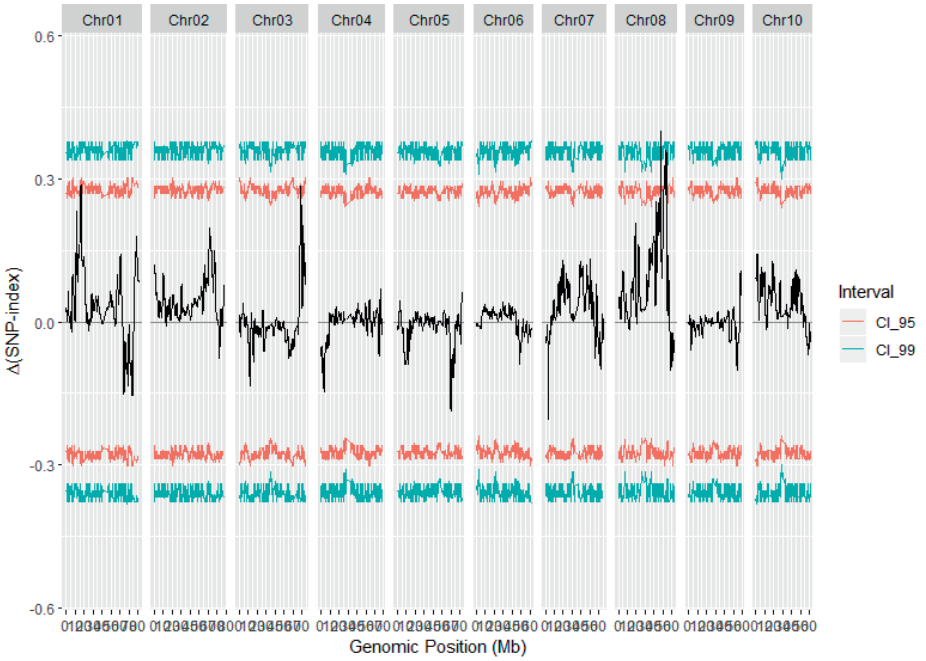


Figure 6. Quantitative trait loci (QTL) for the sorgomol locus identified by QTLseqr. Plots produced by the plotQTLStats() function with a 1 Mb sliding window: Distribution of single nucleotide polymorphisms (SNPs) in each smoothing window. The tricube smoothed Δ(SNP-index) and corresponding two-sided confidence intervals: 95% (red), and 99% (green).

Because we made the assumption that the "sorgomol gene" for which we searched likely coded for a cytochrome P450, based on the work of Motonami et al. (2013), we looked for genes under the peaks in Figure 6 whose products were annotated as cytochrome P450s and found such candidates at all three loci. These candidates are listed in Table 5 and include one on Chromosome 1, *Sobic.001G137900*, two on Chromosome 3, *Sobic.003G360600* and *Sobic.003G360900*, and one on Chromosome 8, *Sobic.008G106200*. The differentially expressed candidate P450 obtained from the RNAseq comparison described above, *Sobic.001G319900*, does not locate to any of these QTLs. We further assumed that the candidate gene would be expressed in sorghum roots since strigolactone biosynthesis occurs there (Sorefan et al. 2003). All candidates are highly expressed in sorghum roots according to the MOROKOSHI sorghum transcriptome database (Makita et al. 2015), except for *Sobic.003G360900* which only displays low expression in the roots. We further expect serious mutations (frameshifts, deletions or insertions or significant amino acid changes) in the candidate gene coding regions in the lines unable to hydroxylate 5-deoxystrigol to sorgomol (K1597 and the LOW bulk RILs) compared with the lines with sorgomol in their root exudates (KP33-2 and HIGH bulk RILs). Table 5 lists all the exon variations for each candidate gene and their potential function altering mutations, the most numerous and serious occurring in the Chromosome 8 candidate, *Sobic.008G106200*.

Table 3. Cytochrome P450 candidate genes with the amino acid function and changes within the QTLs identified for sorgomol

CHR	REF	ALT	GENEID	Predicted gene function/EXPRESSION	REF codon	VAR codon	REF amino acid	VAR amino acid	CONSEQUENCE	Amino acid change	
Chr01	T	A	*Sobic.001G137900*	Cytochrome P450 CYP2 subfamily HIGH EXPRESSION IN ROOTS*	TTC	TAC	F	Y	nonsynonymous	nonpolar (hydrophobic)c	polar (hydrophilic)
Chr03	T	C	*Sobic.003G360600*	Fatty acid omega-hydroxylase	ATG	ACG	M	T	nonsynonymous	nonpolar (hydrophobic)	polar (hydrophilic)
Chr03	A	C	*Sobic.003G360600*	Cytochrome P450	GAG	GCG	E	A	nonsynonymous	negative (hydrophilic)	nonpolar (hydrophobic)
Chr03	A	G	*Sobic.003G360600*	HIGH EXPRESSION IN ROOTS	ATC	GTC	I	V	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr03	A	G	*Sobic.003G360900*		ATC	GTC	I	V	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr03	A	G	*Sobic.003G360900*	Cytochrome P450 CYP2 subfamily	ATC	GTC	I	V	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr03	C	T	*Sobic.003G360900*	LOW EXPRESSION IN ROOTS	CTC	TTC	L	F	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr03	A	C	*Sobic.003G360900*		AAG	ACG	K	T	nonsynonymous	positive (hydrophilic)	polar (hydrophilic)
Chr08	T	C	*Sobic.008G106200*		ATC	CTC	I	L	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr08	C	A	*Sobic.008G106200*	Cytochrome P450	GCA	ACA	A	T	nonsynonymous	nonpolar (hydrophobic)	polar (hydrophilic)
Chr08	C	A	*Sobic.008G106200*	Cytochrome P450 CYP2 subfamily	GTC	ATC	V	I	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr08	T	T	*Sobic.008G106200*		AAG	ATG	K	M	nonsynonymous	positive (hydrophilic)	nonpolar (hydrophobic)
Chr08	G	G	*Sobic.008G106200*		TCC	TGC	S	C	nonsynonymous	polar (hydrophilic)	polar (hydrophilic)
Chr08	A	A	*Sobic.008G106200*		TCC	ACC	S	T	nonsynonymous	polar (hydrophilic)	polar (hydrophilic)
Chr08	ACTT	T	*Sobic.008G106200*	HIGH EXPRESSION IN ROOTS	GAAGTG	GTG	EV	V	codon deletion		
Chr08	C	C	*Sobic.008G106200*		GGT	CGT	G	R	nonsynonymous	nonpolar (hydrophobic)	positive (hydrophilic)
Chr08	A	G	*Sobic.008G106200*		TTT	TTG	F	L	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr08	A	G	*Sobic.008G106200*		TTT	TGT	F	C	nonsynonymous	nonpolar (hydrophobic)	polar (hydrophilic)
Chr08	T	G	*Sobic.008G106200*		ATC	GTC	I	V	nonsynonymous	nonpolar (hydrophobic)	nonpolar (hydrophobic)
Chr08	T	C	*Sobic.008G106200*		AGT	CGT	S	R	nonsynonymous	polar (hydrophilic)	positive (hydrophilic)
Chr08	CAGT	G	*Sobic.008G106200*		ACTGGT	GGT	TG	G	codon deletion		
Chr08	CCAG	G	*Sobic.008G106200*		TCTGGT	TGT	SG	C	codon deletion		

*Expression according to MOROKOSHI sorghum transcriptome database (Makita et al. 2014) and Phytozome.

As stated in the Introduction, another assumption guiding our candidate gene search is that variations between parental alleles of the candidate gene must be shared by all members of their respective bulks. That is, all RILs of the LOW bulk must have the K1597 allele of the "sorgomol gene" and all HIGH bulk RILs must carry the KP33-2 allele. We were able to eliminate the candidate genes on Chromosomes 1 and 3 by comparing the consensus sequence of the bulks with the sequence of their respective parents. The variations in *Sobic.001G137900* between K1597 and KP33-2 were not consistently shared with reads aligning to this locus from their respective re-sequenced bulks (Figure 7). Figure 7 shows the area where exonic variations occur in *Sobic.001G137900* at the K1597 allele with respect to KP33-2 by Integrated Genomics Viewer (Robinson et al., 2011). Although several of the reads from the LOW bulk share these variations, some do not. Also, some reads from the HIGH bulk contain this mutation. The same occurs at variable regions between parental alleles at *Sobic.003G360600* (Figure 8) and *Sobic.003G360900* (Figure 9). When the reads from the re-sequenced bulks and parent lines are viewed by this software at *Sobic.008G106200*, all bulked reads match their respective parents (Figure 10) indicating that all HIGH bulk RILs contain the KP33-2 allele and all LOW bulk RILs contain the K1597 allele. The full alignment of the consensus sequence from the bulks with the parent lines and reference genomes over the exons of *Sobic008G.106200* are shown in Figure 10 and their predicted proteins are compared in Supplementary Figure 2. The KP33-2 allele matches that carried by the BTx623 reference across the entire gene. The one exception is a SNP at position ~49,983,339 between the second and third introns of the annotated 5'-UTR, shown at the end of Supplementary Figure 1. This difference would not likely cause a difference in gene function or expression. The K1597 allele, in contrast, has multiple variations from BTx623 at this gene within the coding region, introns and UTRs. This allele more closely matches the Rio reference (Supplementary Figure 1). Its predicted product has one more amino acid residue than that of Rio (Supplementary Figure 2) but shares the other deletion near the beginning of the first exon as well as single nucleotide variations throughout the coding sequence as Rio (Supplementary Figure 1).

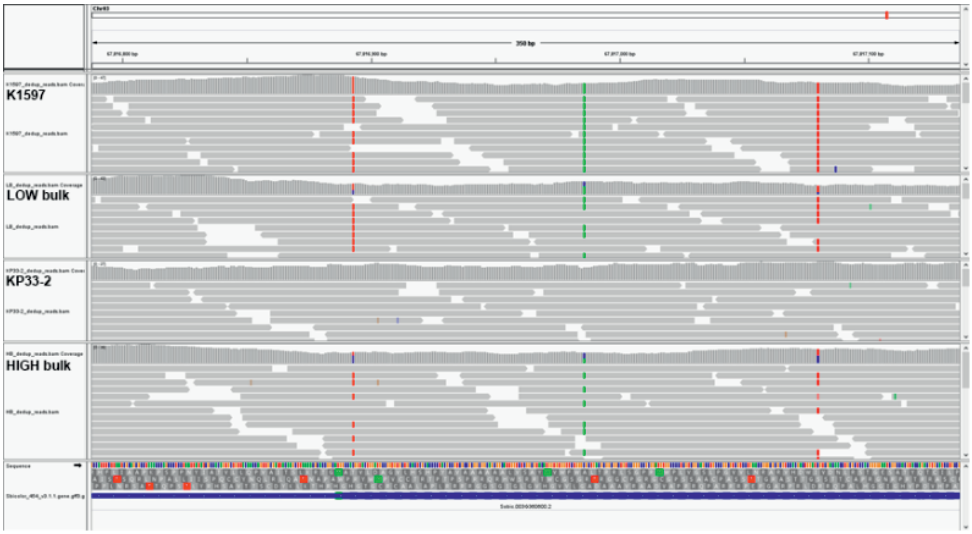


Figure 7. Sequence reads aligned to variant area in *Sobic.001G137900* Chr01:10,950,319..10,951,080 from K1597, LOW bulk, KP33-2 and HIGH bulk. Note the presence of K1597 variations in some reads of the HIGH bulk and absence of these in some LOW bulk reads. This indicates individual RILs in the LOW bulk contain the KP33-2 allele and some RILs in the HIGH bulk carrying the K1597 allele, and therefore this is not a good candidate gene. Reads from whole genome sequences viewed with IGV software (Broad Institute, Robinson et al., 2011).

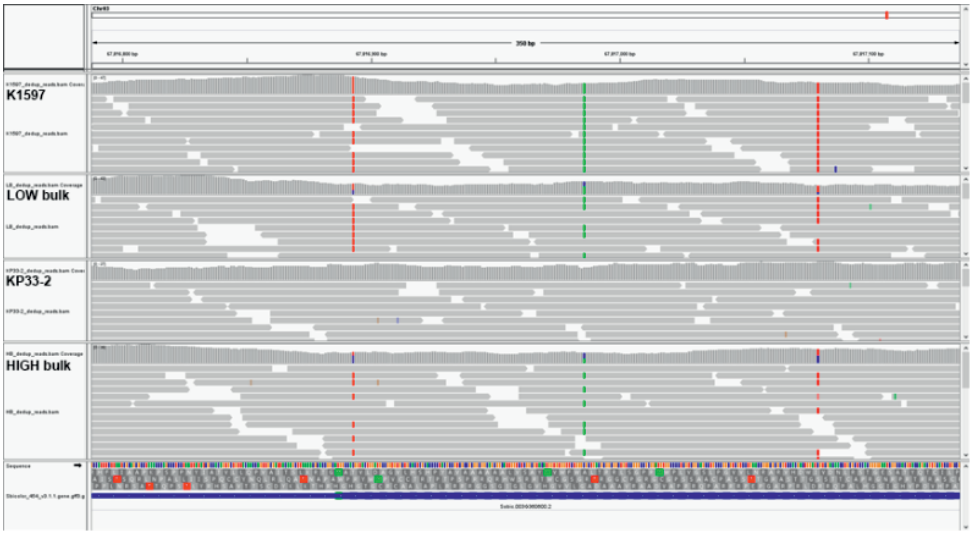


Figure 8. Sequence reads aligned to variant area in *Sobic.003G360600* Chr03:67,816,800..67,817,100 from K1597, LOW bulk, KP33-2 and HIGH bulk. Note the presence of K1597 variations in some reads of the HIGH bulk and absence of these in some LOW bulk reads. This indicates individual RILs in the LOW bulk contain the KP33-2 allele and some RILs in the HIGH bulk carrying the K1597 allele, and therefore this is not a good candidate gene. Reads from whole genome sequences viewed with IGV software (Broad Institute, Robinson et al., 2011).

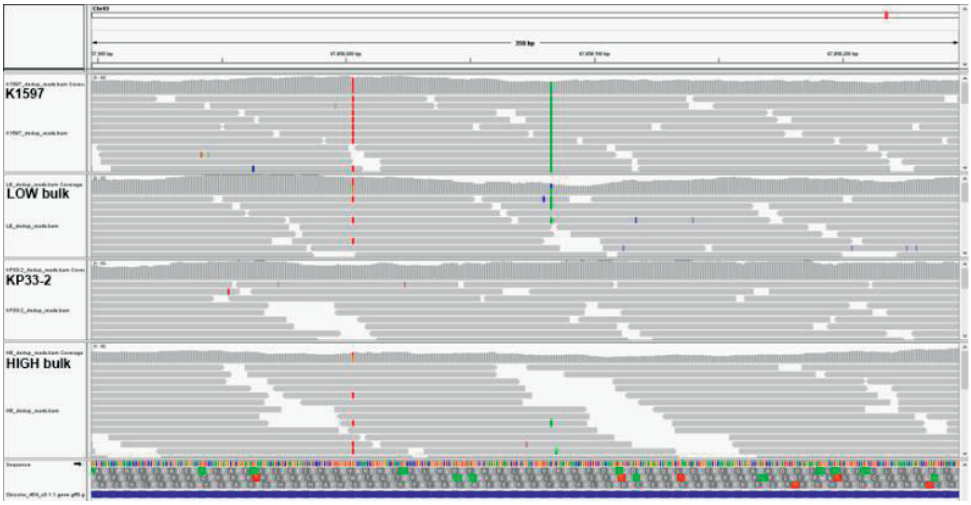


Figure 9. Sequence reads aligned to variant area in *Sobic.003G360900* Chr03:67,857,900..67,858,200 from K1597, LOW bulk, KP33-2 and HIGH bulk. Note the presence of K1597 variations in some reads of the HIGH bulk and absence of these in some LOW bulk reads. This indicates individual RILs in the LOW bulk contain the KP33-2 allele and some RILs in the HIGH bulk carrying the K1597 allele, and therefore this is not a good candidate gene. Reads from whole genome sequences viewed with IGV software (Broad Institute, Robinson et al., 2011).

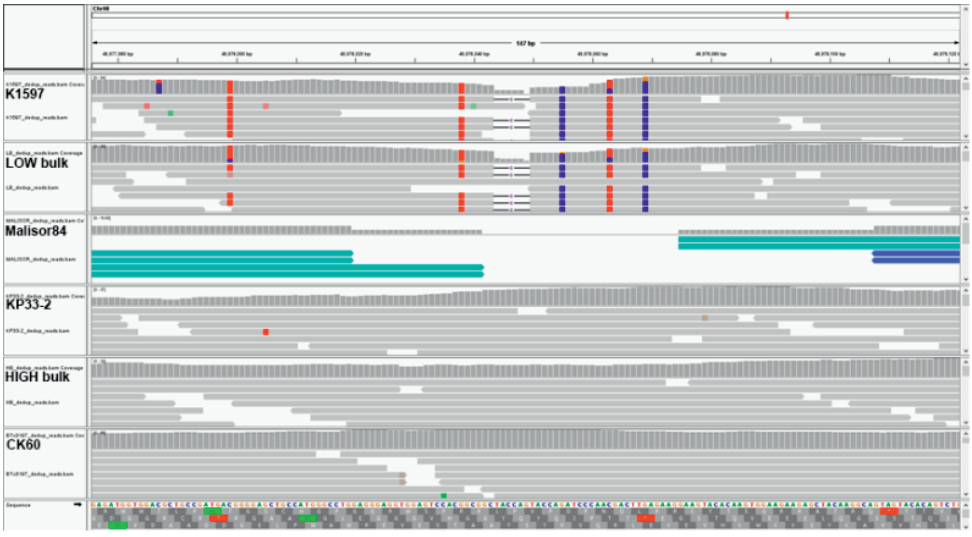


Figure 10. Sequence reads aligned to variant area in *Sobic.008G106200* Chr03: Chr08:49,977,976..49,978,122 from K1597, LOW bulk, Malisor84 (an unrelated low sorgomol line) KP33-2, HIGH bulk and CK60 (an unrelated high sorgomol line). Note the clean match between variations particular to K1597 in reads of the LOW bulk and absence of these variations in the HIGH bulk reads. This indicates RILs in the LOW bulk all carry the K1597 allele and the HIGH bulk RILs all carry the KP33-2 allele, and this gene is therefore a good candidate. Further support comes from sequenced sorghum lines unrelated to this population. CK60, a high sorgomol producer shares the KP33-2 allele. This area of the candidate gene is deleted in Malisor84 which doesn't produce sorgomol. Reads from whole genome sequences viewed with IGV software (Broad Institute, Robinson et al., 2011).

Discussion

Several modifications or substitutions on the A-B rings of strigolactones have been reported and shown to affect their biological activity (Akiyama et al. 2010). In the latter work, it was shown, for example, that introducing a hydroxyl group at different positions (C5 and C9) of 5-deoxystrigol reduced its hyphal branching activity. Hydroxylation of 5-deoxystrigol at the C-9 position is an enzymatic step postulated in the production of sorgomol. As a result, sorgomol becomes the most active strigolactone when compared to other strigolactones to induce *S. hermonthica* germination (Umehara et al. 2015; Cardoso et al. 2014). Interestingly, it is less active towards *Orobanche spp.* (Xie et al. 2008). On the other hand, when compared to other strigolactones, it showed lower activity in inducing hyphal branching activity towards AM fungi *G. margarita* (Akiyama et al. 2010). All in all, there are very few indications that the conversion of 5-deoxystrigol to sorgomol has any effect on the known rhizosphere biological processes affected by strigolactones. Identifying the gene that is responsible for the 5-deoxystrigol to sorgomol conversion would give the tools to study what the biological relevance of this process is. Genetic modification or selection of mutants using TILLING would allow to create identical genotypes that only differ in sorgomol production and could be used to study biological consequences of sorgomol presence or absence. An example of such a biological effect could be the root microbial community composition. Schlemper et al. (2017) showed that orobanchol and 5-deoxystrigol producing sorghum genotypes recruit a slightly different root microbiome. It would be interesting to see if sorgomol also affects the recruitment of bacteria or fungi to the roots of sorghum.

In this study, we come up with a number of candidate genes for the conversion of 5-deoxystrigol to sorgomol, based on bulk segregant analysis of RNAseq and genomic DNA of RILs derived from a cross of two sorghum lines contrasting for the levels of sorgomol in their root exudates, as well as QTL mapping. Reviewing the assumptions we made in our search for candidate genes, we found that *Sobic.008G106200* satisfied most of them. The frequency distribution of the measured sorgomol phenotypes in the recombinant inbred population is consistent with a 1:1 ratio expected for two parental alleles at a single locus and giving one of two phenotypes in the homozygous state. Calling the phenotype as HIGH or LOW sorgomol and associating this with genotypes obtained from polymorphic chromosome anchored SSR markers between K1597 and KP33-2, mapped the causal locus to Chromosome 8. Bulk segregant analysis of whole genome sequencing obtained from the HIGH and LOW bulks also mapped the sorgomol locus to Chromosome 8 but the specific region was difficult to determine when standard variant calls against either BTx623 or Rio references were used as genotypes. This issue was resolved by constructing an alternate reference that replaced the reference bases where both parents varied with the parental bases in areas where neither KP33-2 nor K1597 matched the standard reference assemblies and then aligning the reads of the LOW and HIGH sorgomol bulks to this assembly. This showed three QTLs, the most significant on Chromosome 8.

Each of the QTL peaks detected by bulked segregant analysis of genomic pools on Chromosomes 1, 3 and 8 spanned genomic regions containing cytochrome P450s. We based our assumption that the "sorgomol gene" would code for such an enzyme on the fact that it could catalyze the oxidation step required to add a hydroxyl group at C-9 of 5-deoxystigol to yield sorgomol. It was previously shown that roots of a high sorgomol exuding sorghum cultivar (Sudax) fed with deuterium-labeled 5-deoxystrigol could convert it to sorgomol (D-labeled) and this conversion was mostly eliminated when uniconazole-P, a specific inhibitor of cytochrome P450s, was added to the assay mixture (Motonami et al. 2013). Several P450s have already been shown to play an important role in strigolactone diversification (Zhang et al. 2014; Pandey et al. 2016). A P450 from *Arabidopsis thaliana*, AtMAX1, was reported to produce carlactonoic acid by catalyzing the oxidation of the methyl group of carlactone at C-19 (Abe et al. 2014). In recent work, a tomato MAX1 was also shown to produce carlactonoic acid from carlactone, upon expression in yeast as well as in *Nicotiana benthamiana* (Yoneyama et al. 2018; Zhang et al. 2018). A P450 encoded by rice *Os01g0700900* catalyzes the production of 4-deoxyorobanchol from carlactone, while another P450, encoded by *Os01g0701400*, catalyzes C-4 hydroxylation of 4-deoxyorobanchol to form orobanchol (Zhang et al. 2014). A maize MAX1 was also shown to convert carlactone to carlactonoic acid but could also convert 4-deoxyorobanchol to orobanchol (Yoneyama et al. 2018). In this work, MAX1s were classified into three types. The A1-type catalyzes the conversion of carlactone to carlactonoic acid whereas the conversion of carlactone to 4-deoxyorobanchol (through carlactonoic acid) is catalyzed by A2-type MAX1s. A3-type MAX1s convert carlactone to carlactonoic acid and 4-deoxyorobanchol to orobanchol. In Chapter 4, we performed a transient assay in *N. benthamiana* where we co-infiltrated the four putative sorghum MAX1 homologs together with *D27*, *CCD7* and *CCD8* from rice. In this work, we showed that carlactone is a substrate of the two MAX1 homologs; Sb2310 and Sb2210. In the transient assay, we were not able to check if any of these MAX1 were able to convert 5-deoxystrigol to sorgomol since we did not have the gene that converts carlactone to 5-deoxystrigol. Therefore, we looked at these genes in the RNAseq and bulk segregant DNA data. However, none of the sorghum MAX1s showed any significant difference both in Δ(SNP-index) and transcription level between HIGH and LOW sorgomol bulks suggesting that they can indeed not catalyze the 5-deoxystrigol to sorgomol conversion. Therefore, we excluded these genes from our list of candidate genes.

We were also able to eliminate the cytochrome P450 coding candidate genes from bulk segregant analysis of the genomic pools on Chromosomes 1 and 3 by viewing the reads aligning to these genes and comparing these to the parents (Figures 7-10). We assumed that if the variations in these genes determined sorgomol levels measured in the root exudates then those variations should be held uniformly by all members of each bulk. That is, all LOW bulk members should carry the K1597 allele and all HIGH bulk members should carry the KP33-2 allele. Because we mixed the DNA from all members of a particular pool, the sequence reads when aligned to the reference should uniformly share the variations with the reads from their respective parental types. This was only observed for the Chromosome 8 candidate, *Sobic.008G106200* (Figure 10). There are consistent variations between the KP33-2 and K1597

alleles that impact amino acid residues (Supplementary Figures 1 and 2) and presumably would alter the function of the cytochrome P450 for which they code.

The consensus sequence of *Sobic.008G106200* for the HIGH sorgomol bulk and KP33-2 match the BTx623 reference sequence throughout the coding region, and indeed also throughout the rest of the gene (UTRs, introns, promoter). Indeed, as stated above BTx623 also produces about 15% sorgomol. We also looked at other sorghum lines unrelated to our RIL population or its parents for which we have obtained root exudate strigolactone profiles and have re-sequenced, and examined their genomic sequences at *Sobic.008G106200*. CK60, a sorghum line with root exudate containing both 5-deoxystrigol (653 pmol/plant) and sorgomol (410 pmol/plant) has the KP33-2 allele. Another line, Malisor 84, whose root exudate contains (in the same run as CK60) only 5-deoxystrigol (2645 pmol/plant), has neither the KP33-2 nor the K1597 allele at *Sobic.008G106200* and actually seems to be missing this gene by deletion (Figure S1). While these examples do not provide definitive evidence for this candidate sorgomol gene, it does support what one would expect in a sorghum line able to make 5-deoxystrigol and also possessing a functional allele allowing it to also make sorgomol.

There is another assumption about the "sorgomol gene" that *Sobic.008G106200* violates. Expression of this candidate does not appear to contrast between K1597 and KP33-2 or between their respective bulked pools by RNAseq. Considering that the quantities of the substrates and products we observe in our experiments are in picomoles and our highest HIGH bulk member exudes only just over a thousand of these units of sorgomol over the 48hr collection period, we do not expect a great deal of the transcripts from its biosynthetic genes. We did find, however, among the other transcripts involved in strigolactone biosynthesis, RPKM values ranging from 2-3325 and so we looked for transcript levels in this range for our candidates. The candidate *Sobic.008G106200* fell below this range with the HIGH bulk and KP33-2 RPKMs averaging 0.36 and the LOW bulk and K1597 with average RPKMs 3× greater at 1.17. We based the assumption that we would see differences between the LOW and HIGH bulks on the expectation that the "mutant" allele, that of K1597 and the LOW sorgomol RILs, for our candidate would be non-functional. One might expect a gene with a non-functional product to be silenced. We did not consider an *alternate* function. We see from Table 2 that the strigolactones produced by each parental type have three possible oxidative fates for the 5-deoxystrigol they produce represented by the three hydroxylated strigolactones detected in the root exudates, *ent*-2'-*epi*-orobanchol (OH at C-4), which was negligible in both groups, strigol (OH at C-5) and sorgomol (OH at C-9). The low sorgomol lines generally exude more strigol than high sorgomol lines, on average 2.5× the amount. Root exudates of the low sorgomol lines also contain on average about twice the 5-deoxystrigol as low sorgomol lines. This may be due to perturbations in feedback loops between strigolactone oxidative pathways caused by increased levels of substrate, *i.e.*, more 5-deoxystrigol present due to its inability to convert to sorgomol drives the enzyme that converts it to strigol. Alternatively, given that both the path from 5-deoxystrigol to strigol and sorgomol are likely catalyzed by cytochrome P450s, could it be possible that the altered cytochrome P450 product of the K1597 allele, having lost the ability to oxidize C-9, now oxidizes, however inefficiently, at C-5? Strigol was not detected

in the deuterium labeled 5-deoxystrigol feeding studies that led to sorgomol in Sudax (Motonami et al., 2013), but then we would not expect that in a sorghum line that makes a cytochrome P450 able to efficiently oxidize C-9 leading to sorgomol. As we have measured in KP33-2 and the high sorgomol RILs that inherited its allele, their root exudates contain very little strigol.

Of course, these things are mere speculation, and only a way to explain why our favorite candidate gene is actually expressed *more* in lines that essentially don't make sorgomol. In the bulk segregant RNAseq data, we found another P450 (*Sobic.001G319900*) that was expressed 5-fold higher in high sorgomol producing lines compared to low sorgomol producing lines. However, the bulk segregant genome sequencing did not yield a significant SNP within 20kb up or down-stream of this gene. Saying this, it is also possible that *Sobic.001G319900* might be regulating the conversion of 5-deoxystrigol to sorgomol but is regulated by another gene causing the difference in the expression of *Sobic.001G319900* in high and low sorgomol producing lines, but we were not able to identify a candidate for this. Verification through cloning and testing *in vitro*, *in vivo*, and through mutagenesis should show whether *Sobic.001G319900* is indeed encoding the cytochrome P450 enzyme that catalyzes the conversion of 5-deoxystrigol to sorgomol.

Materials and Methods

Genetic material

Recombinant inbred lines (RILs) used in this study were advanced to the F8 by single seed descent from an F2 of the cross (KP33-2 × K1597) at the Sorghum Program in the Agronomy Department of Purdue University in West Lafayette, Indiana, USA.

Strigolactone extraction from root exudates and extracts

We grew 162 RILs, KP33-2 and K1597 in a climate chamber to collect root exudates and tissue for strigolactone analysis and nucleic acid extraction. Sorghum seeds were surface sterilized with 1% bleach for 30 min, the seeds were washed and germinated for 48hrs in darkness on moistened filter paper. The climate chamber was set at a temperature of 25^{o}C and 16 hrs light and 80% relative humidity. Three germinated seeds were transplanted to 14 cm pots filled with river sand. Seed from some RILs showed low germinability or poor growth and were excluded from the experiment. In the end, data was collected from 96 RILs and the two parents arrayed in a randomized blocked design with three biological replicates. A total of three experiments were performed where in each experiment, one biological replicate for all sorghum lines was grown. The seedlings were grown for two weeks and were watered with half-strength Hoagland's nutrient solution. After two weeks, the pots were flushed with 1 L of phosphate deficient half strength Hoagland's to remove the phosphate from the pots. The plants were then grown for an additional week and watered with half strength Hoagland's without phosphate to induce strigolactone production. After that week, a second flushing was used to rinse out any metabolites, including strigolactones, accumulated in the pot. Another 48 hrs later, 1 L of root exudate was collected by passing 1 L of tap water through each pot. Subsequently, the roots were collected, washed, frozen in liquid nitrogen and stored at -80^{o}C for further analysis.

The collected root exudates were passed through a C18 column (500 mg, Grace pure) that was preconditioned with 3 ml acetone and 3 ml of demineralized water. The strigolactones were eluted with 4 ml acetone. Two ml of the eluent were used for further analysis. 200 µl of internal standard (D6-*epi*-5-deoxystrigol) to a final concentration 0.1 nmol/ml was added, and the acetone was evaporated to dryness. The samples were then re-dissolved in 50 µl of ethyl acetate and 4 ml hexane and passed through a silica column (200 mg) (preconditioned with 2 ml ethyl acetate and 4 ml hexane). The columns were eluted with 2 ml of 10:90 hexane: ethyl acetate (v/v). The eluent was evaporated to dryness and re-dissolved in 200 µl of 25% acetonitrile in water. The samples were filtered through a 0.45 µm Minisart SRP filter (Sartorius, Germany). Strigolactones were then analyzed by comparing the retention times and mass transition with the internal standards using a Waters Xevo tandem quadrupole mass spectrometer. An Acquity UPLC BEH C18 column was used for chromatographic separation by applying a water: acetonitrile gradient with 0.1% formic acid to the column. The starting gradient was 5% acetonitrile for 0.33 min. Then, it was raised to 27% (v/v) acetonitrile in 0.34

min, followed by a 4.33 min gradient to 40% (v/v) acetonitrile, then rising to 65% acetonitrile in 3 min and maintained for 0.67 min. A 90% acetonitrile gradient replaced that in 0.2 min was used and was maintained for 0.46 min. The gradient was then set back to 5% acetonitrile in 0.2 min and maintained for 2.47 min to equilibrate the column prior to the next run. The column temperature was set at 50^{o}C with flow rate of 0.5 mL/min. 10 – 20 µl of samples were injected. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L/hr, respectively. The source temperature was set at 150^{o}C, the capillary voltage was set at 3.0 kV and the desolvation temperature at 650^{o}C. For each standard, the cone voltage was optimized using Waters IntelliStart MS Console. MRM was used for identification of strigolactones. Data were analyzed using MassLynx 4.1 (combined with TargetLynx) software.

To extract strigolactones from roots, 2 ml ethyl acetate containing 0.1 µM internal standard GR24 was added to 500 mg frozen root powder. The samples were sonicated for 10-15 min in a Branson 3510 ultrasonic bath, centrifuged for 10 min at 2000 rpm and the supernatant transferred to a clean vial. An additional 2 ml of ethyl acetate were used to re-extract the pellet which was then centrifuged again for 10 min. The combined supernatants were evaporated to dryness and the residue re-dissolved in 50 µl ethyl acetate and 4 ml hexane. The samples were then loaded on a silica column (200 mg) (preconditioned with 2 ml ethyl acetate and 4 ml hexane). The column was eluted in fractions: fraction 1 (eluted with 4 ml 35:80 ethyl acetate: hexane) and fraction 2 (4 ml 90:20 EtOAc: hexane). The fractions were combined, and the solvent was evaporated to dryness after which the residue was re-dissolved in 200 µl of 25% acetonitrile in water. Strigolactones were then analyzed as described above.

RNA isolation and preparation for RNAseq

Fourteen sorghum lines (2 parental lines and 12 RILs) were selected for RNAseq (Table 1). Six high and six low sorgomol producing RILs were selected and the total RNA from three RILs were pooled together. For each bulk, two technical replicates were used. The RNA from the parental lines of high and low sorgomol producing lines were sequenced separately.

Table 1. List of sorghum lines selected for Bulk sequencing and RNA-seq experiments

Sorghum lines	High sorgomol producing lines		Sorghum lines	Low sorgomol producing lines	
	Bulk sequencing	RNAseq		Bulk sequencing	RNAseq
KP33-2	√	√	K1597	√	√
#3821	√	√	#3932	√	
#3780	√	√	#3777	√	
#3942	√		#3959	√	
#3953	√		#3812	√	
#3817	√		#3804	√	
#3813	√	√	#3832	√	√
#3827	√		#3914	√	√
#3952	√		#3926	√	
#3786	√	√	#3862	√	√
#3869	√	√	#3807	√	√
#3857	√	√	#3930	√	
#3957	√		#3782	√	√

RNA was isolated from 100 mg frozen and ground root material using hot borate method (modified from (Wan, 1994 #229). Ground seeds were re-suspended in 800 µl XT-hot borate buffer (0.2 M Na borate decahydrate (Borax), 30 mM MEGTA, 1% sodium dodecyl sulfate (SDS), 1% Na deoxycholate, 1.5 mM polyvinylpyrrolidone (PVP) and 13 mM dithiothreitol (DTT)). The quality of the RNA was assessed by checking the Absorbance (Abs) at 260 and 280 nm and Abs260 nm/Abs280 nm ratio >1.8 was used as a good quality RNA. 1% agarose gel was used to check the RNA integrity. The RNA samples were stored at – 80^{o}C until sent for sequencing.

Bulk sequencing

Recombinant inbred lines were selected to construct two bulks based on the strigolactone profiles obtained from phosphate starved plants (described above). We avoided lines in which strigolactone content was generally low (major strigolactone < 90 pmol/plant). The 12 RILs chosen for the HIGH sorgomol bulk had sorgomol levels ranging from 556 pmol/plant to 1131 pmol/plant, nearly equal to or exceeding levels in the high sorgomol parent KP33-2. These lines had 5-deoxystrigol levels ranging from 5-85 pmol/plant. The 12 selected RILs for the LOW sorgomol bulk had sorgomol levels between 1 and 4 pmol/plant, in the range of the low sorgomol parent K1597. These LOW bulk members and K1597 had 5-deoxystrigol levels ranging from 93 to 141 pmol/plant.

An area of 0.9 cm^2 was cut from the leaf of one seedling from each of the 12 selected RILs in each bulk. These were combined by the genotype they represent with respect to their sorgomol levels (high or low) into a mortar where they were ground with a pestle together in liquid nitrogen. Two bulks were thus homogenized, the first (HIGH bulk) with plants having ≥ 550 pmol/plant sorgomol in their root exudates, the second (LOW bulk) having ≤ 4 pmol/plant. DNA was isolated in 4 × 100 mg aliquots from each homogenate in separate extractions using the Qiagen DNeasy Plant Mini Kit (Qiagen, USA). Equimolar concentrations of DNA from each

of the four aliquots representing each bulk were combined and purified with the DNA Clean-Up & Concentration Kit (ZYMO Research, USA) before submitting for sequencing at the Purdue Genomics Core Facility, West Lafayette, Indiana.

Construction of libraries and Illumina sequencing

Libraries were prepared using TruSeq DNA PCR-Free LT Library Preparation Kit - Set B, FC-121-3002 (Illumina). Two micrograms of DNA from each sample was sheared using a Covaris S2 (Covaris, Woburn, MA), end repaired, and adapter ligated. Each sample was ligated to differently indexed adapters to allow them all to be run in the sample lane. Size selection of libraries was performed using polyethylene glycol cuts with the aid of magnetic binding beads as described in the kit protocol. This results in a target insert size of 500–600 bp among library molecules that cluster. The quality of resulting DNA libraries was assessed on an Agilent Technologies 2100 Bioanalyzer using a High Sensitivity chip. Final libraries were titred for clustering using a KAPA Library Quantification Kit Illumina, (KAPA Cat # KR045), pooled and clustered in 3.5 lanes of Illumina Hiseq 2500 High Output v3 chemistry (Illumina Inc., San Diego, CA) to generate 100-base pair-end reads. Average coverage was 29× with 233 million reads.

QTL mapping

We used the 'OneMap' statistical package in R for constructing the genetic map (Margarido et al. 2007). Two point recombination fractions were estimated and the markers assigned to linkage groups using the criteria LOD = 3, maximum recombination fraction (max.rf) = 0.255, and map function set to Kosambi. We conducted quality checks that included segregation distortion, correct LG assignment, and genotyping error before building the genetic map. Marker genotypes that were extremely distorted or had high genotyping error were excluded. Markers were then ordered on each linkage group using the rapid chain delineation method (RCD) (Doerge et al. 1996), following the same criterion used for two-point recombination estimates, and rippled with window = 3 and LOD = 3. The generated genetic map was saved with the R function "write.map" for QTL mapping with different software packages.

The marker order and distance inferred by "OneMap" were used to find the QTL using single marker analysis (SMA) (Zeng 1994), implemented in Windows QTL Cartographer version 2.5 (Wang et al. 2007).

NGS read mapping and variant calls

Illumina HiSeq paired-end sequencing reads of K1597 and KP33-2 were mapped to version 3.1.1 of BTx623 sorghum genome reference sequences (Paterson et al. 2009) and RIO version 2.1 (*Sorghum bicolor* Rio, DOE-JGI, http://phytozome.jgi.doe.gov/) sorghum genome references downloaded from the Phytozome (Version 12) web portal (Goodstein et al. 2012). Read alignment was performed using BWA 0.7.15 program for short reads alignment (Li et al. 2009). Variant calls were computed for the parents against the reference genome (BTx623 and

RIO) using a GATK pipeline (Van der Auwera et al. 2013). Highest quality and confidence SNPs were extracted from the VCF files and INDELS were excluded (the indels would shift the sequence positions). The FastaAlternateReferenceMaker (GATK version 3.8-0) was used to replace the reference bases at the variation sites with the parental bases (alternative reference sequence) and an alternate fasta was defined. The Illumina HiSeq paired-end sequencing reads of the low and high sorgomol bulks was mapped to the alternate reference fasta for each parent. Variant calls were computed against the alternate reference fasta.

Bulk segregant analysis

The R statistical package QTLseqr (Mansfeld et al. 2018), an R package that performs next-generation sequencing bulk segregant analysis was used to perform the QTL-seq (Takagi et al. 2013) and G' (Magwene et al. 2011) analysis. QTL were defined as regions with a q-value above the false discovery rate of 0.01 or a Δ(SNP-index) above a confidence interval of 99% for G' or QTL-seq, respectively. In addition, the software Popoolation2 (Kofler et al. 2011) was used to compare allele frequencies for SNPs between the two sorgomol bulks to identify significant differences. The software package VariantAnnotation(Lawrence et al. 2013; Obenchain et al. 2014) was used to annotate variants, compute amino acid changes and predict coding outcomes on gene function within the significant QTL confidence intervals. We used the predicted protein sequences of BTx623 (version 3.1).

Data analysis

The raw data from RNA sequencing was filtered for sequencing adaptors and PhiX, these "contaminations" are of a technical nature. The sorghum V3.1 genome sequence and annotation were downloaded and combined in CLC.

References

Abe, Satoko, Aika Sado, Kai Tanaka, Takaya Kisugi, Kei Asami, Saeko Ota, Hyun Il Kim, et al. 2014. 'Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18084-89.

Agusti, J., S. Herold, M. Schwarz, P. Sanchez, K. Ljung, E. A. Dun, P. B. Brewer, et al. 2012. 'Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants (Proceedings of the National Academy of Sciences United States of America (2011) 108, (20242-20247) DOI: 10.1073/pnas.1111902108)', *Proceedings of the National Academy of Sciences of the United States of America*, 109.

Akiyama, K., K. Matsuzaki, and H. Hayashi. 2005. 'Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi', *Nature*, 435: 824-7.

Akiyama, K., S. Ogasawara, S. Ito, and H. Hayashi. 2010. 'Structural requirements of strigolactones for hyphal branching in AM fungi', *Plant Cell Physiol*, 51: 1104-17.

Al-Babili, Salim, and Harro J. Bouwmeester. 2015. 'Strigolactones, a Novel Carotenoid-Derived Plant Hormone', *Annual Review of Plant Biology*, 66: 161-86.

Alder, Adrian, Muhammad Jamil, Mattia Marzorati, Mark Bruno, Martina Vermathen, Peter Bigler, Sandro Ghisla, et al. 2012. *The Path from -Carotene to Carlactone, a Strigolactone-Like Plant Hormone*.

Arite, Tomotsugu, Hirotaka Iwata, Kenji Ohshima, Masahiko Maekawa, Masatoshi Nakajima, Mikiko Kojima, Hitoshi Sakakibara, et al. 2007. 'DWARF10, an RMS1/MAX4/DAD1 ortholog, controls lateral bud outgrowth in rice', *The Plant Journal*, 51: 1019-29.

Besserer, A., V. Puech-Pages, P. Kiefer, V. Gomez-Roldan, A. Jauneau, S. Roy, J. C. Portais, et al. 2006. 'Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria', *PLoS Biol*, 4: e226.

Booker, Jonathan, Michele Auldridge, Sarah Wills, Donald McCarty, Harry Klee, and Ottoline Leyser. 2004. 'MAX3/CCD7 Is a Carotenoid Cleavage Dioxygenase Required for the Synthesis of a Novel Plant Signaling Molecule', *Current Biology*, 14: 1232-38.

Bouwmeester, H. J., Chr Roux, J. A. Lopez Raez, and G. Bécard. 2007. "Rhizosphere communication of plants, parasitic plants and AM fungi." In *Trends in Plant Science 12 (2007) 5*. 2007.

Cardoso, C., Y. Zhang, M. Jamil, J. Hepworth, T. Charnikhova, S. O. Dimkpa, C. Meharg, et al. 2014. 'Natural variation of rice strigolactone biosynthesis is associated with the deletion of two MAX1 orthologs', *Proc Natl Acad Sci U S A*, 111: 2379-84.

Doerge, R. W., and G. A. Churchill. 1996. 'Permutation tests for multiple loci affecting a quantitative character', *Genetics*, 142: 285-94.

Flematti, G. R., A. Scaffidi, M. T. Waters, and S. M. Smith. 2016. 'Stereospecificity in strigolactone biosynthesis and perception', *Planta*, 243: 1361-73.

Gomez-Roldan, M. V., S. Fermas, P. B. Brewer, V. Puech-Pages, E. A. Dun, J. P. Pillot, F. Letisse, et al. 2008. "Strigolactone inhibition of shoot branching." In *Nature 455 (2008)*. 2008.

Goodstein, D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes, J. Fazo, T. Mitros, et al. 2012. 'Phytozome: a comparative platform for green plant genomics', *Nucleic Acids Res*, 40: D1178-86.

Harrison, M. J. 2005. 'Signaling in the arbuscular mycorrhizal symbiosis', *Annu Rev Microbiol*, 59: 19-42.

Kloeckener-Gruissem, B., J. M. Vogel, and M. Freeling. 1992. 'The TATA box promoter region of maize Adh1 affects its organ-specific expression', *Embo j*, 11: 157-66.

Kofler, R., R. V. Pandey, and C. Schlotterer. 2011. 'PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq)', *Bioinformatics*, 27: 3435-6.

Lawrence, Michael, Wolfgang Huber, Hervé Pagès, Patrick Aboyoun, Marc Carlson, Robert Gentleman, Martin T. Morgan, et al. 2013. 'Software for Computing and Annotating Genomic Ranges', *PLOS Computational Biology*, 9: e1003118.

Li, J., and M. P. Timko. 2009. 'Gene-for-gene resistance in *Striga*-cowpea associations', *Science*, 325: 1094.

Lin, Hao, Renxiao Wang, Qian Qian, Meixian Yan, Xiangbing Meng, Zhiming Fu, Cunyu Yan, et al. 2009. 'DWARF27, an Iron-Containing Protein Required for the Biosynthesis of Strigolactones, Regulates Rice Tiller Bud Outgrowth', *The Plant Cell*, 21: 1512.

Magwene, Paul M., John H. Willis, and John K. Kelly. 2011. 'The Statistics of Bulk Segregant Analysis Using Next Generation Sequencing', *PLOS Computational Biology*, 7: e1002255.

Makita, Y., S. Shimada, M. Kawashima, T. Kondou-Kuriyama, T. Toyoda, and M. Matsui. 2015. 'MOROKOSHI: transcriptome database in Sorghum bicolor', *Plant Cell Physiol*, 56: e6.

Mansfeld, B. N., and R. Grumet. 2018. 'QTLseqr: An R Package for Bulk Segregant Analysis with Next-Generation Sequencing', *Plant Genome*, 11.

Margarido, G. R., A. P. Souza, and A. A. Garcia. 2007. 'OneMap: software for genetic mapping in outcrossing species', *Hereditas*, 144: 78-9.

Mohamed, Abdalla H., T. L. Housley, and Gebisa Ejeta. 2010. 'Inheritance of hyper sensitive response to Striga parasitism in sorghum [Sorghum bicolor (L.) Moench]', *African Journal of Agricultural Research*, 5: 2720-29.

Mohemed, N., T. Charnikhova, E. F. Fradin, J. Rienstra, A. G. T. Babiker, and H. J. Bouwmeester. 2018. 'Genetic variation in Sorghum bicolor strigolactones and their role in resistance against *Striga hermonthica*', *J Exp Bot*, 69: 2415-30.

Motonami, Noriko, Kotomi Ueno, Hitomi Nakashima, Saki Nomura, Masaharu Mizutani, Hirosato Takikawa, and Yukihiro Sugimoto. 2013. 'The bioconversion of 5-deoxystrigol to sorgomol by the sorghum, Sorghum bicolor (L.) Moench', *Phytochemistry*, 93: 41-48.

Nomura, S., H. Nakashima, M. Mizutani, H. Takikawa, and Y. Sugimoto. 2013. 'Structural requirements of strigolactones for germination induction and inhibition of Striga gesnerioides seeds', *Plant Cell Rep*, 32: 829-38.

Nomura, Takahito, Xiaonan Xie, Koichi Yoneyama, Kaori Yoneyama, Takaya Kisugi, Kohki Akiyama, Yoshifumi Nakatani, et al. 2018. 'Which are the major players, canonical or non-canonical strigolactones?', *Journal of Experimental Botany*, 69: 2231-39.

Obenchain, V., M. Lawrence, V. Carey, S. Gogarten, P. Shannon, and M. Morgan. 2014. 'VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants', *Bioinformatics*, 30: 2076-8.

Pandey, A., M. Sharma, and G. K. Pandey. 2016. 'Emerging Roles of Strigolactones in Plant Responses to Stress and Development', *Front Plant Sci*, 7: 434.

Paterson, Andrew H., John E. Bowers, Rémy Bruggmann, Inna Dubchak, Jane Grimwood, Heidrun Gundlach, Georg Haberer, et al. 2009. 'The Sorghum bicolor genome and the diversification of grasses', *Nature*, 457: 551.

Smith, F. Andrew, and Sally E. Smith. 2011. 'What is the significance of the arbuscular mycorrhizal colonisation of many economically important crop plants?', *Plant and Soil : An International Journal on Plant-Soil Relationships*, 348: 63-79.

Sorefan, K., J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, E. Foo, S. Chatfield, et al. 2003. 'MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea', *Genes Dev*, 17: 1469-74.

Sun, Huwei, Jinyuan Tao, Pengyuan Gu, Guohua Xu, and Yali Zhang. 2016. 'The role of strigolactones in root development', *Plant Signaling & Behavior*, 11.

Takagi, H., A. Abe, K. Yoshida, S. Kosugi, S. Natsume, C. Mitsuoka, A. Uemura, et al. 2013. 'QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations', *Plant J*, 74: 174-83.

Toh, S., D. Holbrook-Smith, P. J. Stogios, O. Onopriyenko, S. Lumba, Y. Tsuchiya, A. Savchenko, et al. 2015. 'Structure-function analysis identifies highly sensitive strigolactone receptors in Striga', *Science*, 350: 203-7.

Umehara, Mikihisa, Mengmeng Cao, Kohki Akiyama, Tomoki Akatsu, Yoshiya Seto, Atsushi Hanada, Weiqiang Li, et al. 2015. 'Structural Requirements of Strigolactones for Shoot Branching Inhibition in Rice and Arabidopsis', *Plant and Cell Physiology*, 56: 1059-72.

Van der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, et al. 2013. 'From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline', *Curr Protoc Bioinformatics*, 43: 11.10.1-33.

Wang, R., and M. G. Brattain. 2007. 'The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa', *FEBS Lett*, 581: 3164-70.

Wang, Yanting, and Harro J. Bouwmeester. 2018. 'Structural diversity in the strigolactones', *Journal of Experimental Botany*, 69: 2219-30.

Xie, X., K. Yoneyama, T. Kisugi, K. Uchida, S. Ito, K. Akiyama, H. Hayashi, et al. 2013. 'Confirming stereochemical structures of strigolactones produced by rice and tobacco', *Mol Plant*, 6: 153-63.

Xie, Xiaonan, Kaori Yoneyama, Yuta Harada, Norio Fusegi, Yoichi Yamada, Satoshi Ito, Takao Yokota, et al. 2009. 'Fabacyl acetate, a germination stimulant for root parasitic plants from P*isum sativum*', *Phytochemistry*, 70: 211-15.

Xie, Xiaonan, Kaori Yoneyama, Dai Kusumoto, Yoichi Yamada, Yasutomo Takeuchi, Yukihiro Sugimoto, and Koichi Yoneyama. 2008. 'Sorgomol, germination stimulant for root parasitic plants, produced by Sorghum bicolor', *Tetrahedron Letters*, 49: 2066-68.

Yoneyama, K., N. Mori, T. Sato, A. Yoda, X. Xie, M. Okamoto, M. Iwanaga, et al. 2018. 'Conversion of carlactone to carlactonoic acid is a conserved function of MAX1 homologs in strigolactone biosynthesis', *New Phytol*, 218: 1522-33.

Yoneyama, K., X. Xie, K. Yoneyama, and Y. Takeuchi. 2009. 'Strigolactones: structures and biological activities', *Pest Manag Sci*, 65: 467-70.

Yoneyama, Koichi, Ayman A. Awad, Xiaonan Xie, Kaori Yoneyama, and Yasutomo Takeuchi. 2010. 'Strigolactones as Germination Stimulants for Root Parasitic Plants', *Plant and Cell Physiology*, 51: 1095-103.

Zeng, Z. B. 1994. 'Precision mapping of quantitative trait loci', *Genetics*, 136: 1457-68.

Zhang, Y., X. Cheng, Y. Wang, C. Diez-Simon, K. Flokova, A. Bimbo, H. J. Bouwmeester, et al. 2018. 'The tomato MAX1 homolog, SlMAX1, is involved in the biosynthesis of tomato strigolactones from carlactone', *New Phytol*, 219: 297-309.

Zhang, Y., A. D. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, et al. 2014. 'Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis', *Nat Chem Biol*, 10: 1028-33.

Zou, Junhuang, Shuying Zhang, Weiping Zhang, Gang Li, Zongxiang Chen, Wenxue Zhai, Xianfeng Zhao, et al. 2006. 'The rice HIGH-TILLERING DWARF1 encoding an ortholog of Arabidopsis MAX3 is required for negative regulation of the outgrowth of axillary buds', *The Plant Journal*, 48: 687-98.

Supplementary data

Supplementary tables

Table S1. RPKM values of putative strigolactone biosynthetic genes in sorghum obtained from RNAseq data

Gene code	Gene	(GE) – RPKM value							
		High sorgomol Bulk				Low sorgomol Bulk			
		Rep. 1	Rep. 2	KP33-2	Mean	Rep. 1	Rep. 2	K1597	Mean
Sobic.009G030800	D27	3.45	2.39	2.84	2.89	2.64	3.52	3.35	3.17
Sobic.005G105700	CCD8a	1.56	1.47	1.73	1.59	1.08	0.37	0.91	0.79
Sobic.002G168800	CCD8d	31.42	15.56	17.51	21.50	22.54	15.41	53.73	30.56
Sobic.007G170300	CCD8d-like	1.43	5.68	4.53	3.88	2.40	1.66	1.47	1.84
Sobic.006G170300	CCD7	22.14	14.40	15.15	17.23	18.53	17.76	19.89	18.73
Sobic.004G095500	MAX1 (Sb7880)	10.85	10.15	7.50	9.50	8.62	8.89	9.17	8.89
Sobic.010G170400	MAX1 (Sb2310)	2.19	1.50	1.44	1.71	1.29	3.17	2.07	2.18
Sobic.003G269500	MAX1 (Sb2210)	187.20	153.32	118.16	152.89	194.67	203.09	258.4	218.72
Sobic.003G269600	MAX1 (Sb2220)	9.29	7.69	7.24	8.08	9.37	9.28	10.47	9.71

Supplemental Table 2. Summary of SSR marker scores obtained from all genotyped and phenotyped lines. RILs included in the HIGH bulk highlighted in blue and those in LOW bulk RIL members highlighted in red.
Gray shaded areas are allele calls at SSR markers nearest the candidate genes: *Xtxp11* near *Sobic.001G319900* (Chr01:60,775,998-60,780,816 forward) and *SB4510* near *Sobic.008G106200* (Chr08:49,975,917-49,977,617 reverse)
Double vertical line marks estimated threshold between parental phenotypes with HIGH sorgomol RILs and KP33-2 above and LOW sorgomol RILs and K1597 below.

Genotype calls from SSR markers polymorphic between K1597 (LOW sorgomol parent; A alleles) and KP33-2 (HIGH sorgomol parent, B alleles). Marker locations from Phytozome, *Sorghum bicolor* v3.1.1

Line #	Chr01				Chr02						Chr03		Chr04			Chr05					Chr06		Chr07			Chr08																Chr09	Chr10					
	Xtxp71 (Chr01_6,685,894)	***Xtxp11*** (Chr01_60,140,171)	***Xtxp279*** (Chr01_68,337,429)	***Xtxp284*** (Chr01_72,274,249)	***Xtxp197*** (Chr02_1,462,149)	***Xtxp297*** (Chr02_4,219,592)	***Xtxp4*** (Chr02_13,148,350)	***Xtxp298*** (Chr02_57,081,385)	***Xtxp56*** (Chr02_61,533,433)	***Xtxp8*** (Chr02_75,950,383)	***Xtxp285*** (Chr03_67,787,890)	***Xtxp98*** (Chr03_73,937,261)	***Xcup05*** (Chr04_3,832,450)	***Xtxp343*** (Chr04_49,456,787)	***Xtxp29*** (Chr04_58,230,842)	***Xtxp303*** (Chr05_5,624,128)	***Kaf2*** (Chr05_67,367,100)	***SBKAFGK1*** (Chr05_67,367,113)	***PDStriga1b*** (Chr05_69,914,207)	***SSR1_UP26580*** (Chr05_70,061,854)	***Xtxp45*** (Chr06_48,362,620)	***Xtxp265*** (Ch06_50,259,846)	***Xtxp312*** (Chr07_4,559,206)	***Sb6_342*** (Chr07_37,531,281)	***Xcup57*** (Chr07_65,332,960)	***Xtxp28*** (Chr08_3,852,232)	***SB4510*** (Chr08_54,320,841)	***SB4511*** (Chr08_54,331,038)	***SB4513*** (Chr08_54,587,082)	***Ch08_TG(26)*** (Chr08_54,622,317)	***Ch08_TA(15)*** (Chr08_54,661,800)	***Ch08_AT(59)*** (Chr08_54,682,733)	***SB4516*** (Chr08_54,754,925)	***Xtxp354*** (Chr08_55,558,543)	***SB4524*** (Chr08_55,636,531)	***SB4528*** (Chr08_56,314,466)	***SB4529*** (Chr08_56,379,542)	***SB4531*** (Chr08_56,617,759)	***SB4534*** (Chr08_56,753,010)	***SB4536*** (Chr08_56,797,675)	***SB4540*** (Chr08_57,136,056)	***Xtxp107*** (Chr09_57,742,430)	***Xcup49*** (Chr10_233,352)	***Xtxp331*** (Chr10_11,742,556)	**Phenotype Call** (B=HIGH sorgomol; A=LOW sorgomol)	**Sorgomol** (pMol/plant/48hrs)	**5-deoxystrigol** (pMol/plant/48hrs)	**Sorgomol : 5-deoxystrigol**
KP33-2	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	593	36	17
#3821	B	B	A	B	B	B	B	B	B	B	B	B	B	B	A	A	B	B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	1131	51	22
#3780	B	B	B	A	A	A	B	B	B	A	A	A	B	A	B	B	B	B	B	B	B	B	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	1096	85	13
#3942	A	A	A	B	A	A/B	A	A	B	A	B	B	A	B	A	B	_	B	B	B	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	987	47	21
#3953	B	A	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	818	40	20
#3817	B	A	B	A	B	A	B	A	B	A	A	B	B	B	B	A	A	A	A	A	B	B	A	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	792	54	15
#3813	A	B	B	B	B	B	B	B	B	A	B	B	A	B	A	B	B	B	B	B	B	B	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	779	59	13
#3827	B	B	A	B	B	B	B	B	B	_	A	A	A	B	B	B	A	A	A	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	742	51	15
#3952	A	B	B	B	A	A	A	A	B	B	A	A	A	A	B	A	B	B	B	B	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	678	85	8
#3786	B	B	B	B	B	B	A	A	B	B	A	B	B	B	B	B	A	A	A	A	A	A	A/B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	662	27	25
#3869	B	B	B	B	A	A	A	B	A	A	B	B	A	B	A	B	B	B	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	B	A	B	640	5	139
#3857	A	B	_	A	B	B	B	B	B	A	B	A	A	B	B	A	B	B	A	B	A	A	A	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	580	24	25
#3957	A	A	A	A	B	B	B	B	B	A	A	B	A	B	B	A	B	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	556	79	7
#3966	B	A	A	A	B	B	A	A	A	A	B	B	A/B	A	B	B	A/B	B	A	A	B	B	A	A	B	A	A	A	A	A	A	A	A	A	A	A/B	A/B	A/B	B	A/B	A/B	A	A/B	A	B	1148	144	8
#3967	B	A	A/B	A/B	B	B	B	B	B	A	B	B	A	B	B	B	A	A	B	B	A	A	A	A	A	A	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	A	A	B	934	26	36
#3771	B	A	B	A	A	A	A	A	A	B	B	B	A	B	B	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	B	B	A	B	850	142	6
#3943	_	_	_	_	A/B	B	_	B	_	_	B	B	B	A	_	A	_	B	B	B	A	_	_	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	_	B	629	72	9
#3840	B	A/B	B	B	A	B	B	A	A	B	B	A	A	B	A	B	A/B	A/B	A/B	B	A	A	A	A	A/B	B	B	B	B	B	B	B	B	B	B	B	_	B	B	B	B	A/B	A	_	B	541	65	8
#3795	A	A	B	B	A	A	A	A	A	B	B	A	B	B	A	B	B	B	B	B	A	A	A	A	A	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	A	B	A	B	530	43	12
#3835	B	A	B	B	A	A	A	B	A	B	B	B	A	B	A	B	B	B	A	B	A	A	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	519	44	12
#3842	A	A	_	A	A	B	A	A	B	A	A	A	A	B	A	A	B	B	B	B	A	A	B	A	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	500	17	30
#3829	B	A	B	B	A	B	B	B	A	B	A	A	B	B	B	B	A	A	A	A	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	465	23	20
#3960	B	B	A	A	B	A	A	B	B	A	B	A	B	B	A	A	A	A	A	A	A	A	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	463	103	5
#3818	A	B	A	A	A	A	A	A	B	B	A	B	A	A	A	B	B	B	B	B	A	A	B	B	A	B	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	A	B	A	B	437	28	15
#3819	B	A	B	B	A	A	A	A	B	A	B	A/B	B	B	A	A	A	A	A	B	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	419	22	19
#3961	B	B	A	A	B	A	A	B	B	B	A	A	A	B	B	B	A	B	A	A	A	A	B	B	A	A	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B	A	B	392	26	15
#3826	B	A	A/B	A/B	B	B	A	A	B	B	A	A	B	B	A	B	B	B	A	A	A	A	A	A	B	A	A/B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	B	391	86	5
#3935	A	A/B	A/B	B	B	A/B	B	B	A	A	A	B	A	B	A	A	_	B	B	B	A	A	A/B	B	B	B	A/B	A	A/B	B	B	A	B	B	B	B	B	B	B	B	B	A	A/B	A	B	387	48	8
#3800	B	B	B	B	B	B	A	A	A	A	A	B	A	B	B	A	A	A	_	B	A	A	A	B	A	B	B	B	B	B	B	B	B	A	B	A	A	A	A	A	A	B	B	A	B	365	26	14
#3896	A	B	B	B	B	B	A	A	A	A	A	A	B	B	B	A	A	A	_	B	A	A	A	B	A	A	B	B	B	B	B	B	B	B	B	B	A/B	B	B	B	B	A/B	A	A	B	363	8	46
#3860	A	A		B	A	A	A	A	A	A	A/B	A	B	B	B	B	B	B	A	B	A	A	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	354	22	16
#3903	A	B	B	B	B	B	B	B	A	B	B	B	B	A	A	A	B	B	A	A	A	A	A	B	B	A	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	B	A	B	351	14	24
#3928	A/B	B	A	A	A	A	A	B	B	A	A	A	A	A	B	A	B	B	B	B	A	A	A	B	A	A	B	B	B	B	B	B	B	B	B	A	A	A	A	A	B	B	B	A	B	329	27	12
#3891	B	B	B	B	A	A	A	B	B	A	B	A	A	A	A	A	B	B	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	314	15	21
#3773	A	B	B	B	B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	A	A	A	A	A	A	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	-	B	A	B	307	33	9
#3919	B	A	B	A	B	A	A	A	A	A	B	A	B	B	B	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	282	11	25
#3775	A	A	B	A	B	B	A	A	A	A	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	257	24	11
#3913	B	A	B	B	B	B	B	A	B	B	A	A	B	A	A	B	B	B	A	A	B	B	A	A	A	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	B	235	19	12
#3796	B	B	A	B	A	A	B	A	B	B	B	B	B	B	A	B	A	A	A	A	B	B	B	B	A	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	A	A	A	B	225	42	5
#3831	B	A	A	A	A/B	A/B	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	B	212	30	7
#3866	A	B	A	A	A	B	A	A	A	B	B	A	B	B	A	B	B	B	B	B	A	A	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	A	B	190	14	14
#3902	B	B	A/B	A	A	A	B	A	A	A	A	B	B	B	B	B	B	B	A	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	178	10	17
#3941	B	B	A	B	B	B	B	B	B	B	A	B	A	B	A	B	_	B	B	B	B	B	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	B	165	97	2
#3939	B	A	A	B	A	A	B	B	B	A	B	B	B	B	A	B	_	B	A	A	B	B	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	161	18	9
#3909	A	A	A	A	A	A	B	A	A	A	C	B	B	C	A	B	B	B	A	B	A	A	B	B	A	A	B	B	B	B	B	B	B	A	A	A	B	B	A	_	A	B	B	B	B	152	22	7
#3824	A/C	B	A/B	A/B	A/B	C	B	A	A/B	_	A/B	A/B	A	A/B	B	C	A	A	B	B	A/B	A	A	A	A/B	A	B	B	B	B	B	B	B	A/B	B	B	B	A/B	A	_	A/B	B	A	A/B	B	120	51	2
#3785	A	A	B	B	B	B	A	B	B	B	B	B	B	B	B	B	A	A	A	A	B	B	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	_	B	B	A	A	B	117	84	1
#3776	B	B	A/B	A	B	B	B	A	A	A	B	A	A	B	A	A	A	A	A	B	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	99	8	12
#3947	A	B	A	A	B	B	A	A	A	B	B	B	A	B	A	A	B	B	B	B	A	B	B	B	B	B	A	A	A	A	A	A	A	B	B	B	B	A	A	A	A	B	A	A	B	91	62	1
#3904	B	B	A	B	A	A	B	B	B	A	B	B	_	B	B	B	B	B	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A/B	A/B	A	B	85	13	6
#3875	B	A	A	A	B	A	B	B	B	A	A	A	A	A	B	B	B	B	A	A	B	B	A	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	83	6	14
#3774	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	X	65	39	2
#3834	A	B	B	B	B	B	A	A	A	B	A	B	A	A	A	B	A	A	B	B	B	B	A	A	A	A	A	C	B	A	C	B	C	A	A	B	B	A	A	A	A	B	B	A	B	65	6	11
K1597	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	4	124	0.03
#3932	B	A	A	A	A	B	B	B	B	B	B	B	A	A	B	B	_	A	A	A	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	B	B/A	B	B	B	B	A	B	A	A	4	94	0.05
#3777	A	A	A/B	A/B	A	A	A	A	B	B	B	B	B	B	A	A	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	A	A	A	4	130	0.03
#3959	B	A	A	A	A	A	A	A	A	A	A	A	B	B	A	B	B	B	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	4	141	0.03
#3812	A	B	B	B	A	A	A	A	B	A	A	B	A	A	A	B	A	A	A	A	A	A	B	B	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	A	4	112	0.04
#3804	A	A	B	B	A	A	A	A	A	A	A	A	A	A	B	B	A	A	B	B	B	B	B	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	4	114	0.03
#3832	A	A	B	B	A	A	B	B	B	A	B	B	B	B	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	4	93	0.04
#3914	A	B	A	B	B	B	B	B	A	A	A	A	B	B	B	B	B	B	A	A	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	4	100	0.04
#3926	B	B	A	A	B	B	A	B	A	B	A	A	A	B	B	A	A	A	A	A	B	B	A	A	A/B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	A	A	4	102	0.03
#3862	A	B	_	B	B	A	A	B	A	B	A	A	B	B	A	B	B	B	B	B	A	A	B	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	3	117	0.03
#3807	B	A	B	B	A	A	B	A	A	A	A	B	A	B	B	B	A	A	A	A	A	A	B	A	A	B	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	A	2	97	0.02
#3930	B	B	B	B	B	B	A	A	A	A	B	B	B	B	B	A	B	B	B	B	A	A	A	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	A	A	A	2	169	0.01
#3782	B	A	B	B	B	B	B	B	B	A	A	A	B	B	A	A	A	A	_	A	A	A	B	B	B	_	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	B	B	A	A	1	132	0.01
#3924	A	B	B	B	B	B	A	B	B	B	A	B	A	B	B	A	A	A	A	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	37	101	0.37
#3880	A	B	A	A	B	B	A	A	B	B	A	A	A	B	B	B	A	A	B	B	B	B	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	20	79	0.25
#3936	A	A	B	A	A	A	B	A	A	B	B	B	A	B	B	B	_	B	B	B	A	A	B	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	13	101	0.13
#3856	A	B	_	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	10	90	0.12
#3828	A	A	A	A	A	B	B	B	B	B	B	A	B	B	A	B	B	B	B	B	B	B	A	A	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	8	88	0.09
#3964	B	A	B	B	B	B	A	A	B	A	B	B	B	B	B	A	B	B	A	B	B	B	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	7	96	0.08
#3772	B	A	B	B	A	B	B	B	A	A	A	B	B	A	A	B	A	A	B	B	A	A	B	A	B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	A	7	161	0.05
#3858	A/B	B	_	A	B	B	B	B	A	B	B	B	B	A	A	A	A	A	B	B	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	6	191	0.03
#3803	B	B	B	B	A	A	A	A	B	B	A	A	B	A	A	B	B	B	B	B	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	X	5	21	0.25
#3815	B	A	B	B	B	A	B	A	A	A	B	B	A	B	A	B	A	A	_	B	B	B	B	B	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	A	A	5	140	0.04
#3920	A	B	B	B	B	B	B	B	A	B	B	B	B	B	A	A	A	A	_	B	A	A	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	5	87	0.06
#3852	B	A		B	A	A	B	B	B	A	B	B	B	B	B	A	A	A	A	A	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	5	93	0.06
#3893	B	B	B	B	B	A	A	A	B	B	A	B	A	B	B	B	A	A	B	B	A	A	B	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	5	98	0.05
#3806	A	A	B	B	B	B	A	A	A	A	A	A	A	B	A	A	B	B	A	A	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	5	149	0.03
#3790	B	B	A	A	B	B	A	A	A	B	B	B	A	B	A	B	B	B	B	B	B	B	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	5	142	0.03
#3839	B	B	B	B	A	B	B	A	_	B	B	A	B	B	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	_	X	4	54	0.07
#3931	B	B	A	A	B	B	B	B	B	A	B	A	A	B	A	B	A	A	A	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	A	X	4	61	0.06
#3933	_	_	_	_	A	A	_	A	_	_	A	A	B	B	_	B	_	A	A	A	A	_	_	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	_	X	3	42	0.08
#3962	B	A	A	A	B	B	A	A	A	A	B	B	B	A	B	B	A	A	_	A	B	B	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	X	3	65	0.05
#3787	B	B	A	A	B	B	B	B	A	A	A	A	A	B	A	B	B	B	A	A	B	A	A	A	B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	A	A	A	A	3	80	0.04
#3955	A/B	A/B	A/B	A/B	A/B	A/B	B	B	B	A	A/B	B	A	A/B	B	A	A/B	A	A/B	B	A/B	A	A/B	A	A	B	A/B	A	A/B	B	B	A	B	A/B	A	A/B	A/B	A/B	B	B	A/B	B	B	A	X	3	34	0.08
#3925	A	A	A	A	A	A	A	A	A	A	B	B	B	A	A	A	B	B	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	X	2	6	0.36
#3845	A	B	_	B	A	A	A	A	A	A	B	A	B	A	A	A	B	B	B	B	B	B	B	A	B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	A	A	X	2	72	0.03
#3792	B	A	B	B	B	B	A	A	A	B	A	A	B	B	A	B	B	B	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	X	2	62	0.03
#3876	B	B	A/B	B	A	A	A	A	B	B	B	A	A	B	A	B	A	A	B	B	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	X	1	15	0.10
#3927	B	A	A	A	B	B	B	A	A	A	A	A	B	B	B	A	B	B	A	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	X	1	41	0.03
#3870	A	B	A/B	A	A	A	B	B	B	A	A	B	B	A	A	B	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	A	A	A	X	1	12	0.06
#3814	A	B	A	A	A	A	B	B	B	A	A	A	B	B	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	X	1	8	0.07

Supplementary figure

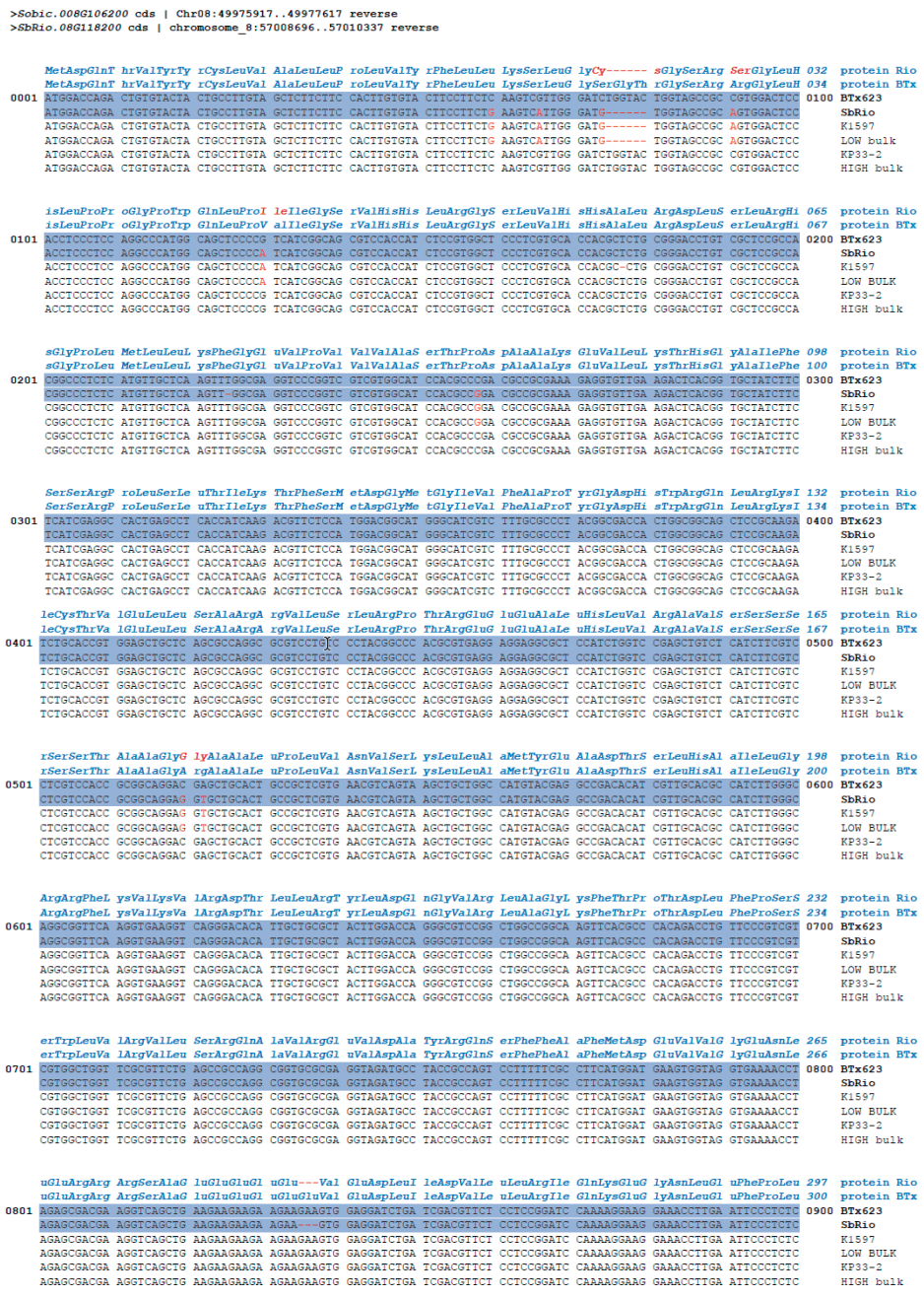


Figure S1. Consensus sequence of *Sobic.008G106200* when compared to HIGH and LOW sorgomol producing lines, parental lines (KP33-2, high sorgomol producing and K1597, low sorgomol producing) of the RIL population and two reference sequences, BTx623 and SbRio.

```
>Sobic.008G106200.1 (516AAs)
MDQTVYYCLVALLPLVYFLLKSLGSGIGSRRGLHLPPGPWQLPTIGSVHHLRGSLVHHALRDLSLRHGPLMLLKFGEVPVVVASTPDAAKEVLKTHGAIFSSRPLSLTI
KTFSMDGMGIVFAPYGDHWRQLRKICTVELLSARRVLSLRPTREEEALHLVRAVSSSSSSTAAGRAALPLVNVSKLLAMYEADTSLHAILGRRFKVKVRDTLLRYLDQG
VRLAGKFTPTDLFPSSWLVRVLSRQAVREVDAYRQSFFAFMDEVVGENLERRRSAEEEEEEVEDLIDVLLRIQREGNLEFPLSMRIIQAVIFDLIGGGIETATTTLQWA
MAELMRNPGIMAKAQAEVRGVFMGQTKVTEDRLGELSYLQLVIKETLRLHVPGPLLVPRECQEQCRILGYDVPKGAMVIVNAWAIARSPDYWVEPDTFHPDRFLSDTRD
FKGNDFEFIPFGAGRRICPGMAFGLANVELGLASLLFHFDWSLPEGMVPSELDMTEATEITARRKADLLLSATPCVKLPS*

>SbRio.08G118200.1 (513AAs)
MDQTVYYCLVALLPLVYFLLKSLGC__GSRRGLHLPPGPWQLPTIGSVHHLRGSLVHHALRDLSLRHGPLMLLKFGEVPVVVASTPDAAKEVLKTHGAIFSSRPLSLTI
KTFSMDGMGIVFAPYGDHWRQLRKICTVELLSARRVLSLRPTREEEALHLVRAVSSSSSSTAAGGAALPLVNVSKLLAMYEADTSLHAILGRRFKVKVRDTLLRYLDQG
VRLAGKFTPTDLFPSSWLVRVLSRQAVREVDAYRQSFFAFMDEVVGENLERRRSAEEEEEEVEDLIDVLLRIQREGNLEFPLSMRIIQAVIFDLIGGGIETATTTLQWA
MAELMRNPGIMAKAQAEVRGVFMGQTKVTEDRLGELSYLQLVIKETLRLHVPGPLLVPRECQEQCRILGYDVPKGAMVIVNAWAIARSPDYWVEPDTFHPDRFLSDTRD
FKGNDFEFIPFGAGRRICPGMAFGLANVELGLASLLFHFDWSLPEGMVPSELDMTEATEITARRKADLLLSATPCVKLPS*

>K1597 (514AAs)
MDQTVYYCLVALLPLVYFLLKSLGC___GSRRGLHLPPGPWQLPTIGSVHHLRGSLVHHALRDLSLRHGPLMLLKFGEVPVVVASTPDAAKEVLKTHGAIFSSRPLSLTI
KTFSMDGMGIVFAPYGDHWRQLRKICTVELLSARRVLSLRPTREEEALHLVRAVSSSSSSTAAGGAALPLVNVSKLLAMYEADTSLHAILGRRFKVKVRDTLLRYLDQG
VRLAGKFTPTDLFPSSWLVRVLSRQAVREVDAYRQSFFAFMDEVVGENLERRRSAEEEEEEVEDLIDVLLRIQREGNLEFPLSMRIIQAVIFDLIGGGIETATTTLQWA
MAELMRNPGIMAKAQAEVRGVFMGQTKVTEDRLGELSYLQLVIKETLRLHVPGPLLVPRECQEQCRILGYDVPKGAMVIVNAWAIARSPDYWVEPDTFHPDRFLSDTRD
FKGNDFEFIPFGAGRRICPGMAFGLANVELGLASLLFHFDWSLPEGMVPSELDMTEATEITARRKADLLLSATPCVKLPS*

>LOW bulk (514AAs)
MDQTVYYCLVALLPLVYFLLKSLGC___GSRSGLHLPPGPWQLPTIGSVHHLRGSLVHHALRDLSLRHGPLMLLKFGEVPVVVASTPDAAKEVLKTHGAIFSSRPLSLTI
KTFSMDGMGIVFAPYGDHWRQLRKICTVELLSARRVLSLRPTREEEALHLVRAVSSSSSSTAAGGAALPLVNVSKLLAMYEADTSLHAILGRRFKVKVRDTLLRYLDQG
VRLAGKFTPTDLFPSSWLVRVLSRQAVREVDAYRQSFFAFMDEVVGENLERRRSAEEEEEEVEDLIDVLLRIQREGNLEFPLSMRIIQAVIFDLIGGGIETATTTLQWA
MAELMRNPGIMAKAQAEVRGVFMGQTKVTEDRLGELSYLQLVIKETLRLHVPGPLLVPRECQEQCRILGYDVPKGAMVIVNAWAIARSPDYWVEPDTFHPDRFLSDTRD
FKGNDFEFIPFGAGRRICPGMAFGLANVELGLASLLFHFDWSLPEGMVPSELDMTEATEITARRKADLLLSATPCVKLPS*

>KP33-2 (516AAs)
MDQTVYYCLVALLPLVYFLLKSLGSGIGSRRGLHLPPGPWQLPTIGSVHHLRGSLVHHALRDLSLRHGPLMLLKFGEVPVVVASTPDAAKEVLKTHGAIFSSRPLSLTI
KTFSMDGMGIVFAPYGDHWRQLRKICTVELLSARRVLSLRPTREEEALHLVRAVSSSSSSTAAGRAALPLVNVSKLLAMYEADTSLHAILGRRFKVKVRDTLLRYLDQG
VRLAGKFTPTDLFPSSWLVRVLSRQAVREVDAYRQSFFAFMDEVVGENLERRRSAEEEEEEVEDLIDVLLRIQREGNLEFPLSMRIIQAVIFDLIGGGIETATTTLQWA
MAELMRNPGIMAKAQAEVRGVFMGQTKVTEDRLGELSYLQLVIKETLRLHVPGPLLVPRECQEQCRILGYDVPKGAMVIVNAWAIARSPDYWVEPDTFHPDRFLSDTRD
FKGNDFEFIPFGAGRRICPGMAFGLANVELGLASLLFHFDWSLPEGMVPSELDMTEATEITARRKADLLLSATPCVKLPS*

>HIGH bulk (516AAs)
MDQTVYYCLVALLPLVYFLLKSLGSGIGSRRGLHLPPGPWQLPTIGSVHHLRGSLVHHALRDLSLRHGPLMLLKFGEVPVVVASTPDAAKEVLKTHGAIFSSRPLSLTI
KTFSMDGMGIVFAPYGDHWRQLRKICTVELLSARRVLSLRPTREEEALHLVRAVSSSSSSTAAGRAALPLVNVSKLLAMYEADTSLHAILGRRFKVKVRDTLLRYLDQG
VRLAGKFTPTDLFPSSWLVRVLSRQAVREVDAYRQSFFAFMDEVVGENLERRRSAEEEEEEVEDLIDVLLRIQREGNLEFPLSMRIIQAVIFDLIGGGIETATTTLQWA
MAELMRNPGIMAKAQAEVRGVFMGQTKVTEDRLGELSYLQLVIKETLRLHVPGPLLVPRECQEQCRILGYDVPKGAMVIVNAWAIARSPDYWVEPDTFHPDRFLSDTRD
FKGNDFEFIPFGAGRRICPGMAFGLANVELGLASLLFHFDWSLPEGMVPSELDMTEATEITARRKADLLLSATPCVKLPS*
```

Figure S2. Amino acid sequence of *Sobic.008G106200* in HIGH and LOW sorgomol bulk sorghum lines, parental lines (KP33-2, high sorgomol producing and K1597, low sorgomol producing) of the RIL population and two reference sequences, BTx623 and SbRio.

Chapter 5

Characterization of sorghum MAX1 homologs in strigolactone biosynthesis

Mahdere Z. Shimels[1], Carolien Ruyter-Spira[1,] Harro Bouwmeester[1,2]

[1]Laboratory of Plant Physiology, Wageningen UR, PO Box 658, 6700 AR Wageningen, the Netherlands
[2]Present address: Plant Hormone Biology group, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands.

Abstract

Strigolactones are a plant hormone that regulates plant architecture. When exuded to the rhizosphere, they serve as signaling molecules that stimulate the symbiotic interaction with arbuscular mycorrhizal fungi and induce germination of root parasitic weeds. In several plant species, different enzymatic steps in the biosynthesis of strigolactones have been shown to be catalyzed by MORE AXILLARY GROWTH 1 (MAX1) homologs, members of the cytochrome P450 monooxygenase class CYP711A. *Arabidopsis* MAX1 catalyzes the conversion of carlactone to carlactonoic acid and a rice MAX1 homolog converts carlactone to 4-deoxyorobanchol which is further converted to orobanchol by another MAX1 homolog of rice. However, many genes involved in diversification of the strigolactones have not been elucidated yet. For instance, biosynthesis of 5-deoxystrigol, the predominant strigolactone produced by sorghum, is still a mystery. In this study, we characterized the four *MAX1* homologs of sorghum; *Sb04g007880* (Sb7880), *Sb10g022310* (Sb2310), *Sb03g032210* (Sb2210) and *Sb03g032220* (Sb2220) (Challis et al. 2013). We show that the expression level of these sorghum *MAX1* homologs is differentially regulated in response to phosphate starvation. Analysis of the expression pattern of these genes in different plant parts such as root, lower stem, auxiliary bud and flower head revealed that Sb2210 and Sb2220 exhibited a similar expression pattern, with the highest expression in lower stems and roots. Phylogenetic analysis showed their evolutionary relationship with other functionally characterized MAX1 homologs. *Sb2210* did not cluster with any of the functionally characterized rice or maize MAX1 homologs. From these analyses, we predicted that *Sb2310* is the likely ortholog of *Os06g0565100* and *ZmMAX1a* while *Sb7880* is of *Os02g0221900* and *ZmMAX1c*. Furthermore, using a transient assay in *Nicotiana benthamiana*, we showed that *Sb2310* and *Sb2210* consumed carlactone, however, this conversion did not result in the production of any of the known canonical strigolactones.

Introduction

In the past two decades, several studies have revealed the dual role of strigolactones as plant hormones and as signaling molecules in the rhizosphere (Akiyama et al. 2005; Besserer et al. 2006; Gomez-Roldan et al. 2008; Umehara et al. 2015; Cook et al. 1966). As plant hormones, strigolactones regulate shoot and root architecture (Gomez-Roldan et al. 2008; Umehara et al. 2008; Brewer et al. 2013; Koltai 2011). When exuded to the rhizosphere, they serve as signaling molecules for beneficial organisms such as arbuscular mycorrhizal (AM) fungi (Besserer et al. 2006; Akiyama et al. 2005). The production and exudation of strigolactones is boosted during nutrient-deficient conditions-especially of phosphate. Their secretion to the rhizosphere stimulates the symbiotic relationship of the plant with AM fungi. In this relationship, the AM fungi enhance phosphorus uptake of their host plants while relying on the carbon that these plants provide to them to stimulate their growth and development in a classic symbiotic relationship (Bouwmeester et al. 2007; Lopez-Raez et al. 2008).

The secretion of strigolactones into the rhizosphere is not always beneficial to the plant. They can be hijacked by parasitic plants and be used as host detection signaling molecules (Cook et al. 1966). Only upon the detection of a strigolactone, the seeds of these parasitic plants germinate and attach to a host within days after germination. Upon contact with the host root, the parasite will initiate a haustorium to be able to penetrate the host root and acquire water, assimilates and nutrients (Cui et al. 2016; Hibberd et al. 2001). These devastating weeds negatively affect the yield of cereals such as rice, maize, millet and sorghum (Berner et al. 1998).

So far, more than 20 strigolactones and strigolactone-like compounds have been identified in different plant species. Plants can produce and exude a blend of strigolactones that differs both in amount and type (Gobena et al. 2017; Yoneyama et al. 2008; Awad et al. 2006). Strigolactones have a common tricyclic structure (ABC-rings) connected to a butenolide D-ring in 2'R configuration via an enol ether bridge. At the B-C ring junction, they can be either β- or α-oriented, leading to two distinct groups, the strigol- and orobanchol-type strigolactones, respectively (Figure 1). Upstream of this in strigolactone biosynthesis, carlactone is the common precursor of all strigolactones, both canonical and non-canonical. The difference between the latter two classes is the ABC-ring which is present in canonical strigolactones but only in a (strongly) modified from in the non-canonical strigolactones (Xie 2017; Wang et al. 2018). Further diversification of strigolactones can be achieved by modifications such as oxidation of carlactonoic acid at C18 and C19 which results in the formation of the canonical strigolactone, 4-deoxyorobanchol (Zhang et al. 2014). Additional modifications such as hydroxylation, acetylation, demethylation, esterification, epoxidation, oxidation and decarboxylation on the AB-rings of the strigolactones results in further diversification (Motonami et al. 2013; Zhang et al. 2014; Al-Babili et al. 2015). For instance, hydroxylation of 4-deoxyorobanchol at C4 results in the formation of orobanchol (Zhang et al. 2014). Hydroxylation at C9 of 5-deoxystrigol – as demonstrated through 5-deoxystrigol feeding assays - results in the formation of sorgomol (Motonami et al. 2013). Then, further oxidation of the

hydroxy group to a carboxylic acid and subsequent decarboxylation probably result in the formation of sorgolactone (Motonami et al. 2013).

Many of the above described enzymatic conversions have been postulated or shown to be catalyzed by members of the CYP711A class of P450s, MAX1 homologs (Yoneyama et al. 2018; Zhang et al. 2018; Abe et al. 2014; Zhang et al. 2014). Plant Cytochrome P450 monooxygenases (P450s) are a large superfamily of enzymes involved in many different pathways and are responsible for the diversification of many primary and secondary metabolites (Mizutani et al. 2011; Sezutsu et al. 2013). They catalyze a vast array of reactions such as oxidation, hydroxylation, isomerization and epoxidation often in a stereoselective manner (Schuler et al. 2003). Based on phylogenetic analysis, plant P450s are grouped into single- and multi-family clans. Single-family clans were shown to be involved in catalyzing enzymatic reactions of similar compounds (Nelson et al. 2011). For instance, CYP711, a single-family P450 that has one gene copy in most dicots but which is duplicated in monocots (Bak et al. 2011), has been shown to be involved in strigolactone biosynthesis (Abe et al. 2014). A single copy of CYP711, CYP711A1, was identified in *Arabidopsis*; a mutation in this gene causes an increased shoot branching phenotype (Booker et al. 2005), and the gene was therefore called *more axillary branched 1 (max1).* This P450 (MAX1) was later shown to use carlactone, the precursor of strigolactones, as a substrate and converts it to carlactonoic acid (Figure 1) (Abe et al. 2014; Booker et al. 2005; Stirnberg et al. 2002).

Carlactone is derived from all-*trans*-β-carotene, which is first converted to 9-*cis*-β-carotene by DWARF27 (D27), a *cis-trans* isomerase, followed by two subsequent enzymatic steps catalyzed by MORE AXILLARY GROWTH (MAX) genes; MAX3 and MAX4. The latter two are carotenoid dioxygenases CCD7 and CCD8 respectively. CCD7 cleaves 9-*cis*-β-carotene to 9-*cis*-β-apo-10'-carotenal that is further cleaved by CCD8 to form carlactone (Figure 1) (Nomura et al. 2018; Al-Babili et al. 2015; Xie et al. 2013; Wang et al. 2018). In rice, five *MAX1* homologs have been identified of which two, *Os01g0700900 (*Os900*)* and *Os01g0701400* (Os1400), have been functionally characterized. Os900 is a carlactone oxidase that catalyzes the conversion of carlactone to 4-deoxyorobanchol which is subsequently converted to orobanchol by Os1400 (Zhang et al. 2014).

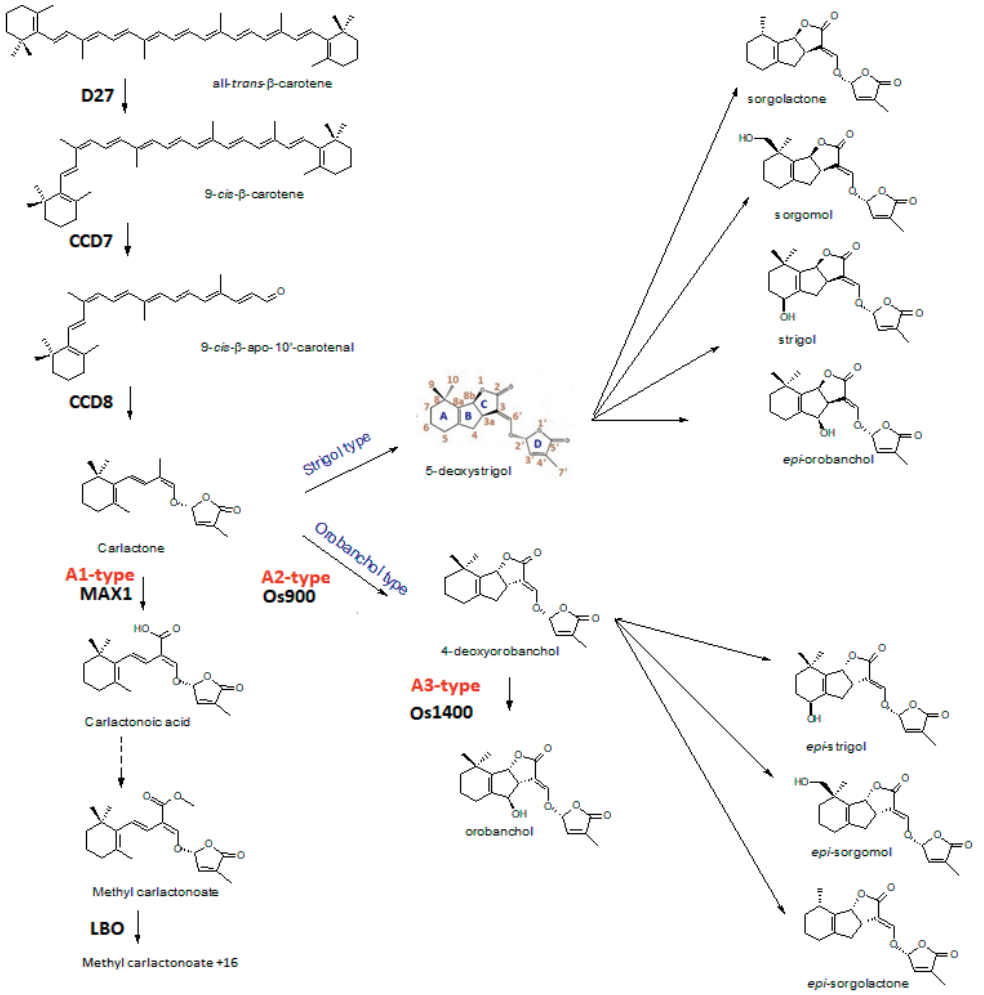


Figure 1. Schematic representation of part of the (postulated) strigolactone biosynthetic pathway. The A1-, A2 and A3-type of MAX1s that are postulated to be catalyzing several steps are marked in red adopted from (Yoneyama et al. 2018)

In recent work, MAX1 proteins were grouped into three types based on the enzymatic steps they catalyze. The A1-type catalyzes the conversion of carlactone to carlactonoic acid whereas the conversion of carlactone to 4-deoxyorobanchol (through carlactonoic acid) is catalyzed by A2-type MAX1s. A3-type MAX1s convert carlactone to carlactonoic acid and 4-deoxyorobanchol to orobanchol (Figure 1) (Yoneyama et al. 2018). Though several genes have been identified, the strigolactone biosynthetic pathway is far from completely elucidated. The formation of 5-deoxystrigol, for example, from carlactone has not been elucidated, nor have the genes been identified that convert 5-deoxystrigol to the strigol-type strigolactones (Figure 1). These type of strigolactones occur in most sorghum genotypes that are highly susceptible to *Striga*. Identifying these genes is crucial not only to understand the role of these genes in strigolactone diversification but also to use this knowledge as a tool to combat the important

parasitic weed *Striga.* In recent work, it has been shown that sorghum produces different types of strigolactones in a stereochemical manner, which, with the right type, results in low germination stimulant activity towards *Striga* (Gobena et al. 2017). Furthermore, more studies show that *Striga* species differ in their sensitivity to different strigolactones in a stereochemistry-specific manner. *S. gesnerioides*, for example, are more sensitive to orobanchol-type than strigol-type strigolactones while the reverse is true for *S. hermonthica*. The present study aims at characterizing the four sorghum MAX1 homologs and to identify/predict their substrates and role in strigolactone biosynthesis in sorghum. We characterized these genes by checking the expression profile in different parts of the plant. Furthermore, we looked at their response to phosphate starvation and their substrate preference using a transient gene expression assay in *Nicotiana benthamiana*.

Results

Phylogenetic tree of MAX1 homologs in sorghum

The five rice *MAX1* homologs (*Os01g0700900*, *Os01g0701400*, *Os01g0701500, Os02g0221900* and *Os06g0565100*) were used as baits for BLAST to find putative homologs in selected monocot plant species including *Sorghum bicolor (Challis et al. 2013)*. The amino acid sequences were blasted against the sorghum database on NCBI. This yielded four *MAX1* homologs, *Sb04g007880* (Sb7880), *Sb10g022310* (Sb2310), *Sb03g032210* (Sb2210) and *Sb03g032220* (Sb2220) (Challis et al. 2013). The corresponding AA sequences exhibited between 40 to 70% identity with the five rice MAX1 proteins. For instance, *Os01g0701400* displays 65% and 57% identity with Sb2220 and Sb2210, respectively and 52% and 45% identity with Sb2310 and Sb7880, respectively. The other functionally characterized MAX1 *Os01g0700900* displays 78%, 64%, 71% and 52% identity with Sb2220, Sb2310, Sb2210 and Sb7880, respectively. Homologs from other monocots were selected based on the highest hits obtained with BLAST. A phylogenetic tree was constructed using MAFFT after the AA sequences were aligned using MUSSLE.

In the phylogenetic tree, Sb7880 and Sb2310 occur in two different branches of the phylogenetic tree and cluster with *Os02g0221900* and *Os06g0565100,* respectively (Figure 2). The first clade also contains one of the three maize MAX1s, ZmMAX1c while the latter branch contains ZmMAX1a. The 3rd maize MAX1, ZmMAX1b, shares a clade with Sb2220 (Figure 2). *Sb2210* clusters separately from the other three in a clade with two *Brachypodium distachyon* and one *Hordeum vulgare* MAX1 homolog.

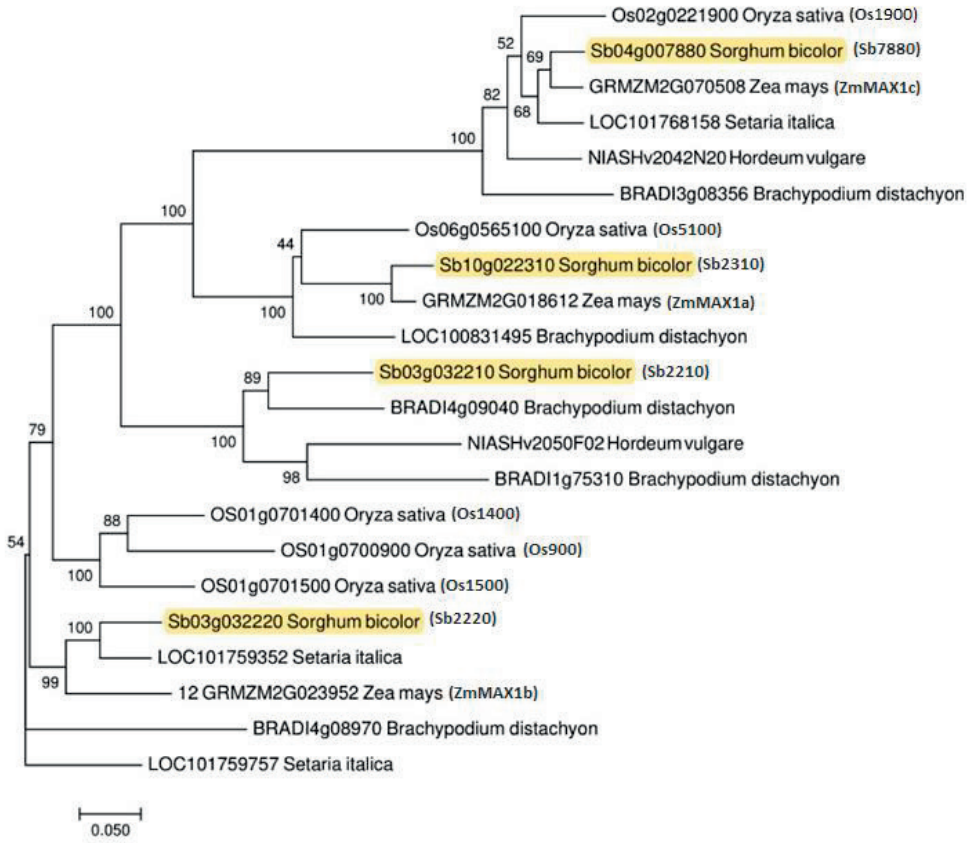


Figure 2. Phylogenetic tree of MAX1s in different monocot plants. Amino acid sequences were obtained from NCBI (https://www.ncbi.nlm.nih.gov/). They were aligned using MEGA 5 and the phylogenetic tree was constructed using IQ-TREE with neighbor-joining method with 1000 bootstraps. The four sorghum MAX1 homologs are shaded in yellow.

Response of sorghum MAX1 genes to phosphate starvation

Previous studies have shown that the production of strigolactones is upregulated by phosphate starvation and this is also reflected in the upregulation of biosynthetic genes. Therefore, we checked the effect of phosphate starvation on the expression level of the *MAX1* genes in roots of SRN39 plants. Plants were grown for the first two weeks watered with half-strength Hoagland's nutrient solution. To induce strigolactone production, the seedlings were exposed to phosphate-deficient conditions for one week. As shown in Figure 3, the expression of *Sb2310* was very low in both phosphate-starved and none-starved root materials. The expression of all *MAX1s* was significantly upregulated by phosphate starvation except *Sb2310*. (Figure 3C and D). Although not conclusively, these results suggest that Sb7880, Sb2210 and Sb2220 may be involved in strigolactone production in sorghum roots.

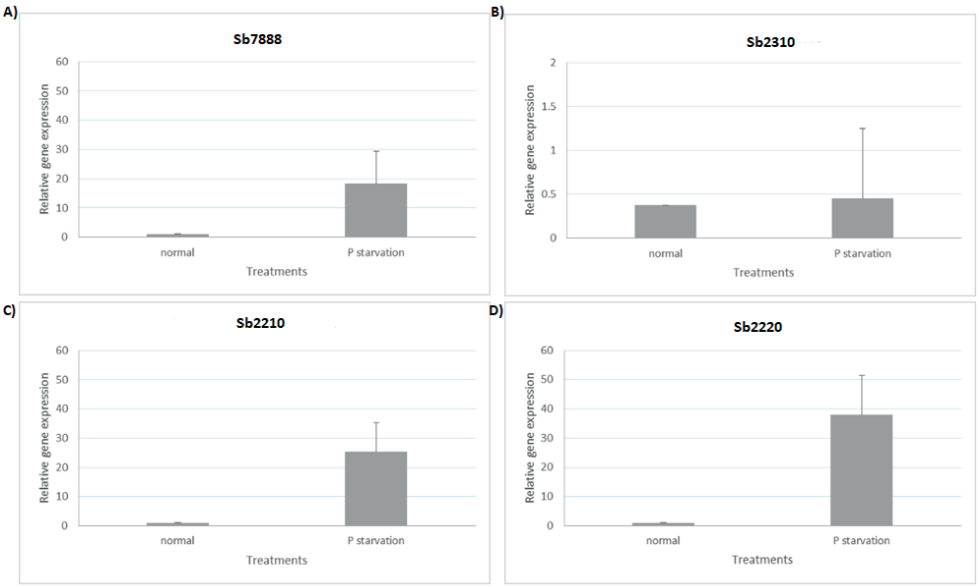


Figure 3. Relative gene expression of sorghum MAX1 homolog in roots of SRN39 that have been grown under normal and phosphate starvation conditions. A) *Sb04g007880* (*Sb7880*), B) *Sb10g022310* (*Sb2310*), C) *Sb03g032210* (*Sb2210*), D) *Sb03g032220* (*Sb2220*). Ubiquitin was used as a house keeping gene. The bars represent the average of three technical and three biological replicates + SE.

Expression patterns of MAX1 homologs in different plant parts of sorghum

To obtain more evidence supporting the involvement of the sorghum MAX1 homologs in strigolactone biosynthesis we looked at the expression of the four genes in different plant parts. Hereto, plants were grown until the head developed while watered with ½ strength Hoagland solution. Then, they were grown for one week under phosphate starvation conditions. We harvested the flower head, peduncle, leaf sheet, stem, axillary bud, lower stem and roots of SRN39, separately, and analyzed gene expression using qPCR. Figures 4A-D show the relative expression of each gene relative to the flower head.

In general, *Sb2220* had the highest relative expression followed by *Sb2210* (Figure 4). These two expressed higher in the lower stem than the other plant parts tested. The relative expression of these two was also significantly higher than the other *MAX1s*. On the other hand, *Sb2310* and *Sb7880* had a more than 10-fold lower expression than *Sb2220* and *Sb2210* in all plant parts tested. In general, *Sb7880* and *Sb2310* were mainly expressed in the above-ground parts of the plant while the other two were expressed higher in lower stem and, to some extent, roots (Figure 4A-D). These results suggest that *Sb2220* and *Sb2210* are the likely *MAX1* genes that catalyze the formation of sorghum strigolactones in the roots.

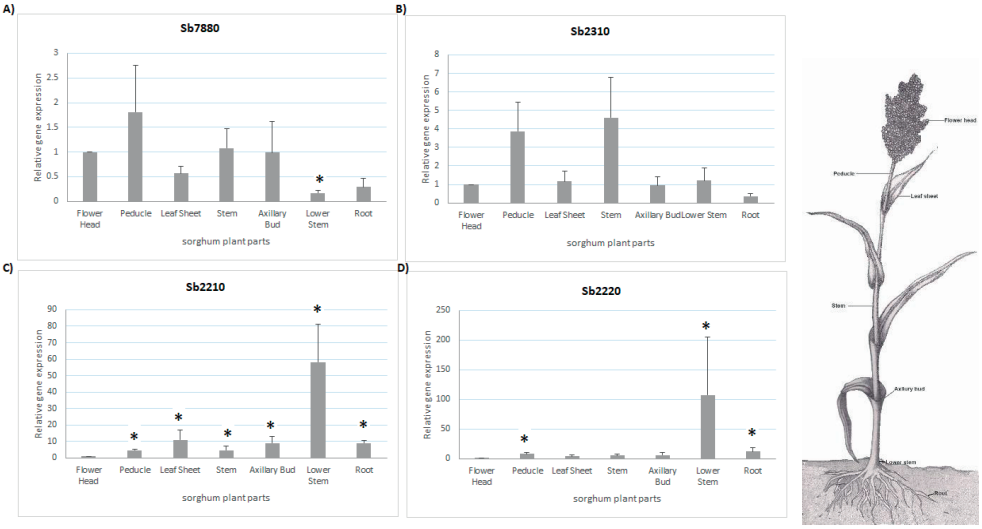


Figure 4. Relative gene expression of sorghum *MAX1* homologs in different plant parts of SRN39 that was grown under phosphate starvation for a week. A) *Sb04g007880* (*Sb7880*), B) *Sb10g022310* (*Sb2310*), C) *Sb03g032210* (*Sb2210*), D) *Sb03g032220* (*Sb2220*). Ubiquitin was used as a house keeping gene. "*" indicates if the relative expression of the gene is significantly different compared to its expression in the flower head. The bars represent average of three technical and three biological replicates + SE.

To further characterize the four sorghum MAX1 homologs, we looked at publicly available data to find highly connected genes in a co-expression network. We first looked at the MORKOSHI sorghum transcriptome database. In this database, the top 20 genes with similar expression to a gene of interest are given. Furthermore, a gene co-expression network is presented using RNAseq data obtained from publicly available datasets (Makita et al. 2015). The four sorghum MAX1s did not show co-expression with each other nor did they share a common co-expression network (Figure 5 and Table 1). Genes that were co-expressed with the four *MAX1s* are, for example, *Sb03g033290*, an ABC transporter family protein and two genes, *Sb06g025820* and *Sb03g040250*, annotated as cytochrome P450 that displayed co-expression with Sb2310 (Table 1). A gene annotated as 2-oxoglutarate and Fe (II)-dependent oxygenase (2-ODD) superfamily protein, *Sb05g005720*, was one of the genes that exhibited a similar expression pattern with Sb2220. In addition, *Sb03g040250*, cytochrome P450, exhibited a similar expression pattern with Sb2210. None of these genes nor any of the ones listed in Table 1 and Figure 5 have been shown to be involved in strigolactone biosynthesis nor the regulation of their expression. Even the putative strigolactone biosynthetic related genes from sorghum were not shown to have a similar expression with the MAX1s or appear in the co-expression analysis. To find more information on genes that are co-expressed with the four MAX1s homologs we surveyed Phytozome. In this database, the list of co-expressed genes is short usually one or two and the results found were different from the one obtained with the MORKOSHI sorghum transcriptome database. In the case of Sb2210, it is co-expressed with *Sb04g036900*, an O-methyltransferase while Sb2310 is co-expressed with *Sb01g050000*, a transcription factor, and another uncharacterized protein, *Sb07g023970*. Sb2220 was shown to be co-expressed with *Sb02g035990* which is annotated as X-BOX transcription factor

related. Since none of these genes were co-expressed with putative strigolactone biosynthesis genes, it did not help to identify the best candidate MAX1 homolog.

Table 1. Description of the top ten similarly expressed genes with the four *MAX1* homologs. The genes are listed in descending rank based on the JSD score described in http://sorghum.riken.jp/morokoshi/Data/Sobic.003G269500.html

	Functional annotation	KEGG pathway	GO (gene orthology)
Sb04g007880	cytochrome P450, putative, expressed		AT2G26170, OS02G0221900, GRMZM2G070508, GRMZM2G147119, BRADI3G08360
Sb01g033190	STRUBBELIG-receptor family 7		
Sb03g008970	C-8,7 sterol isomerase	Steroid biosynthesis Metabolic pathways Biosynthesis of secondary metabolites	
Sb08g018990	ARM repeat superfamily protein		
Sb09g023295			
Sb10g020820			
Sb01g030220	Major facilitator superfamily protein		
Sb09g000800	nodulin MtN21 /EamA-like transporter family protein		
Sb02g036220	UDP-glucosyl transferase 88A1		
Sb05g023240	UDP-Glycosyltransferase superfamily protein		
Sb08g016930	NB-ARC domain-containing disease resistance protein		
Sb10g022310	cytochrome P450, putative, expressed		AT2G26170, OS06G0565100, GRMZM2G018612, BRADI1G37730
Sb03g033290	ABC transporter family protein	ABC transporters	
Sb07g008540	Protein kinase superfamily protein		
Sb04g029170	NSP-interacting kinase 1		
Sb02g033000	basic helix-loop-helix (bHLH) DNA-binding superfamily protein		
Sb02g031610	glutathione S-transferase PHI 10		
Sb07g023970	Walls Are Thin 1		
Sb09g029750	O-acetyltransferase family protein		
Sb03g013750			
Sb01g038070	SCAR homolog 2		
Sb04g000860			
Sb03g032210	cytochrome P450, putative, expressed		AT2G26170, BRADI4G09040
Sb03g028190	UDP-Glycosyltransferase superfamily protein		
Sb04g010760	xyloglucan endotransglucosylase/hydrolase 12		
Sb06g025820	cytochrome P450, family 87, subfamily A, polypeptide 2		
Sb10g005530	Plant protein of unknown function (DUF868)		
Sb01g037520	basic region/leucine zipper motif 53		
Sb03g001530			
Sb02g023380	Major facilitator superfamily protein		
Sb03g040250	cytochrome P450, family 86, subfamily A, polypeptide 1		Cutin, suberine and wax biosynthesis
Sb06g000920		Zinc finger (C3HC4-type RING finger) family protein	
Sb05g003860		Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	
Sb03g032220	cytochrome P450, putative, expressed		AT2G26170, OS01G0701400, OS01G0701300, OS01G0700900, OS01G0701500, GRMZM2G023952, BRADI4G08970
Sb05g005720	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	Cysteine and methionine metabolism Metabolic pathways Biosynthesis of secondary metabolites	
Sb10g021077	RING/U-box superfamily protein		
Sb02g023370	Major facilitator superfamily protein		
Sb01g038120			
Sb03g029760	TRICHOME BIREFRINGENCE-LIKE 21		
Sb03g003545			
Sb04g009050	Predicted AT-hook DNA-binding family protein		
Sb05g022480	Laccase/Diphenol oxidase family protein		
Sb04g034440	WRKY DNA-binding protein 9		
Sb02g023380	Major facilitator superfamily protein		

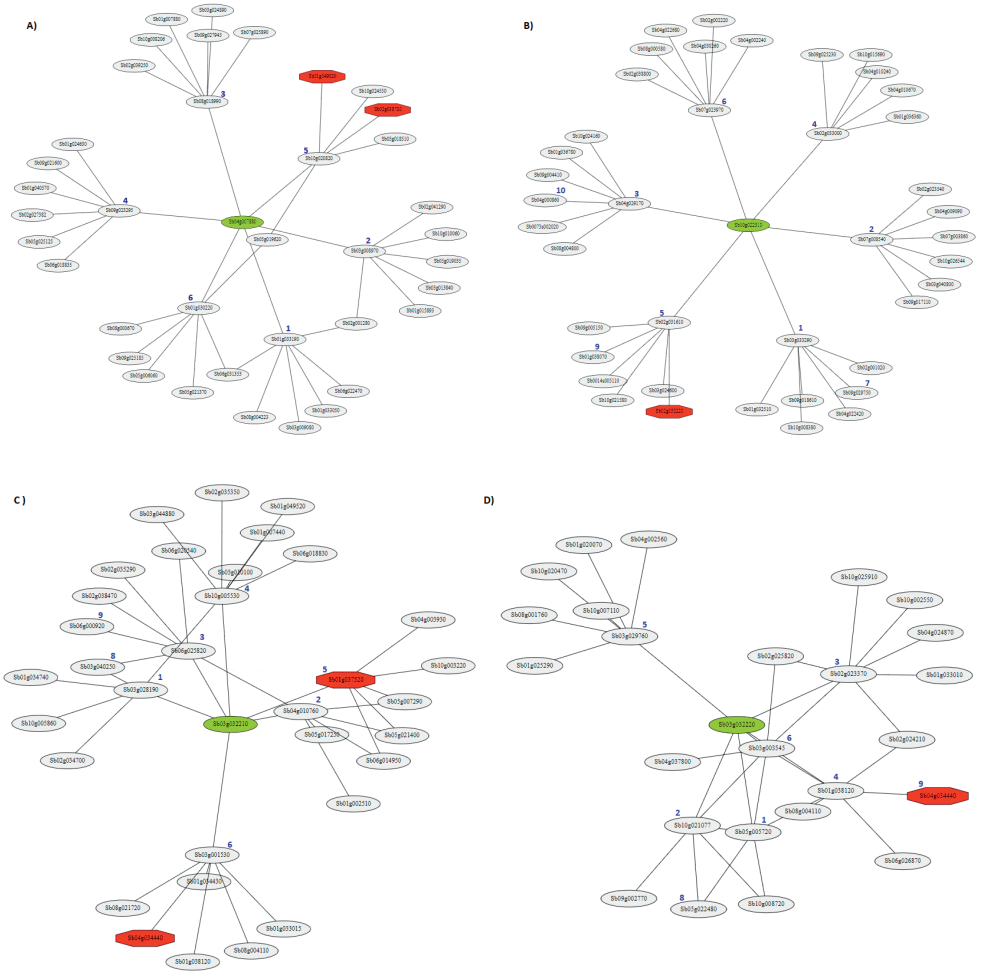


Figure 5. Gene co-expression network of A) *Sb04g007880* (Sb7880), B) *Sb10g022310* (Sb2310), C) *Sb03g032210* (Sb2210) and D) *Sb03g032220* (Sb2220). *Sb04g007880*, *Sb10g022310*, *Sb03g032210* and *Sb03g032220* are shaded in green. Transcription factors are shown in octagon. Similarly expressed genes are numbered in blue and these numbers indicates the rank of similarity in expression of the genes with either Sb04g007880 or *Sb10g022310* (see Table 1 for the complete list of genes that are similarly expressed with *Sb04g007880* and *Sb10g022310*). The picture was modified from (Makita et al. 2015).

Cloning and functional analysis of sorghum MAX1 homologs in N. benthamiana leaves.

In order to address the enzymatic functionality of the four sorghum MAX1 homologs, the genes were cloned from root RNA obtained from SRN39 that was exposed to phosphate starvation. For the transient assay in *N. benthamiana*, we overexpressed the rice strigolactone biosynthetic genes (OsD27, OsCCD7 and OsCCD8) using *A. tumefaciens* to produce carlactone which is the likely precursor for the sorghum MAX1s. Sorghum *MAX1* genes were added to this mix either individually or in combination. Three days after infiltration, the leaves of *N. benthamiana* were harvested and analyzed for product formation.

As shown in Figure 5, the transient expression of the rice carlactone pathway indeed resulted in the production of a large peak of carlactone. When the rice carlactone pathway was co-infiltrated with sorghum *MAX1s*, the carlactone level was strongly reduced by Sb2210 when compared to carlactone produced by the rice genes. Furthermore, a significant amount of carlactone was consumed by Sb2310. However, Sb2220 and Sb7880 did not affect the level of carlactone when compared to carlactone produced by the rice genes. Interestingly, the combination of Sb2310 and Sb7880 or combining one of these two genes with any of the other two *MAX1*s resulted in a significant reduction of carlactone (Figure 5). In this experiment, we also looked at possible products of sorghum MAX1 homologs such as 4-deoxyorobanchol, 5-deoxystrigol, orobanchol, sorgomol and sorgolactone. In addition, we also looked at carlactonoic acid. However, none of these strigolactones were found as an end product of any of the combinations tested. In separate experiments, we added Os900, that converts carlactone to 4-deoxyorobanchol (Zhang et al. 2014), in the blend to rule out the possibility of the MAX1s not having the right precursor for production of some strigolactones such as orobanchol, *epi*-strigol and *epi*-sorgomol (Figure 1). 4-deoxyorobanchol was indeed produced with this combination of genes (data not shown). However, none of the MAX1 homologs from sorghum were able to produce any known strigolactones from 4-deoxyorobanchol.

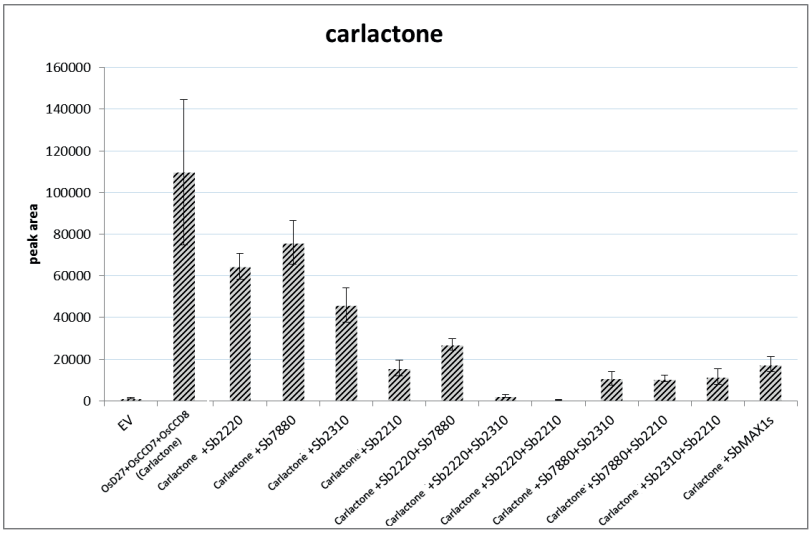


Figure 5. Carlactone profile in *N. benthamiana* leaves. For each combination, *OsD27*, *OsCCD7* (*Os04g550600*) and *OsCCD8* (*Os01g0706400*) from rice were co-infiltrated to produce carlactone. Single or combined sorghum MAX1 homologs were also added to this mix to check the consumption of carlactone. *Sb04g007880* (*Sb7880*), *Sb10g022310* (*Sb2310*), *Sb03g032210* (*Sb2210*) and *Sb03g032220* (*Sb2220*). The bars represent average of three biological replicates + SE.

Discussion

Several steps in the strigolactone biosynthesis pathway have been shown to be catalyzed by a single-family clan cytochrome P450, CYP711. Our study on the four sorghum MAX1 homologs shows that Sb2210 consumed carlactone (Figure 5). MAX1s can be group into three types, A1-, A2- and A3-type based on their enzymatic function (Yoneyama et al. 2018). The A1-type catalyzes the conversion of carlactone to carlactonoic acid whereas the conversion of carlactone to 4-deoxyorobanchol (through carlactonoic acid) is catalyzed by A2-type MAX1s. A3-type MAX1s convert carlactone to carlactonoic acid and 4-deoxyorobanchol to orobanchol (Yoneyama et al. 2018). In our transient assay, we combined Sb2210 with rice genes to produce precursors carlactone and 4-deoxyorobanchol. However, we did not find 4-deoxyorobanchol nor orobanchol when these rice genes were co-infiltrated with any of the sorghum MAX1s showing that there are no A2 and A3-type MAX1s in sorghum. The decrease in carlactone upon co-infiltration of the rice carlactone pathway with Sb2210 suggests that carlactone is converted to carlactonoic acid (see below), which suggests that Sb2210 is an A1-type MAX1.

The involvement of Sb2210 in strigolactone biosynthesis is supported by the fact that its expression is upregulated in response to phosphate deficiency (Figure 3). The symbiotic relationship between plants and AM fungi is a very old and important interaction. The presence of strigolactones in the rhizosphere, exuded by the plant, initiates hyphal branching in AM fungi (Lopez-Raez et al. 2011). The AM fungi provide the plant with accessible phosphate while relying on the plant for its carbon source. This interaction is induced when the plant is suffering from nutrient deficiency, especially phosphate. To achieve this, the level of strigolactones in both the plant and the rhizosphere is boosted through the upregulation of the expression of the biosynthetic genes. We also looked at the expression of the sorghum MAX1s in different plant parts. The expression of strigolactone biosynthetic genes such as D27, CCD7 and CCD8 is higher in stem than in roots (Wu et al. 2019). That is also what we observed for the expression of the two MAX1 homologs, Sb2210 and Sb2220 (Figure 4). Based on the localization of strigolactone biosynthesis genes, their production is also predicted to occur in the cortical elongation zone of the root, root vascular tissue and axillary buds (Ruyter-Spira et al. 2013). The expression of Sb2210 and Sb2220 was also significantly higher in axillary buds when compared to the flower head. This result suggests that they are good candidates for a role in phosphate-regulated strigolactone biosynthesis. The co-expression analysis of sorghum MAX1 homologs using publicly available data was intended to find more supporting evidence for the involvement of these genes in strigolactone biosynthesis. In many studies, genes that are co-expressed in a similar pattern (possibly because they are co-regulated by the same transcription factor) have been shown to be functionally related. Surprisingly, the list of genes obtained from different databases did not include any of the strigolactone biosynthesis-related genes. Moreover, none of the MAX1s were co-expressed with each other or show a similar expression pattern throughout the plant. One possible explanation is that the data collected in these databases are obtained from experiments that were not aimed at affecting

strigolactone production. For instance, if a similar analysis is conducted using RNAseq data from the roots of sorghum plants that were treated with phosphate starvation the outcome would likely be completely different. Likely all the genes that are involved in stress response of phosphate shortage, including strigolactone biosynthesis, would be co-expressed.

The other interesting candidate which behaves like Sb2210 in terms of expression pattern is Sb2220. This gene was upregulated by phosphate starvation and also expressed in lower stem and root (Figure 3 and 4). However, its expression in axillary buds was not significantly different and the corresponding protein did not consume carlactone in our transient assay. Surprisingly, when co-infiltrated with either Sb2310 or Sb2210, Sb2220 significantly increased the efficiency of carlactone consumption by the enzymes encoded by the former two genes when compared to their infiltration alone (Figure 5). This could be explained by a feedback regulation of carlactonoic acid on carlactone conversion and activity of Sb2220 downstream so in the conversion of carlactonoic acid to something else. In the case of Sb2210, it was not possible to find supporting evidence for this hypothesis in the phylogenetic tree since Sb2210 did not cluster with any of the functionally characterized rice or maize MAX1 homologs (Figure 2). On the other hand, Sb2220 was in the same clade with *ZmMAX1b* which is an A3-type MAX1 homolog from maize (Figure 2) (Yoneyama et al. 2018). Even though ZmMAX1b is the closest homolog to Os900, it was not able to convert carlactone to 4-deoxyorobanchol, but rather it converts 4-deoxyorobanchol to orobanchol (Yoneyama et al. 2018). The Sb2220, however, did not seem to catalyze any of these reactions.

From the recent work on MAX1s, it is clear that it is not always possible to rely on predicting the function of these genes based on a phylogenetic analysis. For instance, the three rice MAX1 homologs (Os900, Os1400 and Os1500) clustered in the same clade while they are A1-, A2- and A3-type, respectively (Yoneyama et al. 2018). Based on our analysis, we can predict that Sb2310 is the likely ortholog of *Os06g0565100* and *ZmMAX1a* while *Sb7880* is to *Os02g0221900* and *ZmMAX1c*. These rice and maize genes are A1-type MAX1s that can convert carlactone to carlactonoic acid. However, in our transient assay, Sb2310 and Sb7880 did not consume carlactone when infiltrated individually. Despite our effort to identify end-products using targeted strigolactone analysis, in our experiments, none of the sorghum MAX1s were able to produce canonical strigolactones. Our attempt to hunt products of these genes using untargeted metabolomics was not successful for three major reasons. First, by the time we were conducting the experiments, authentic standards of intermediate products such as carlactonoic acid, hydroxycarlactone and other unknown possible products of these genes were not available to confirm candidate peaks. Second, it is possible that we have not included key components for the enzymatic activity to happen. For instance, adding the sulfotransferase gene and a sulfate donor might result in producing the right precursor for the right MAX1 (see Chapter 3). The third possible explanation is that the intermediate product or products produced by these genes are conjugated by endogenous *N. benthamiana* genes. Several studies have shown similar observations during the characterization of biosynthetic genes using transient expression in *N. benthamiana* (Wang et al. 2016; Hofer et al. 2013; Miettinen et al. 2014). Particularly, acid products (such as the CLA that we expect to be

produced) were prone to conjugation by glycosylation. Metabolomics analyses of the products of the transient expression assays in *N. benthamiana* are underway to test this assumption. In conclusion, this study provides the basis for the further characterization of the sorghum MAX1s. More work is needed to identify the end products and with that the enzymatic functionality of these MAX1s. This should pave the way to a better understanding of the biological role of MAX1 duplication and – by that - strigolactone diversification in sorghum.

Materials and Methods

Plant material

Sorghum seeds were surface sterilized and germinated at 25^{o}C for 48 hrs in darkness on moistened filter paper. The germinated seeds were transferred to 14 cm pots (three plants per pot) filled with river sand. The seedlings were grown for two weeks in a climate chamber with temperature setting of 28^{o}C and 16 hrs light. Plant materials used to check the response of the sorghum MAX1 homologs to phosphate starvation were grown for a total of three weeks. The first two week, the seedlings were watered with half-strength Hoagland's nutrient solution. To check gene expression of sorghum *MAX1* genes in different plant parts, the sorghum plants were grown until the flower head was developed. In both cases, the pots were flushed with 1 L of phosphate deficient half strength Hoagland's to remove the phosphate from the pots. The plants were then grown for an additional week and watered with half strength Hoagland's without phosphate. After that week, a second flushing was used to rinse out any metabolites, including strigolactones, accumulated in the pot. Another 48 hrs later, the roots were collected, washed, frozen in liquid nitrogen and stored at -80^{o}C for further analysis.

Real-time quantitative RT-PCR (qRT-PCR)

To examine the relative expression of the four sorghum MAX1 homologs in different plant parts and to check their response to phosphate starvation, a total RNA was extracted from sorghum plants using RNeasy Mini Kit (Qiagen) and DNase-I kit (Qiagen). Subsequently, 2 μg of the total RNA was reverse transcribed to cDNA using Vio-Rad iScript cDNA Synthesis kit using the manufacture's instruction. qRT-PCR was performed using SYBR Green Supermix Bio-RAD using specifying primers (Table 1). Three technical and three biological replicates were used, and the relative expression of the genes were determined using comparative C_t value (Livak et al. 2001). A sorghum *Ubiquitin (Ubi)* gene and a phosphate marker gene was used as internal control and to normalize the data.

Plasmid construction

The full-length ORFs of the sorghum MAX1 homologs were obtained from NCBI database. The fragments were amplified by Phusion polymerase (New England Biolabs) from cDNA synthesized from sorghum genotype SRN39 using gene specific primers. Subsequently, the

fragments were transformed to pJET1.2 (Fermentas) vector and sent for sequencing. Then, it was subcloned to an entry vector PIVIA_2.1 containing CaMV35S promoter by digestion (with XhoI and BamHI enzymes) and ligation (Ting et al. 2013). LR reaction were performed to transfer the fragments to the binary vector pBin-Plus, using Gateway LR clonase II enzyme mix (Invitrogen) according to the manufacturer's instructions (van Engelen et al. 1995). Subsequently, the binary vector was introduced to *A. tumefaciens* AGL0 using electroporation. The rice biosynthetic pathway genes, OsD27, OsCCD7(*Os04g550600*), OsCCD8 (*Os01g0706400*) and Os900 (*Os01g0700900*) were kindly provided by Y. Zhang (Laboratory of plant physiology, Wageningen university, The Netherland)

Table 1. primer sequences used for cloning and gene expression analysis.

Primer name	Experiment	Sequences
XhoI_Sb7880-FW	Cloning	5'-CCG**CTCGAG**ATGGAGATTGCACTCACAGTT-3'
Sb7880_BamHI-RV	Cloning	5'-CG**GGATCC**TCAGCTTCTCCTGATGGC -3'
XhoI_Sb2310-1-FW	Cloning	5'-CCG**CTCGAG**ATGGAGATGGCAGGCGCC-3'
Sb2310_BamHI-RV	Cloning	5'-CG**GGATCC**TACGACATGGCAGCGC-3'
XhoI_Sb2210-1-FW	Cloning	5'-CTAG**CTCGAG**ATGGGGTGGGGAGAAATC-3'
Sb2210_BamHI-RV	Cloning	5'-CG**GGATCC**TCAGTTTTTCCTCTCGATGA-3'
XhoI_Sb2220-1-FW	Cloning	5'-CCG**CTCGAG**ATGGAGATGGGCACGGT-3'
Sb2220_BamHI-RV	Cloning	5'-CGC**GGATCC**TTAATTATTTTCCACATGCCTC-3'
Sb7880-FWP2	Gene expression	5'-AGGACTGCATCCCCTTCTG-3'
MAX1-RV	Gene expression	5'-GAGGAACTCCCTGACGTTGT-3'
Sb2310-FW	Gene expression	5'-ACGAAGGACTTCCTGTCAGC-3'
Sb2310-RV	Gene expression	5'-AGCCAGTAGCTTGGCCTCTA-3'
Sb2210-FW	Gene expression	5'-AGCGAAAGATCCCAAGGACT-3'
Sb2210-RV	Gene expression	5'-AGTTGCGGTAGAGGTGGATG-3'
Sb2220-FW	Gene expression	5'-GTGGAGATCGGTGGCTATGT-3'
Sb2220-RV	Gene expression	5'-AAAAGGGATGAACGCGTATG-3'
Ubi-FW	Gene expression	5'-GCTGTACCTGCGTTTGTCTG -3'
Ubi-RV	Gene expression	5'-ACACACGGGACACAAGACAC-3'

Transient assay in N. benthamiana leaves

In order to functionally characterize the four sorghum MAX1 homologs we use a transient assay in *N. benthamiana* leaves. For this assay, cells were harvested from liquid media culture that contains Agrobacterium strain cultures (AGL0) grown 48 hrs at 28°C at 200 rpm. The cells were resuspended to a final OD_{600} of 0.5 by 10 mM MES-KOH buffer (pH 5.7) containing 10 mM MgCl2 and 100 mM acetosyringone (49-hydroxy-3',5'-dimethoxyacetophenone; Sigma), followed by rolling at room temperature for 4 hrs. Different gene combinations were made by mixing equal concentrations of the agrobacterium strains containing different constructs. In each combination, strains carrying an empty vector were used to compensate for the dosage of each gene. An agrobacterium strain containing a gene encoding the TBSV P19 protein was added to each combination to maximize protein production by suppression of gene silencing (Voinnet et al. 2003).

For agroinfiltration, five weeks old *N. benthamiana* plants were used in greenhouse with 16 h light at 25°C and 8 h dark at 22°C. The construct was infiltrated into abaxial side of *N. benthamiana* leaves by using a 1 ml syringe. Three leaves per plant, as a technical replicate and four individual plants, as biological replicates were infiltrated per each combination of

constructs. The infiltrated leaves were harvested after six days and were frozen in liquid nitrogen and stored at -80^{o}C for further analysis.

Carlactone and strigolactone extraction and detection from N. benthamiana leaves

To extract carlactone and strigolactones, 500mg of agro-infiltrated *N. benthamiana* leaves were grinded and 2ml of ethyl acetate containing 0.1 µM internal standard GR24 was added. The samples were sonicated for 10-15 min in a Branson 3510 ultrasonic bath, centrifuged for 10 min at 2000 rpm and the supernatant transferred to a clean vial. An additional 2 ml of ethyl acetate were used to re-extract the pellet which was then centrifuged again for 10 min. The combined supernatants were evaporated to dryness and the residue re-dissolved in 50 µl ethyl acetate and 4 ml hexane. The samples were then loaded on a silica column (200 mg) (preconditioned with 2 ml ethyl acetate and 4 ml hexane). The column was eluted in fractions: fraction 1 (eluted with 4 ml 35:80 ethyl acetate: hexane) and fraction 2 (4 ml 90:20 EtOAc: hexane). Part of fraction 1 was used to measure carlactone, carlactonoic acid. The rest of fraction 1 and 2 were combined to measure strigolactones. The solvent was evaporated to dryness after which the residue was re-dissolved in 200 µl of 25% acetonitrile in water. The samples were filtered through a 0.45 µm Minisart SRP filter (Sartorius, Germany).

Strigolactones were then analyzed by comparing the retention times and mass transition with the internal standards using a Waters Xevo tandem quadrupole mass spectrometer. An Acquity UPLC BEH C18 column was used for chromatographic separation by applying a water: acetonitrile gradient with 0.1% formic acid to the column. For strigolactones, gradient was started with 5% acetonitrile for 0.33 min. Then, it was raised to 27% (v/v) acetonitrile in 0.34 min, followed by a 4.33 min gradient to 40% (v/v) acetonitrile, then rising to 65% acetonitrile in 3 min and maintained for 0.67 min. A 90% acetonitrile gradient replaced that in 0.2 min was used and was maintained for 0.46 min. The gradient was then set back to 5% acetonitrile in 0.2 min and maintained for 2.47 min to equilibrate the column prior to the next run. For carlactone detection, the gradient started from 5% acetonitrile or 0.15 min and raised to 60% (v/v) acetonitrile in 2.0 min, followed by a 5.25-min gradient to 90% (v/v) acetonitrile, which was maintained for 0.75 min before going back to 5% acetonitrile using a 0.25-min gradient and maintained for 1.85 min to equilibrate the column before the next run.

The column temperature was set at 50^{o}C with flow rate of 0.5 mL/min. 10 – 20 µl of samples were injected. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L/hr, respectively. The source temperature was set at 150^{o}C, the capillary voltage was set at 3.0 kV and the desolvation temperature at 650^{o}C. For each standard, the cone voltage was optimized using Waters IntelliStart MS Console. MRM was used for identification of strigolactones. Data were analyzed using MassLynx 4.1 (combined with TargetLynx) software.

References

Abe, Satoko, Aika Sado, Kai Tanaka, Takaya Kisugi, Kei Asami, Saeko Ota, Hyun Il Kim, et al. 2014. 'Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18084-89.

Akiyama, K., K. Matsuzaki, and H. Hayashi. 2005. 'Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi', *Nature*, 435: 824-7.

Al-Babili, Salim, and Harro J. Bouwmeester. 2015. 'Strigolactones, a Novel Carotenoid-Derived Plant Hormone', *Annual Review of Plant Biology*, 66: 161-86.

Awad, Ayman, Daisuke Sato, Dai Kusumoto, Hiroaki Kamioka, Yasutomo Takeuchi, and Koichi Yoneyama. 2006. 'Characterization of Strigolactones, Germination Stimulants for the Root Parasitic Plants Striga and Orobanche, Produced by Maize, Millet and Sorghum', *Plant Growth Regulation*, 48: 221-27.

Bak, Søren, Fred Beisson, Gerard Bishop, Björn Hamberger, René Höfer, Suzanne Paquette, and Danièle Werck-Reichhart. 2011. 'Cytochromes p450', *The arabidopsis book*, 9: e0144-e44.

Besserer, A., V. Puech-Pages, P. Kiefer, V. Gomez-Roldan, A. Jauneau, S. Roy, J. C. Portais, et al. 2006. 'Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria', *PLoS Biol*, 4: e226.

Booker, J., T. Sieberer, W. Wright, L. Williamson, B. Willett, P. Stirnberg, C. Turnbull, et al. 2005. 'MAX1 encodes a cytochrome P450 family member that acts downstream of MAX3/4 to produce a carotenoid-derived branch-inhibiting hormone', *Dev Cell*, 8: 443-9.

Bouwmeester, H. J., Chr Roux, J. A. Lopez Raez, and G. Bécard. 2007. "Rhizosphere communication of plants, parasitic plants and AM fungi." In *Trends in Plant Science 12 (2007) 5*. 2007.

Brewer, Philip B., Hinanit Koltai, and Christine A. Beveridge. 2013. 'Diverse roles of strigolactones in plant development', *Mol Plant*, 6: 18-28.

Challis, R. J., J. Hepworth, C. Mouchel, R. Waites, and O. Leyser. 2013. 'A role for more axillary growth1 (MAX1) in evolutionary diversity in strigolactone signaling upstream of MAX2', *Plant Physiol*, 161: 1885-902.

Cook, C. E., Leona P. Whichard, Beverly Turner, Monroe E. Wall, and Grant H. Egley. 1966. 'Germination of Witchweed (Striga lutea Lour.): Isolation and Properties of a Potent Stimulant', *Science*, 154: 1189.

Gobena, D., M. Shimels, P. J. Rich, C. Ruyter-Spira, H. Bouwmeester, S. Kanuganti, T. Mengiste, et al. 2017. 'Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes *Striga* resistance', *Proc Natl Acad Sci U S A*, 114: 4471-76.

Gomez-Roldan, M. V., S. Fermas, P. B. Brewer, V. Puech-Pages, E. A. Dun, J. P. Pillot, F. Letisse, et al. 2008. "Strigolactone inhibition of shoot branching." In *Nature 455 (2008)*. 2008.

Hofer, R., L. Dong, F. Andre, J. F. Ginglinger, R. Lugan, C. Gavira, S. Grec, et al. 2013. 'Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for engineering the early steps of the (seco)iridoid pathway', *Metab Eng*, 20: 221-32.

Koltai, Hinanit. 2011. 'Strigolactones are regulators of root development', *New Phytologist*, 190: 545-49.

Livak, K. J., and T. D. Schmittgen. 2001. 'Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method', *Methods*, 25: 402-8.

Lopez-Raez, J. A., T. Charnikhova, I. Fernandez, H. Bouwmeester, and M. J. Pozo. 2011. 'Arbuscular mycorrhizal symbiosis decreases strigolactone production in tomato', *J Plant Physiol*, 168: 294-7.

Lopez-Raez, J. A., T. Charnikhova, V. Gomez-Roldan, R. Matusova, W. Kohlen, R. De Vos, F. Verstappen, et al. 2008. 'Tomato strigolactones are derived from carotenoids and their biosynthesis is promoted by phosphate starvation', *New Phytol*, 178: 863-74.

Makita, Y., S. Shimada, M. Kawashima, T. Kondou-Kuriyama, T. Toyoda, and M. Matsui. 2015. 'MOROKOSHI: transcriptome database in Sorghum bicolor', *Plant Cell Physiol*, 56: e6.

Miettinen, Karel, Lemeng Dong, Nicolas Navrot, Thomas Schneider, Vincent Burlat, Jacob Pollier, Lotte Woittiez, et al. 2014. 'The seco-iridoid pathway from Catharanthus roseus', *Nature Communications*, 5: 3606.

Mizutani, M., and F. Sato. 2011. 'Unusual P450 reactions in plant secondary metabolism', *Arch Biochem Biophys*, 507: 194-203.

Motonami, Noriko, Kotomi Ueno, Hitomi Nakashima, Saki Nomura, Masaharu Mizutani, Hirosato Takikawa, and Yukihiro Sugimoto. 2013. 'The bioconversion of 5-deoxystrigol to sorgomol by the sorghum, Sorghum bicolor (L.) Moench', *Phytochemistry*, 93: 41-48.

Nelson, D., and D. Werck-Reichhart. 2011. 'A P450-centric view of plant evolution', *Plant J*, 66: 194-211.

Nomura, Takahito, Xiaonan Xie, Koichi Yoneyama, Kaori Yoneyama, Takaya Kisugi, Kohki Akiyama, Yoshifumi Nakatani, et al. 2018. 'Which are the major players, canonical or non-canonical strigolactones?', *Journal of Experimental Botany*, 69: 2231-39.

Ruyter-Spira, C., S. Al-Babili, S. van der Krol, and H. Bouwmeester. 2013. 'The biology of strigolactones', *Trends Plant Sci*, 18: 72-83.

Schuler, Mary A., and Daniele Werck-Reichhart. 2003. 'Functional Genomics of P450S', *Annual Review of Plant Biology*, 54: 629-67.

Sezutsu, H., G. Le Goff, and R. Feyereisen. 2013. 'Origins of P450 diversity', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368: 20120428.

Stirnberg, P., K. van De Sande, and H. M. Leyser. 2002. 'MAX1 and MAX2 control shoot lateral branching in Arabidopsis', *Development*, 129: 1131-41.

Ting, H. M., B. Wang, A. M. Ryden, L. Woittiez, T. van Herpen, F. W. Verstappen, C. Ruyter-Spira, et al. 2013. 'The metabolite chemotype of Nicotiana benthamiana transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage', *New Phytol*, 199: 352-66.

Umehara, M., A. Hanada, S. Yoshida, K. Akiyama, T. Arite, N. Takeda-Kamiya, H. Magome, et al. 2008. 'Inhibition of shoot branching by new terpenoid plant hormones', *Nature*, 455: 195-200.

Umehara, Mikihisa, Mengmeng Cao, Kohki Akiyama, Tomoki Akatsu, Yoshiya Seto, Atsushi Hanada, Weiqiang Li, et al. 2015. 'Structural Requirements of Strigolactones for Shoot Branching Inhibition in Rice and Arabidopsis', *Plant and Cell Physiology*, 56: 1059-72.

van Engelen, Fred A., Jos W. Molthoff, Anthony J. Conner, Jan-Peter Nap, Andy Pereira, and Willem J. Stiekema. 1995. 'pBINPLUS: An improved plant transformation vector based on pBIN19', *Transgenic Research*, 4: 288-90.

Voinnet, O., S. Rivas, P. Mestre, and D. Baulcombe. 2003. 'An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus', *Plant J*, 33: 949-56.

Wang, Bo, Arman Beyraghdar Kashkooli, Adrienne Sallets, Hieng-Ming Ting, Norbert C. A. de Ruijter, Linda Olofsson, Peter Brodelius, et al. 2016. 'Transient production of artemisinin in Nicotiana benthamiana is boosted by a specific lipid transfer protein from A. annua', *Metab Eng*, 38: 159-69.

Wang, Yanting, and Harro J. Bouwmeester. 2018. 'Structural diversity in the strigolactones', *Journal of Experimental Botany*, 69: 2219-30.

Wu, Han, Huihui Li, Hong Chen, Qi Qi, Qiangqiang Ding, Juan Xue, Jing Ding, et al. 2019. 'Identification and expression analysis of strigolactone biosynthetic and signaling genes reveal strigolactones are involved in fruit development of the woodland strawberry (*Fragaria vesca*)', *BMC Plant Biology*, 19: 73.

Xie, X., K. Yoneyama, T. Kisugi, K. Uchida, S. Ito, K. Akiyama, H. Hayashi, et al. 2013. 'Confirming stereochemical structures of strigolactones produced by rice and tobacco', *Mol Plant*, 6: 153-63.

Xie, Xiaonan. 2017. 'Structural diversity of strigolactones and their distribution in the plant kingdom', *Japanese Journal of Pesticide Science*, 42: 10-16.

Yoneyama, K., N. Mori, T. Sato, A. Yoda, X. Xie, M. Okamoto, M. Iwanaga, et al. 2018. 'Conversion of carlactone to carlactonoic acid is a conserved function of MAX1 homologs in strigolactone biosynthesis', *New Phytol*, 218: 1522-33.

Yoneyama, Kaori, Xiaonan Xie, Hitoshi Sekimoto, Yasutomo Takeuchi, Shin Ogasawara, Kohki Akiyama, Hideo Hayashi, et al. 2008. 'Strigolactones, host recognition signals for root parasitic plants and arbuscular mycorrhizal fungi, from Fabaceae plants', *New Phytologist*, 179: 484-94.

Zhang, Y., X. Cheng, Y. Wang, C. Diez-Simon, K. Flokova, A. Bimbo, H. J. Bouwmeester, et al. 2018. 'The tomato MAX1 homolog, SlMAX1, is involved in the biosynthesis of tomato strigolactones from carlactone', *New Phytol*, 219: 297-309.

Zhang, Y., A. D. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, et al. 2014. 'Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis', *Nat Chem Biol*, 10: 1028-33.

Chapter 6

General Discussion

Genetic diversification and host-based Striga resistance mechanisms

Infection by *Striga*, a root parasitic plant, can result in up to 100% loss of crops (Ejeta 2007b). *Striga* infestation is not only causing crop losses but also forcing farmers to abandon their land. In addition, *Striga* is spreading to new areas and broadening its host range, which can make it an even more serious challenge. One of the host plants is sorghum, a cereal crop grown globally both for feed and human consumption. In subtropical, semi-arid regions of Africa and Asia alone, it is a staple food for more than 500 million people (Are et al. 2011). Growing sorghum is convenient in these areas since it can withstand harsh environments and grows in hot and dry areas. However, losses due to *Striga* are a serious constraint for sorghum production, especially in the African continent.

Most of the damage caused by *Striga* occurs between attachment to the host plant and its emergence above-ground. Breaking the life cycle of the parasitic plant before it attaches to the host plant will therefore be the most effective control method. Studying resistance mechanisms is a promising resource in the global effort to find a sustainable solution to eradicate the *Striga* invasion. So far, induction of low germination and haustorium formation, releasing toxic compounds, mechanical barriers, hypersensitive response and incompatibility have been shown to be resistance mechanisms that reduce or prevent *Striga* infection (Ejeta 2007a; Mbuvi et al. 2017; Gurney et al. 2006; Li et al. 2009). These host resistance mechanisms operate either pre-attachment or post-attachment. The first one prevents the germination and attachment of the parasitic plant while the latter reduces the success of the parasitic plant upon attachment (Ejeta 2007b; Yoder et al. 2010). However, these host-resistance *Striga* control mechanisms are not a cure for all. Genetic variation in *Striga* brings about its ability to overcome resistance mechanisms, but there is also variation in infectivity in different environments that have been shown to affect the efficacy of these resistance mechanisms (Koyama 2000; Hearne 2009). Furthermore, the identification of resistance genes is a difficult and time-consuming process. In my thesis work, I focused on identification and characterization of the genes that control the resistance mechanism of low germination. One of the main finding of my thesis, described in Chapter 2, provides evidence that this pre-attachment resistance mechanism in sorghum is regulated by the type of strigolactones exuded into the rhizosphere. Moreover, in Chapter 2 and 3, I show that the major strigolactones present in *Striga* resistant and susceptible lines, that is orobanchol and 5-deoxystrigol, respectively, are stereochemically different.

Because strigolactones have several different biological functions, in nature, they may be under selection pressure to modify the blend of strigolactones in their roots and root exudates, so they remain attractive to arbuscular mycorrhizal (AM) fungi but are resistant to certain species of parasitic plants. All sorghum lines I tested produce and/or exude different blends of strigolactones. This has also been shown in other crops such as rice (Jamil, Charnikhova, et al. 2012; Kohlen et al. 2013). Since the discovery of the first strigolactone, over 20 different types have been identified in many different plant species. Their role in relation to AM fungi symbiosis and *Striga* infestation has also been reported by several studies. Some

studies have also shown the efficiency of individual strigolactones in inducing hyphal branching of AM fungi or inducing *Striga* germination in different concentrations. Therefore, it is crucial to check if a change in strigolactone profile affects other roles of these compounds. In Chapter 2, we showed that though the type of strigolactone exuded by *Striga* resistant and susceptible sorghum lines is different, the colonization percentage of sorghum plants with AM fungi remained the same. This is an important finding that suggests that breeding for low germination induction can be used without affecting the symbiotic relationship of the crop with AM fungi. Another factor affecting plants may be its interaction with rhizosphere microbes. The composition of the root microbial community is affected by different factors including soil type, cultivar and growth stage of the plant. Interestingly, recent work on the orobanchol producing *Striga* resistant sorghum line showed that the relative abundance of some bacterial strains is different from other – *Striga* sensitive - sorghum lines tested (Schlemper et al. 2017). Though the biological relevance and the mechanism behind this selective recruitment of the bacterial community is not clear yet, it is an important tool to consider in breeding programs.

In order to be able to efficiently use low germination as a tool to control *Striga* infection, it is important to identify and understand enzymatic steps and genes that are involved in the strigolactone biosynthesis pathway. So far, a number of enzymatic reactions and genes that encode these enzymes have been identified. The biosynthesis of strigolactones starts with the isomerization of all-*trans*-β-carotene to 9-*cis*-β-carotene by DWARF27 (D27); 9-*cis*-β-carotene is subsequently cleaved by Carotenoid Cleavage Dioxygenase 7 (CCD7) and CCD8, which results in the formation of carlactone, the precursor, as it now seems, for all strigolactones (Sorefan et al. 2003; Alder et al. 2012; Schwartz et al. 2004; Seto et al. 2014). Further diversification of strigolactones has been shown to involve different enzymatic reactions such as hydroxylation, oxidation and methylation (Brewer et al. 2016; Abe et al. 2014; Zhang et al. 2014; Motonami et al. 2013). Sulfation of strigolactones/strigolactone-like compounds is an interesting finding that is discussed in Chapter 3. The gene that regulates this enzymatic activity is identified as SbSOT4A encoded by *Sobic.005G213600* and its mutation (called *lgs*) results in sorghum resistance against *Striga* (Chapter 3). SRN39 is one of the sorghum lines that bares this mutation and therefore is a low germination inducing genotype. As a result of the mutation, it produces and exudes orobanchol as a major strigolactone while in high germination stimulant lines such as Shanqui-red 5-deoxystrigol is dominant (Chapter 2). Linking the role of this sulfotransferase to the strigolactone biosynthesis pathway is addressed in Chapter 3. Based on the current knowledge in plants, sulfotransferases are involved in sulfation of compound classes, such as brassinosteroids, coumarins, flavonoids, gibberellic acid, jasmonate, salicylic acid, glucosinolates, phenolic acids and terpenoids (Hirschmann et al. 2014). In Chapter 3, I postulated that an intact SbSOT4A sulfates C18-hydroxycarlactone that is further converted to sulfated C18-hydroxycarlactonoic acid by oxidation at the C19 position. The ring closure after the removal of the sulfate group could then lead to the production of 5-deoxystrigol (Figure 1). This attractive mechanism needs now to be supported with enzymatic assays and further studies to better understand its biological relevance.

In addition to SbSOT4A, other genes are also involved in the diversification of strigolactones. Previous studies on strigolactone biosynthesis showed that several biosynthetic steps are catalyzed by cytochrome P450s (Sorefan et al. 2003). In Arabidopsis, MAX1 has been shown to convert carlactone to carlactonoic acid. Two rice MAX1 homologs have been shown to catalyze consecutive steps in converting carlactone to 4-deoxyorobanchol and 4-deoxyorobanchol to orobanchol (Zhang et al. 2014). In a recent study, more MAX1s were characterized in tomato and maize (Zhang et al. 2018; Yoneyama et al. 2018). In chapter 5, I characterized the four putative sorghum MAX1 homologs. I used *Agrobacterium*-mediated transient expression in *N. benthamiana* leaves to reconstruct the strigolactone pathway using already identified rice genes (Zhang et al. 2014). This transient assay was shown in several studies to be a great tool to study the function of biosynthetic pathway genes including in strigolactone biosynthesis (Hofer et al. 2013; Miettinen et al. 2014; Reed et al. 2018; Zhang et al. 2014). It is a fast and effective tool to reconstruct a biosynthetic pathway and to produce the possible substrate for a candidate gene of interest. As a result, I was able to show that carlactone is a substrate for two of the four sorghum MAX1 homologs.

In addition, using this approach facilitates the co-infiltration of different combinations of genes of interest in a single experiment. As shown in Chapter 5, the co-infiltration of Sb2210 or Sb2310 with the carlactone pathway resulted in the consumption of carlactone. Though it is not possible to rule out the role of Sb2310 and the other MAX1 homologs, based on our result it is likely that Sb2210 is involved in the strigolactone biosynthetic pathway. In general, MAX1s can be grouped into three types, A1, A2 and A3 based on their enzymatic activity. The combination of the results from the phylogenetic and gene (co-)expression analysis and the transient assay suggests that Sb2210 is an A1-type MAX1 that converts carlactone to carlactonoic acid (Figure 1). However, despite our effort to detect it, we were not able to show the production of carlactonoic acid in the *N. benthamiana* leaves that were infiltrated with Sb2210 with the carlactone pathway. A possible explanation is the further conjugation or conversion of carlactonoic acid to anther compound by *N. benthamiana* enzymes (see below). We also looked at the production of other possible products such as 4-deoxyorobanchol, orobanchol, 5-deoxystrigol, sorgomol and strigol. However, none of these were detected in our experiment. In case of sorgomol and strigol, their postulated precursor 5-deoxystrigol might not be present in our *N. benthamiana* leaves since it was not produced by any of the *MAX1* gene combinations tested. On the other hand, we also co-infiltrated Os900 with the carlactone pathway. Os900 was demonstrated to produce 4-deoxyorobanchol from carlactone, in rice the precursor for orobanchol formation by a second MAX1 homolog, Os1400 (Zhang et al. 2014). However, none of the sorghum MAX1 combinations we tried resulted in the conversion of 4-deoxyorobanchol to orobanchol. The hunt to find strigolactone biosynthetic genes in sorghum is not yet over.

Unlike other experimental setups such as expression of a gene of interest in *E. coli* or yeast and then feeding the substrate, one of the challenges of the *N. benthamiana* assay is, it is only applicable if all genes upstream of the gene of interest have already been identified. In chapter 4, for example, I studied sorgomol biosynthesis, one of the main strigolactones produced by

sorghum and a strong inducer of *Striga* germination. However, we were not able to use the *N. benthamiana* system to test if any of the candidate genes catalyze the production of sorgomol since the gene that converts carlactone to the putative, precursor of sorgomol, 5-deoxystrigol, has not yet been identified. A possible solution to this problem would be to infiltrate the precursor, 5-deoxystrigol, into *N. benthamiana* leaves a number of days after infiltration of the candidate gene of which we anticipate the encoded enzyme may use it as a precursor. Using the transient assay in *N. benthamiana* has more challenges. For instance, endogenous enzymes of *N. benthamiana* can conjugate products of agroinfiltrated pathways. As a result, it will not be possible to detect a product with the expected mass transitions. This was observed in similar studies on the seco-iridoid biosynthesis pathway using transient expression in *N. benthamiana*, where they found that certain pathway products are conjugated by endogenous *N. benthamiana* enzymes (Wang et al. 2016; Hofer et al. 2013; Miettinen et al. 2014). In case we are using this system to identify the end product of two or more genes that possibly catalyzes consecutive steps, such as the MAX1 homologs in Chapter 5, the intermediates could be converted/conjugated to another product by *N. benthamiana* which would potentially take away the substrate of the next gene in the biosynthesis pathway. However, in the similar studies mentioned above, it was demonstrated that the affinity of the real pathway enzymes is so much higher than that of the aspecific *N. benthamiana* enzymes that the biosynthesis pathway proceeded with the next step, thus preventing conjugation, at least to a large extent.

Mapping to identify genes that regulate certain traits is a time-consuming process that might result in an extended list of candidate genes. Techniques such as genome sequencing with bulk segregant analysis and RNAseq are considered as a new way to accelerate this process. In chapter 4, I applied both of these techniques to identify candidate genes that regulate sorgomol biosynthesis in sorghum. Using the RNAseq approach, I searched for differentially expressed genes in high- and low-sorgomol producing lines. The replicates in each sample were obtained from a pool of three sorghum lines. Considering that the sorghum lines used for the bulk were RIL lines that were obtained from parental lines that strongly differ in sorgomol production, a reliable and shortlist of candidate genes was expected. However, more than 30 candidate genes were found of which most were not functionally annotated. Therefore, further study was needed to narrow down candidate genes, which is through bulk segregant genome resequencing. Though puzzled by the none-overlapping list of candidate genes that came out of these two approaches, combining these techniques was a successful approach to identify two candidate genes. The result strongly suggests that both candidate genes might affect sorgomol production and likely *Sobic.001G319900* regulates the level of production while *Sobic.008G106200* is catalyzing the conversation of 5-deoxystrigol to sorgomol (Figure 1).

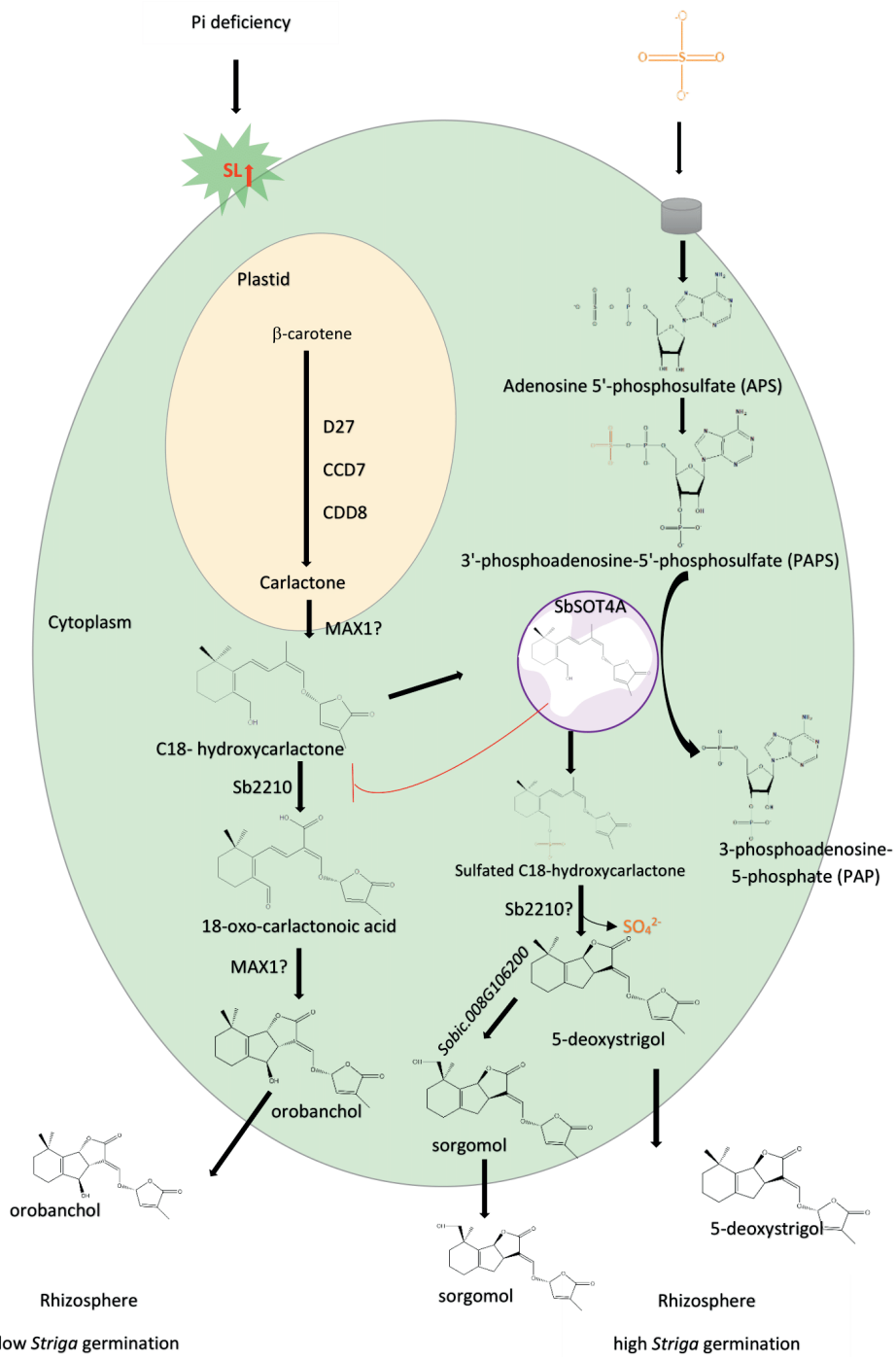


Figure 1. Proposed model of strigolactone biosynthesis in sorghum. In strigolactone biosynthesis pathway, carlactone is produced in the plastid from β-carlactone after consecutive steps catalyzed D27, CCD7 and CCD8 (yellow circle). Once carlactone is in the cytoplasm, it is converted to C18-

hydroxycarlactone possibly by a MAX1 homolog or other cytochrome P450. Then, it is either further converted to 18-oxo-carlactonoic acid or sulfated by SbSOT4A. The former will spontaneously catalyzed to orobanchol, resulting in Striga resistant sorghum. In the latter case, when C18-hydroxycarlactone is sulfated (shown in purple circle) upon loss of the sulphate group ring closure will result in 5-deoxystrigol formation, which makes sorghum susceptible to Striga. 5-Deoxystrigol also serves as precursor for formation of sorgomol by *Sobic.008G106200*.

Future perspectives

My finding that a sorghum sulfotransferase is regulating the strigolactone profile and that this results in resistance of the crop towards *Striga* infestation can be used as a tool to improve control of the parasitic weed. One application is to use this finding in breeding programs. Producing transgenic sorghum lines is difficult due to several reasons such as low frequency of regeneration of transgenic plants. Nevertheless, a number of groups and companies can now transform sorghum. Recent developments in genome editing such as RNA-silencing, Targeting Induced Local Lesion IN Genomes (TILLING) and CRISPR-cas can be used to confirm the model proposed in Chapter 3 (Xin et al. 2008; Che et al. 2018; Liu et al. 2019). Sulfation of strigolactones can also be translated into a direct control measure to control *Striga* infection. For instance, as shown in Chapter 2, an exogenous application of a sulfotransferase inhibitor can alter the strigolactone profile. As a result of treatment with Triclosan, the level of orobanchol increased significantly compared to plants that were not treated. This finding and identification of other inhibitors needs further investigation to be developed into a tool to control *Striga*. In addition, its possible application in other crops needs further study. In sorghum, 25 sulfotransferase homologs have been identified (Chapter 3). However, their function is not known yet. Therefore, it is important to carefully study the impact of exogenous sulfotransferase inhibitor application on different biological processes. A wide range of substrate acceptance can also be used as an advantage where we can exogenously apply other compounds. In this way, we can prevent the binding of strigolactones/-like compounds to the SbSOT4A. However, to apply this methodology further study on the function and the substrates of the other sulfotransferase homologs in sorghum needs to be conducted.

Application of integrated *Striga* management methods is key to having long-lasting control. This ranges from improving soil fertility to cultural practice and using resistant cultivars. For this, additional resources can be obtained from wild sorghum. Recently, wild sorghum accessions have been identified to be resistant against *Striga* (Mbuvi et al. 2017). These sorghum lines have been shown to display delayed *Striga* haustorium penetration. In addition, one of the wild sorghum accessions produced secondary metabolites in response to *Striga* infestation (Mbuvi et al. 2017). These findings show that resistance mechanisms are present in nature, which have not been uncovered yet. Integrating these findings into breeding programs can be a tool to develop a long-term strategy for *Striga* control.

In addition to these aspects, the success of *Striga* control depends on the success of adoption of newly released cultivars by farmers. Therefore, the opinion of the farmers in the proposed *Striga* control is crucial. My field visit in Ethiopia in *Striga* infested areas where farmers used *Striga* resistant sorghum lines showed me how this is a very important factor in *Striga* control.

The selection of *Striga* resistant cultivars should consider farmers preferred traits that considers their culture and also the main purpose of why they grow sorghum. For instance, in some areas, farmers grow sorghum for human consumption, but feed is also an important use. Therefore, *Striga* resistant cultivars with high biomass are preferred. In other areas, sorghum is grown intercropped with other cash crops such as soybean, haricot bean and cowpea. In this case, leafy sorghum genotypes are not preferred due to their shading effect that will reduce the performance of the intercropped crops. In some regions depending on what they will make from the sorghum flour, the grain color is another important factor for farmers to be willing to grow a sorghum line.

Another aspect of integrated *Striga* control can be obtained from *Striga* itself or from the soil environment. In the case of *Striga*, understanding the mechanism and identifying the genes behind its parasitism could lead to additional tools that can be used in control measures. For instance, the strigolactone receptors in *Striga* seeds have been identified and characterized in different studies (Hamiaux et al. 2012; Yao et al. 2016; Tsuchiya et al. 2018). Understanding the sensitivity and selectivity of these receptors can be used as a tool to develop methods that can interfere with the signaling process. For instance, ShHTL7, a strigolactone receptor in *Striga hermonthica*, has been shown to be the most sensitive receptor that plays an important role in the germination of the *Striga* seeds (Yao et al. 2017). Inhibitors of this receptor were shown to successfully inhibit *Striga* germination (Mashita et al. 2016; Hamiaux et al. 2018). When it comes to the soil environment, the key component will be the microbes present, their composition and their role in the interaction between the parasitic plant and the host plant. In a study on growth-promoting bacteria of sorghum, two strains of *Azospirillum brasilense* that were isolated from the sorghum rhizosphere reduced *Striga* germination (Miché et al. 2000). (Ahonsi et al. 2002) also showed a significant reduction of *Striga hermonthica* germination by 15 *Pseudomonas fluorescens /Pseudomonas putida* strains isolated from weed suppressive soils. The recent work on bacterial community compositions in soil revealed that SRN39 – the *Striga* resistant sorghum line - recruited a higher relative abundance of some bacteria such as *Cidobacteria GP1*, *Burkholderia, Cupriavidus (Burkholderiaceae), Acidovorax* and *Albidiferax* (*Comamonadaceae*) compared to susceptible lines (Schlemper et al. 2017). These results suggest an additional signaling role for the strigolactones in recruiting other microbes in addition to AM fungi. These findings need further research on the importance of the rhizosphere community and the mechanism behind it, also under different environmental conditions. Understanding the effect of strigolactones on the soil microbiome can be a useful tool that can be manipulated and integrated into the effort to control *Striga* infection. Therefore, fundamental research on the diversification of strigolactones and their biological role is crucial.

References

Abe, Satoko, Aika Sado, Kai Tanaka, Takaya Kisugi, Kei Asami, Saeko Ota, Hyun Il Kim, et al. 2014. 'Carlactone is converted to carlactonoic acid by MAX1 in Arabidopsis and its methyl ester can directly interact with AtD14 in vitro', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 18084-89.

Ahonsi, Monday O., Dana K. Berner, Alphonse M. Emechebe, and Segun T. Lagoke. 2002. 'Selection of rhizobacterial strains for suppression of germination of *Striga hermonthica* (Del.) Benth. seeds', *Biological control*, 24: 143-52.

Alder, Adrian, Muhammad Jamil, Mattia Marzorati, Mark Bruno, Martina Vermathen, Peter Bigler, Sandro Ghisla, et al. 2012. *The Path from -Carotene to Carlactone, a Strigolactone-Like Plant Hormone*.

Are, Ashok, Belum Venkata, Subba Reddy, H. C. Sharma, Jr Charles Hash, Srinivasa Rao Pinnamaneni, Bhavanasi Ramaiah, et al. 2011. *Recent advances in sorghum genetic enhancement research at ICRISAT*.

Brewer, P. B., K. Yoneyama, F. Filardo, E. Meyers, A. Scaffidi, T. Frickey, K. Akiyama, et al. 2016. 'LATERAL BRANCHING OXIDOREDUCTASE acts in the final stages of strigolactone biosynthesis in Arabidopsis', *Proc Natl Acad Sci U S A*, 113: 6301-6.

Che, P., A. Anand, E. Wu, J. D. Sander, M. K. Simon, W. Zhu, A. L. Sigmund, et al. 2018. 'Developing a flexible, high-efficiency Agrobacterium-mediated sorghum transformation system with broad application', *Plant Biotechnol J*, 16: 1388-95.

Ejeta, Gebisa. 2007a. 'Breeding for Striga Resistance in Sorghum: Exploitation of an Intricate Host-Parasite Biology', *Crop Science*, 47: S216-S27.

Ejeta, Gebisa. 2007b. *The Striga scourge in Africa: A growing pandemic*. in, *Integrating new technologies for Striga control: towards ending the witch-hunt* (World Scientific).

Gurney, A. L., J. Slate, M. C. Press, and J. D. Scholes. 2006. 'A novel form of resistance in rice to the angiosperm parasite Striga hermonthica', *New Phytol*, 169: 199-208.

Hamiaux, Cyril, Revel S. M. Drummond, Bart J. Janssen, Susan E. Ledger, Janine M. Cooney, Richard D. Newcomb, and Kimberley C. Snowden. 2012. 'DAD2 is an α/β hydrolase likely to be involved in the perception of the plant branching hormone, strigolactone', *Current Biology*, 22: 2032-36.

Hamiaux, Cyril, Revel S. M. Drummond, Zhiwei Luo, Hui Wen Lee, Prachi Sharma, Bart J. Janssen, Nigel B. Perry, et al. 2018. 'Inhibition of strigolactone receptors by N-phenylanthranilic acid derivatives: structural and functional insights', *Journal of Biological Chemistry*, 293: 6530-43.

Hearne, Sarah J. 2009. 'Control—the Striga conundrum', *Pest Manag Sci*, 65: 603-14.

Hirschmann, F., F. Krause, and J. Papenbrock. 2014. 'The multi-protein family of sulfotransferases in plants: composition, occurrence, substrate specificity, and functions', *Front Plant Sci*, 5: 556.

Hofer, R., L. Dong, F. Andre, J. F. Ginglinger, R. Lugan, C. Gavira, S. Grec, et al. 2013. 'Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for engineering the early steps of the (seco)iridoid pathway', *Metab Eng*, 20: 221-32.

Jamil, Muhammad, Tatsiana Charnikhova, Benyamin Houshyani, Aad van Ast, and Harro J. Bouwmeester. 2012. 'Genetic variation in strigolactone production and tillering in rice and its effect on *Striga hermonthica* infection', *Planta*, 235: 473-84.

Kohlen, Wouter, Tatsiana Charnikhova, Ralph Bours, Juan A. López-Ráez, and Harro Bouwmeester. 2013. 'Tomato strigolactones: a more detailed look', *Plant Signaling & Behavior*, 8: e22785-e85.

Koyama, M. L. 2000. "In BIG Haussmann, et al., Eds., Breeding for Striga Resistance in Cereals." In.: Margraf Verlag, Weikersheim, Germany.

Li, J., and M. P. Timko. 2009. 'Gene-for-gene resistance in Striga-cowpea associations', *Science*, 325: 1094.

Liu, G., J. Li, and I. D. Godwin. 2019. 'Genome Editing by CRISPR/Cas9 in Sorghum Through Biolistic Bombardment', *Methods Mol Biol*, 1931: 169-83.

Mashita, Okishi, Hikaru Koishihara, Kosuke Fukui, Hidemitsu Nakamura, and Tadao Asami. 2016. 'Discovery and identification of 2-methoxy-1-naphthaldehyde as a novel strigolactone-signaling inhibitor', *Journal of Pesticide Science*: D16-028.

Mbuvi, D. A., C. W. Masiga, E. Kuria, J. Masanga, M. Wamalwa, A. Mohamed, D. A. Odeny, et al. 2017. 'Novel Sources of Witchweed (Striga) Resistance from Wild Sorghum Accessions', *Front Plant Sci*, 8: 116.

Miché, L., M. L. Bouillant, R. Rohr, G. Sallé, and R. Bally. 2000. 'Physiological and Cytological Studies on the Inhibition of Striga Seed Germination by the Plant Growth-promoting Bacterium Azospirillum brasilense', *European Journal of Plant Pathology*, 106: 347-51.

Miettinen, Karel, Lemeng Dong, Nicolas Navrot, Thomas Schneider, Vincent Burlat, Jacob Pollier, Lotte Woittiez, et al. 2014. 'The seco-iridoid pathway from Catharanthus roseus', *Nature Communications*, 5: 3606.

Motonami, Noriko, Kotomi Ueno, Hitomi Nakashima, Saki Nomura, Masaharu Mizutani, Hirosato Takikawa, and Yukihiro Sugimoto. 2013. 'The bioconversion of 5-deoxystrigol to sorgomol by the sorghum, Sorghum bicolor (L.) Moench', *Phytochemistry*, 93: 41-48.

Reed, James, and Anne Osbourn. 2018. 'Engineering terpenoid production through transient expression in Nicotiana benthamiana', *Plant Cell Rep*, 37: 1431-41.

Schlemper, Thiago R., Márcio F. A. Leite, Adriano R. Lucheta, Mahdere Shimels, Harro J. Bouwmeester, Johannes A. van Veen, and Eiko E. Kuramae. 2017. 'Rhizobacterial community structure differences among sorghum cultivars in different growth stages and soils', *FEMS Microbiology Ecology*, 93.

Schwartz, S. H., X. Qin, and M. C. Loewen. 2004. 'The biochemical characterization of two carotenoid cleavage enzymes from Arabidopsis indicates that a carotenoid-derived compound inhibits lateral branching', *J Biol Chem*, 279: 46940-5.

Seto, Yoshiya, Aika Sado, Kei Asami, Atsushi Hanada, Mikihisa Umehara, Kohki Akiyama, and Shinjiro Yamaguchi. 2014. 'Carlactone is an endogenous biosynthetic precursor for strigolactones', *Proceedings of the National Academy of Sciences*, 111: 1640.

Sorefan, K., J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, E. Foo, S. Chatfield, et al. 2003. 'MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea', *Genes Dev*, 17: 1469-74.

Tsuchiya, Yuichiro, Masahiko Yoshimura, and Shinya Hagihara. 2018. 'The dynamics of strigolactone perception in Striga hermonthica: a working hypothesis', *Journal of Experimental Botany*, 69: 2281-90.

Wang, Bo, Arman Beyraghdar Kashkooli, Adrienne Sallets, Hieng-Ming Ting, Norbert C. A. de Ruijter, Linda Olofsson, Peter Brodelius, et al. 2016. 'Transient production of artemisinin in Nicotiana benthamiana is boosted by a specific lipid transfer protein from A. annua', *Metab Eng*, 38: 159-69.

Xin, Zhanguo, Ming Li Wang, Noelle A. Barkley, Gloria Burow, Cleve Franks, Gary Pederson, and John Burke. 2008. 'Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population', *BMC Plant Biology*, 8: 103.

Yao, Ruifeng, Zhenhua Ming, Liming Yan, Suhua Li, Fei Wang, Sui Ma, Caiting Yu, et al. 2016. 'DWARF14 is a non-canonical hormone receptor for strigolactone', *Nature*, 536: 469.

Yao, Ruifeng, Fei Wang, Zhenhua Ming, Xiaoxi Du, Li Chen, Yupei Wang, Wenhao Zhang, et al. 2017. 'ShHTL7 is a non-canonical receptor for strigolactones in root parasitic weeds', *Cell Research*, 27: 838.

Yoder, John I., and Julie D. Scholes. 2010. 'Host plant resistance to parasitic weeds; recent progress and bottlenecks', *Curr Opin Plant Biol*, 13: 478-84.

Yoneyama, K., N. Mori, T. Sato, A. Yoda, X. Xie, M. Okamoto, M. Iwanaga, et al. 2018. 'Conversion of carlactone to carlactonoic acid is a conserved function of MAX1 homologs in strigolactone biosynthesis', *New Phytol*, 218: 1522-33.

Zhang, Y., X. Cheng, Y. Wang, C. Diez-Simon, K. Flokova, A. Bimbo, H. J. Bouwmeester, et al. 2018. 'The tomato MAX1 homolog, SlMAX1, is involved in the biosynthesis of tomato strigolactones from carlactone', *New Phytol*, 219: 297-309.

Zhang, Y., A. D. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, et al. 2014. 'Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis', *Nat Chem Biol*, 10: 1028-33.

Summary

Striga hermonthica (*Striga*) is a parasitic plant that attaches to the roots of a host plant from which it drains nutrients and water to complete its life cycle. Sorghum is one of the host plants that is greatly affected by *Striga* infestation which can result in up to 70-100% yield losses. In **Chapter 1** of this thesis, I discuss the challenges of controlling *Striga* infestation, the different *Striga* resistance mechanisms, with the emphasis on *Low Germination Stimulant* activity (*LGS*) and factors affecting this trait, strigolactones. Strigolactones are the key player in inducing the germination of *Striga* seeds by serving as a signaling molecule for host presence when exuded into the rhizosphere by the host plant. The current knowledge on their evolution, biosynthesis and diversification is extensively discussed. Furthermore, their positive role in the rhizosphere, to induce a symbiotic relationship with *arbuscular mycorrhizal* (AM) fungi, and in regulating plant architecture is addressed.

So far, more than 20 strigolactones have been identified from different plant species. Different blends of strigolactones can be produced by a single plant species. The amount and/or type of strigolactones produced/exuded differs from one plant species to the other. Moreover, the blend can also differ between different cultivars of the same species. In **Chapter 2**, we investigated the correlation of these differences with *Striga* resistance in sorghum. We showed that not the level but the type of strigolactone, in a stereospecific manner, determines the resistance of sorghum lines. Sorghum lines with high *Striga* germination stimulant activity predominantly produce 5-deoxystrigol while the low germination stimulant lines produce orobanchol. Since the purpose of sorghum to exude strigolactones into the rhizosphere is to attract AM-fungi, we looked at how the stereospecific difference of strigolactones affects this symbiotic relationship. I showed that the colonization by three AM fungi species was similar in the high- and low- germination stimulant sorghum lines. Furthermore, we provided evidence for the functional loss of an enzyme annotated as a sulfotransferase (*Sobic.005G213600*, SbSOT4A) and hypothesized it is responsible for the stereospecific difference of strigolactones between low- and high- germination stimulant sorghum lines.

In **Chapter 3**, I further investigated the role of SbSOT4A in the total strigolactone profile of low-and high- germination stimulant lines. We provided evidence on how a sulfotransferase can possibly be involved in strigolactone biosynthesis. We showed that SbSOT4A is localized in the cytosol, suggesting it sulfates small molecules such as hormones. Using protein modeling and substrate docking, we showed the enzyme has good affinity to C18-hydroxycarlactone and proposed a new model on strigolactone biosynthesis in sorghum. In summary, this model proposes that in high germination stimulant lines, SbSOT4A is intact; after sulfation of C18-hydroxycarlactone it is further oxidized at the C19 position to form a carboxy group and upon the loss of the sulfate group ring closure occurs which results in the formation of 5-deoxystrigol. The loss of SbSOT4A function results in the lack of the sulfated intermediate; rather further oxidation of the C18 hydroxyl group results in a carbonyl; upon ring closure orobanchol will be produced which will lead to low germination stimulant activity

towards *Striga*. We further showed that inhibition of sulfotransferases using Triclosan gave a similar phenotype which can be integrated as a tool to control *Striga*.

These findings emphasize the importance of strigolactone diversification. Therefore, in **Chapter 4**, we further studied the production of sorgomol, a strigolactone produced by sorghum that can induce a higher level of *Striga* germination than 5-deoxystrigol. Using Recombinant Inbred Lines (RILs) derived from parents contrasting for the presence of sorgomol, we identified the locus that correlates with sorgomol production in sorghum. With further investigation using RNAseq and bulk segregant analysis, we narrowed down the list of candidate genes and presented evidences on the involvement of two *priori* candidate genes in sorgomol production in sorghum. I proposed the role of *Sobic.008G106200*, which encodes a cytochrome P450, in catalyzing the conversion of 5-deoxystrigol to sorgomol while *Sobic.001G319900* is regulating the level of production.

Strigolactone diversification is achieved by different modifications such as hydroxylation, acetylation, demethylation, esterification, decarboxylation, epoxidation and oxidation. The key players in catalyzing these steps are MAX1, CYP711A homologs that belong to the cytochrome P450 super family of enzymes. Sorghum has four MAX1 homologs and in **Chapter 5**, we characterized their response to phosphate starvation and their expression pattern in different sorghum parts such as root, lower stem, axillary buds and the flower head. Using phylogenetic tree analysis with functionally characterized MAX1 homologs from different plants, I predicted the part of the biosynthesis that are likely to be catalyzed by these MAX1 homologs from sorghum. I also showed their affinity to use carlactone as a substrate using a transient assay in *Nicotiana benthamiana*.

In **Chapter 6**, I discuss the main highlights of the thesis, the challenges and future perspectives. Based on the fact that the success of *Striga* infestation is dependent on the type of strigolactones exuded by sorghum plants, I propose possible tools that can be used to eradicate *Striga*. I also address the concept of integrated *Striga* management with the emphasis on the cultural aspects of the farmers based on my personal observation during a field trip that gave me an opportunity to hear the farmer's side of the *Striga* control measures.

Curriculum vitae

Mahdere Z. Shimels was born on the 12th of March 1985 in Addis Ababa Ethiopia. After completing high school studies at Akaki Adventist school, she joined Jimma University College of Agriculture and Veterinary medicine (JUCAVM) to study Horticulture and obtained her bachelor's degree in 2006. Since 2007 she worked at the Ethiopian Horticulture Producer Exporters' Association (EHPEA) as an on-farm trainer. She provided on-farm trainings and consultancy services to commercial flower farms in Ethiopia to help them meet a good agricultural practice standard compliance and be certified of EHPEA code of practice' (EHPEACoP).

In 2010, Mahdere was awarded a scholarship from the Netherlands Fellowship Program (NFP) to study her MSc. Degree in plant sciences at Wageningen University, the Netherlands. The work she conducted in an MSc thesis on NB-LRR Required for Hypersensitive Response-Associated Cell Death-1 (NRC1) (Laboratory of Plant Pathology and Entomology, supervised by Dr. Daniela J. Sueldo and Dr. Wladimir I.L. Tameling) was vital in her motivation to continue a career in science. As part of the MSc. study, she joined the laboratory of Plant Physiology to conduct a minor thesis supervised by Yanxia Zhang. This was a door that introduced her to the world of *Striga* which created a great motivation to study it further and contribute her part to the efforts to control the damage caused by this parasitic weed. In late 2012, she started as a PhD candidate at the Laboratory of Plant Physiology, Wageningen University to study "The Mechanism Underlying Strigolactone Diversification in Sorghum and Its Role in Resistance Against the Parasitic Weed *Striga hermonthica*" The fruits of this project are presented in this thesis.

Publications

Schlemper R.T, Leite F.A.M., Lucheta R.A., **Shimels** M., Harro J. Bouwmeester., van Veen A.J. and Kuramae E.E. (2017), Rhizobacterial community structure differences among sorghum cultivars in different growth stages and soils, FEMS Microbiology Ecology, 93:8

Gobena D., **Shimels** M., Rich P., Ruyter-Spira C.P., Bouwmeester H., Kanuganti S., Mengiste T. and Ejeta G. (2017), Mutation in sorghum LOW GERMINATION STIMULANT 1(LGS1) alters strigolactones and causes *Striga* resistance, PNAS, 114: 4471-4476.

Sueldo, D. J., **Shimels**, M., Spiridon, L. N., Caldararu, O., Petrescu, A.-J., Joosten, M. H. A. J. and Tameling, W. I. L. (2015), Random mutagenesis of the nucleotide-binding domain of NRC1 (NB-LRR Required for Hypersensitive Response-Associated Cell Death-1), a downstream signaling nucleotide-binding, leucine-rich repeat (NB-LRR) protein, identifies gain-of-function mutations in the nucleotide-binding pocket. New Phytologist, 208: 210-223.

Acknowledgements

From the first moment I saw the beautiful *Striga* flowers framed in Harro's office in May 2010, I learned a lot about *Striga*. My field trip to Ethiopia was truly overwhelming. Seeing *Striga*'s impact on sorghum production changed the way I see the research I do. I am still overwhelmed by the turns and twists it took before we had the current hypothesis in understanding the mechanism behind low germination stimulant activity in sorghum. I was lucky to have the finest people in the field as my supervisors to guide me through the process. I wish to first express my profound thankfulness to my promotor professor Harro Bouwmeester. Dear Harro, thank you for your critical questions, discussion and suggestions during our meetings. Your guidance and support during the entire journey, especially, during the last stretch of the writing was very valuable to me. The other important contributor for the thesis to take this shape was Caroline. Caroline, I enjoyed every meeting we had. Knowing you helped me to grow both in my professional and personal life. Your support was not only bound during my creative and enthusiastic moments but also during days of frustrations and weaknesses. I also owe Sander gratitude for the critical questions he raised during our weekly meeting. Sander, your suggestion helped me to think and see different aspects of every topic raised.

This thesis was done in collaboration with Perdue University, Prof. Gebisa Ejeta's group. Dear Gabisa, you are my inspiration and the person I always look up to. I am honored to get the chance to work with you and the trip to Ethiopia was extra special because you were there. I hope we will find another opportunity to work together. I wish you all the best in life. Daniel, Tesfaye, Patrick, all the brainstorming sessions and discussions during the Skype meetings were amazing. Thank you all for your contributions! Patrick, I always admire how you can tell and also beautifully write our work. I am grateful to get the chance to work with you. I hope we will keep in touch and publish some of the work.

Plant Physiology was filled with many great people. I would like to thank all of you for the nice company and discussions. Rina, you always have the right answers for my questions. I also enjoyed your welcoming and warm personality. Lidiya, you are the most kind and wonderful person. I wish you all the best in your life. Wilco, I am one of many that enjoys your joy. Your office was always open to us and your humbleness shaped me to think of you when I have a critical question on my experimental setups. Tanya, Carmen and Krystina, thank you for your help with the data analysis of the metabolomics work. Krystina, Bea and Giovanni, thank you for your nice friendship! Gonda thank you for being the great person you are. Mariana and Nafiseh all the coffee breaks we had together given me a boost to go back on track. It was great to have friends like you who could listen to my challenges and give me excellent advice. Most importantly, the laughter we shared was uplifting and priceless. Mariana, I am so happy that you are still here to be my paranymph! Arman, Nafiseh, Elise, Natalia, Gonda, the writing room was enjoyable because you guys were there. I wish you all great success in life. The very last stretch of the writing process was barrable because of my colleagues and my friends at NIOO. Dominika, Sewinet, Getahun, Francisco, Raul and Desu thank you for welcoming me to the group with your great personalities. Dominika and Sewinet, we have known each other

only recently but it feels like our friendship lasted longer. Thank you both for the nice chats and your support. Jos, thank you for opening the door for me to join your group and also for giving me the opportunity to work on the subject matter that I truly love and enjoy.

Of course, fun was also part of the PhD journey. ሸንጎ (Shengo) had it all! Without it, these years would have felt even much longer. Yes, if I say ሸንጎ, I should immediately mention you Desu. Your friendship means a lot to me. I am glad that you are still staying in Wageningen with your beautiful family. Abelo, Ruta, Sol, Getch, Duguma it was great to see you all during ሸንጎ after a stressful week. Zenu, Hanani, Betim, Frezer, Abe, Tg, Tade, Sabi, Eske, Mahi, Adam, Beshir, Netsi, Aziz, Mohammed, Meski, Mebre, Tsi, Mikinaye, Mekdi, Fre, Engu, Mekuanint and the rest of Ethiopian community, it felt like home when we organized a get together during holidays and for that, Thank you!

The other beautiful souls I met in Wageningen are Banchiti, Mahi and Kali. It was not only the gym ;-) but also our fun during lunch breaks, the Hema visits and the none stop laughter was refreshing. I am truly thankful for having you in my life. Kali, my dear paranymph, my friend. You are like a sister to me and a great አክስት to my kids. Your great personality is inspiring. You were there when I needed you most and your comforting words kept me going. Thank you for everything and I wish you and Aman all the best in life. Lucie, Neli, Mike, Susu, Mo. Thank you all for your great friendship! Habtish, my first years in Wageningen was so enjoyable because of you and your family. I truly miss you! I am still waiting for the next opportunity to be in one city and continue our great friendship beyond video calls.

All was possible because of unconditional love and support coming from the other part of the world, from my support system, my family. Emeye end Eluye, the two strong women I know, one phone call was enough to make me feel like I can do anything. Fitsume, Nebaye, Bezaye and chapuye, I am who I am today because I grew up with you guys. Every time I hear your voice on the phone, I feel sad that I am not there to support you during your millstones. I am so proud to see how you are handling the challenges in your life and accomplish a lot in life. Most importantly I am proud to be your sister! I wish you all the best in life. Zeku, አባቴ thank you ተማሪ ስላልከኝ። ምክር ግን ብቻ እልነበረም እጄን ይዘህ እዚህ እድርሰኽኛል። አመሰግናለሁ! Living away from you and the rest of my family is hard. I lost count of the times I call you to complain because I am homesick or another challenge in life was there. Every single time, you listened, comforted and advised me as if you never heard of it before. You were always strategic in choosing the right words that keeps me moving. Every phone call with you was a constant reminder of why I am here. Thank you አባ! Roye, እናቴ፣ እኑዬ፣ ሆዴ፣ ያንቺን ጉዳይ ቀሪውን ህይወቴን በሙሉ በየቀኑ ዝርዝር ባወጣ እንኳን ተናግሬ አልጨርሰውም። እንዲያው ባጠቃላይ ለልፋትሽ እና ግሩም ለሆነው ጥንካሬሽ፣ እኛን ብለሽ ላሳለፍሽው ሁሉ ብድርሽን መላሽ ያርገኝ።

Yimegn, Tsedi, Bitu and Bruke, living abroad didn't make it easy to get together that often but the opportunities we had were priceless both for me and for Edna. Yimegnu, I truly admire your strength and I wish you all the best in life. Dear Abiy, thank you for designing and drawing the cover page. Ok, here comes the hardest part but let me try. Abiyu, ዓብዩ የኔ ድንቅ ሰው ምን

ብዬ አመስግኜህ፣ ምን ቃላት ፅፌ ያንተን በጎነት እገልፅ ይሆን? I just have to thank Yimegn for raising a gentle man like you! Abiyo, my husband, my best friend, my Love, every single second of my life journey was possible because I had you next to me. Knowing you is the best thing that happened to me and that was by chance. Marrying you, that is the best decision I have ever made in my life. It is amazing how you make my life so calm. I am fulfilling my ambitions with your great support. At times, it gets tough, but you always make me feel like I am handling it smoothly. You listen, you listen good and also know the right words to say and make me forget it all. You believed in me more than I believed in myself. And for that, Thank you! ሒወቴ ዓብዩ there were days/weeks, evenings and weekends that I had to work and was not possible to spend time together. Guess what, we made it! You share in the accomplishment of this thesis through immeasurable sacrifices. We both accomplished a lot together, but I am most proud the moment you decided to change your carrier to be a programmer. It was not easy but you manage to find the time and make it happen. I am super happy that you are working in the field you love and enjoy. Ednaye and Heranu, my beautiful girls, mommy is always proud of you and love you to the moon and back. ትልቅ ሆናቹህ ምን አይነት አባት እንዳላቸሁ እስከምታውቁ እኔ ለናንተ ጓጓሁ ነው ያለው ፍፃ? እኔስ ብትሉ። እድግ በሉልኝ የኔ እንቁዎች!

The research described in this thesis was financially supported by Bill and Melinda Gates Foundation Grant OPP1006216 and Netherlands Organization for Scientific Research Vici Grant 865.06.002 and Equipment Grant 834.08.001 (to H.B.).

Financial support from the Laboratory of Plant Physiology of Wageningen University, the G. Schieter Foundation and Biomoney for printing this thesis is gratefully acknowledged.

Cover design: Abiy Hailu Sahle

Printing: ProefschriftMaken.nl || DigiForce