2011

Replication protein A and its role in the iron stress response of soybean

Sarah Atwood
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Agronomy and Crop Sciences Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/10362

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Replication protein A and its role in the iron stress response of soybean

by

Sarah Elizabeth Atwood

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Genetics

Program of Study Committee:
Randy C. Shoemaker, Major Professor
Michelle A. Graham
Steven A. Whitham

Iowa State University
Ames, Iowa
2011

Copyright © Sarah Elizabeth Atwood, 2011. All rights reserved.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** iii  
**ABSTRACT** iv  

**CHAPTER 1. GENERAL INTRODUCTION**  
Thesis Organization 1  
Literature Review 2  
References 16  

**CHAPTER 2. REPLICATION PROTEIN A GENE EXPRESSION IN TWO NEAR ISOGENIC LINES (NILS) OF SOYBEAN DURING THE EARLY IRON STRESS RESPONSE** 24  
Abstract 24  
Introduction 25  
Materials and Methods 28  
Results 34  
Discussion 40  
References 45  

**CHAPTER 3. SILENCING REPLICATION PROTEIN A SUBUNIT 3** 60  
Abstract 60  
Introduction 61  
Materials and Methods 64  
Results 70  
Discussion 72  
References 75  

**CHAPTER 4. GENERAL CONCLUSIONS** 87  
Conclusions 87  
Future Research 88  
References 89  

**APPENDIX A. SUPPLEMENTARY TABLES** 90
ACKNOWLEDGEMENTS

I would like to thank my advisor, Randy Shoemaker, for the opportunity to learn and grow in his lab. I thank Dr. Shoemaker, as well as Dr. Steven Whitham and Dr. Michelle Graham, for serving on my committee and guiding me in my research. I would like to thank the members of the Shoemaker, Graham, and Whitham laboratories for their help and support in my research. Finally, I would like to thank my parents, Bruce and Carri Eckhart, my brother Adam Eckhart, and my husband Matt Atwood for their patience and support.
ABSTRACT

Soybean (*Glycine max* (L.) Merr.) is the second most abundant crop in the United States, with a total crop value of nearly $39 billion in 2010. U.S. soybean exports are used as an important protein and oil source in over 80 countries. Abiotic stress, such as iron deficiency chlorosis (IDC), threatens soybean yields, which affect farmers’ profits in the United States as well as nutrition of consumers all over the world.

IDC occurs when iron is unavailable to soybean roots, often in the calcareous soils of the upper Midwest United States. Iron stress causes a decrease in chlorophyll production, resulting in interveinal chlorosis (yellowing between veins), stunting, and yield losses of up to 80%, which cost U.S. farmers over $120 million in 2004. Improving yields benefits farmers as well as those in developing countries, who consume beans as a main source of proteins and nutrients, such as iron. Understanding the genetic basis behind iron efficiency in soybean will aid in breeding programs to decrease economic losses from IDC and improve global nutrition.

A recent microarray study found many DNA replication and repair genes to be differentially expressed during iron stress, but their role has not been studied. This work studies one such gene family, single-stranded DNA binding protein replication protein A (RPA), during iron stress. The first study found RPA genes had opposite expression patterns in two near-isogenic lines (NILs) of soybean, differing only in their iron efficiency, after 24 hours of iron stress. The second study finds silencing of *RPA subunit 3 (RPA3)* improved IDC symptoms in PI 547430 (Isoclark, iron-inefficient) during iron stress. The expression patterns of *RPA* genes during iron stress and
improvement in IDC symptoms upon *RPA3* silencing provide intriguing details to our understanding of the mechanisms of iron efficiency in soybean.
CHAPTER 1. GENERAL INTRODUCTION

Thesis Organization

This following text is divided into four chapters. The first chapter contains a literature review covering iron homeostasis in soybean, the role of replication protein A (RPA) in the cell as well as stress response, and the use of virus-induced gene silencing (VIGS) for functional analyses. The second and third chapters are two manuscripts currently in preparation for submission. The fourth chapter concludes information gained from the two studies and suggests future research.

Chapter 2, “Replication protein A gene expression in two near-isogenic lines (NILs) of soybean during the early iron stress response”, is a manuscript currently being prepared for submission. Replication Protein A homologs were identified in soybean, and gene expression was determined in PI 548533 (Clark) and PI 547430 (Isoclark) at 1, 6, and 24 hours post iron stress. RPA genes were down-regulated in Clark after 24 hours of iron stress, while Isoclark had the opposite pattern. Clark and Isoclark are near-isogenic lines (NILs) that differ only in their iron efficiency. The opposing gene expression patterns suggest RPA may play a role in the iron stress response. Co-authors include Jamie O’Rourke, Gregory Peiffer, Steven Whitham, Michelle Graham, and Randy Shoemaker. Jamie O’Rourke first identified the potential role of replication protein A in the iron stress response in her 2009 microarray study, and assisted in homolog identification. Gregory Peiffer generously provided the tissue for RNA extraction. Michelle Graham provided assistance with homolog identification and phylogenetic analysis. All other experimentation, analysis, and writing was done by
Sarah Atwood under the guidance of Randy Shoemaker. Michelle Graham, Steven Whitham, and Randy Shoemaker were involved in the editing process.

Chapter 3, “Silencing replication protein A subunit 3 suggests a role in the iron stress response”, is a manuscript currently being prepared for submission. A homolog of replication protein A subunit 3 (RPA3), Glyma20g24590, was targeted for silencing using virus-induced gene silencing (VIGS). Silencing of RPA3 in PI 547430 (Isoclark, iron-inefficient) improved iron deficiency chlorosis (IDC) symptoms during iron stress. Co-authors include Jamie O’Rourke, Chunquan Zhang, John Hill, Steven Whitham, Michelle Graham, and Randy Shoemaker. Jamie O’Rourke first identified replication protein A subunit 3 as differentially expressed between Clark and Isoclark under iron stress in her 2009 microarray study. Chunquan Zhang and John Hill generously provided the Bean pod mottle virus (BPMV) VIGS vector for manipulation. Michelle Graham and Steven Whitham provided assistance and guidance in vector creation. All other experimentation, analysis, and writing was done by Sarah Atwood under the guidance of Randy Shoemaker. Michelle Graham, Steven Whitham, and Randy Shoemaker were involved in the editing process.

**Literature Review**

**Iron Deficiency Chlorosis**

Iron is an essential micronutrient required for proper photosynthesis, respiration, and other essential metabolic processes in plants. However, an overabundance of iron is toxic to cells, as free iron can cause reactions that damage
DNA, proteins, and lipids (Winterbourn, 1995). Iron homeostasis is therefore very important for proper growth, and iron uptake, transport, and storage is tightly regulated (Guerinot et al., 1994).

Iron can be found in two forms: ferric (Fe$^{3+}$) and ferrous (Fe$^{2+}$) iron. Legumes, such as soybean, uptake iron in the ferrous state through a root membrane metal transporter known as IRT1 (Vert et al., 2002). Iron has a high redox potential, meaning it can easily accept and donate electrons during chemical reactions of the cell, making it a popular co-factor for enzymes. However, an overabundance of iron can lead to DNA damage via the Fenton reaction, where free radicals are produced by iron’s reaction with oxygen (Guerinot et al., 1994). Iron is chelated with citrate in the xylem and nicotianamine (NA) in the phloem during transport to prevent free iron from producing damaging free radicals (Briat et al., 2007). Iron is stored in the vacuole of the cell, and is found in high abundance in the chloroplast and mitochondria, as it plays essential roles in photosynthesis and respiration (Briat et al., 2007).

Iron deficiency occurs when iron is unavailable to the plant, either by a lack of iron or a lack of iron in the correct (ferrous) form. Iron deficiency chlorosis (IDC) is a problem for soybeans in the upper Midwest where fields may contain alkaline, calcareous soils. Though iron is usually not limiting in the soil, the plants ability to uptake the ferrous form of iron is hindered by various soil properties, such as high moisture content, high pH, and an abundance of soluble salts (Hansen et al., 2003). IDC symptoms include yellow leaves with green veins (interveinal chlorosis), stunting, and
significant yield losses. These symptoms are a result of a decrease in chlorophyll production and photosynthetic rate under iron stress (Spiller et al., 1980; Terry, 1980).

Iron deficiency chlorosis in soybean is scored on a scale from 1 to 5, with 1 being green and healthy leaves and 5 being yellow and necrotic leaves. Yield losses have been estimated to be ~20% per unit, and economic losses are often quite large. A farmer survey in Minnesota found an average chlorosis score of 3.1-4.42 and yield losses averaging 0.8 Mg ha\(^{-1}\) in chlorotic field positions (Hansen et al., 2003). With the reported 52.8 hectares affected by IDC, this likely resulted in an average loss of nearly $12,000 for each farmer surveyed based on soybean bushel prices in 2003 (www.soystats.com). With 79 farmers surveyed, this amounts to nearly one million dollars lost. In 2004, Hansen et al. (2004) estimated the soybean grain production loss from IDC in the United States was over $120 million. It’s clear that economic losses are great, and a solution to iron deficiency is desirable.

**Combating IDC**

Several solutions have been proposed to combat IDC, but so far variety selection appears to be the most effective at preventing devastating yield losses. Treatments with iron chelates, such as soil and foliar sprays, can improve yields (Schenkeveld et al., 2010). However, other studies have found application of iron chelates do not improve yields and interact with some herbicides, preventing proper weed control (Franzen et al., 2003). Planting IDC resistant varieties across entire fields increases yield in areas of the field vulnerable to iron deficiency, but can lead to overall reduction in yield (Helms
et al., 2010). A combined approach of planting IDC resistant varieties in iron deficiency prone areas of the field while planting non-resistant, high-yielding varieties in other locations appears to be the best solution for maximizing yield potential (Helms et al., 2010). In addition to field management strategies, plant geneticists are researching the molecular mechanisms behind iron acquisition in an effort to find genes contributing to IDC resistance.

Under iron stress, soybeans use an iron uptake system known as Strategy I. This strategy is used in non-grasses under iron insufficiency, and several reviews have recently been published (Hell et al., 2003; Jeong et al., 2009; Walker et al., 2008). Upon recognition of insufficient iron availability by the plant, ATPases pump hydrogen ions across the root membrane into the soil, acidifying the rhizosphere and therefore creating a more favorable environment for reducing ferric to ferrous iron. A ferric chelate reductase, FRO2, pumps electrons to the surface of the root membrane and reduces ferric iron to ferrous iron. Iron regulated transporter 1 (IRT1) then transports this ferrous iron across the membrane and into the plant for use.

A microarray study performed on two near-isogenic lines (NILs) under iron insufficient conditions discovered additional genes important in the response to iron stress in soybean (O’Rourke et al., 2009). O’Rourke et al. (2009) found that a number of genes involved in DNA replication and repair were over-represented among those differentially expressed during iron stress in an iron-efficient line, Clark. In addition, a homolog of replication protein A subunit 3 (RPA3) was found to have greater expression in iron-inefficient line Isoclark (PI 547430) than in iron-efficient line Clark (PI 548533).
Interestingly, this homolog of \textit{RPA3} was located within a known iron efficiency QTL on soybean chromosome 20 thought to be contributing to chlorophyll concentration variation in IDC mapping populations (Lin et al., 1997). The results of the microarray study suggest that DNA metabolism is altered by iron stress in soybean, and perhaps RPA plays a role in iron homeostasis.

\textbf{Role of Replication Protein A in Non-plant Species}

Replication Protein A was first identified in human tumor (HeLa) cell extracts. It was determined to be a protein required for the replication of simian virus 40 DNA \textit{in vitro} (Wold et al., 1988). Initial studies identified human RPA (hRPA) to be a single-stranded DNA binding protein (SSB), binding single-stranded DNA and promoting the rapid unwinding of SV40 DNA at the origin of replication (Wold et al., 1988; Kenny et al., 1989). RPA was also found to have a putative protein interaction role, as hRPA stimulated DNA polymerases alpha and delta, resulting in a higher amount of SV40 DNA replication when compared to single-stranded binding proteins from other organisms (Kenny et al., 1989). Since this discovery, the role of RPA in DNA replication has been well studied in humans because of the possibility of targeting RPA in cancer treatment.

RPA is a heterotrimer involved in DNA replication, repair, and recombination. RPA is made up of three tightly associated subunits, at molecular weights 70-kDa (RPA70), 32-kDa (RPA32), and 14-kDa (RPA14) (Wold et al., 1988). Though RPA is known as the eukaryotic SSB, the importance of RPA in DNA metabolism is greater than just its SSB activity. RPA also interacts with other proteins involved in DNA
metabolism, and is often required for correct recruitment or function of these proteins. In organisms with only one copy of RPA, such as yeast, a knock out of any of the three subunits is lethal, indicating its requirement for life (Brill et al., 1991).

The role of RPA in DNA replication is extensive, from DNA unwinding and chain elongation to maintenance of DNA fidelity. During S-phase of the cell cycle, RPA localizes to the replication foci (Iftode et al., 1999; Wold, 1997). Once there, RPA aids in the unwinding of double-stranded DNA by destabilizing the helix, and then binds ssDNA with high affinity (Wold et al., 1988; Kenny et al., 1989). During chain elongation, RPA interacts with DNA polymerase α and regulates primer creation (Collins et al., 1991; Fortune et al., 2006; Kenny et al., 1989). RPA supports DNA fidelity in several ways, including interacting with DNA polymerase λ during replication of damaged DNA, known as translesion synthesis (Krasikova et al., 2008). RPA also prevents the creation of incorrect base pair mismatches by DNA polymerase δ in yeast (Fortune et al., 2006). The post-translational modification of RPA can also stall replication during times of genotoxic or other stresses, perhaps preventing further DNA damage (Wang et al., 1998).

RPA can prevent DNA damage at the replication step, but it also plays a role in DNA repair after damage has occurred. RPA is involved in nucleotide excision repair (NER), which repairs double-helix distorting lesions, such as those induced by UV radiation. NER involves incision of the DNA at the point of damage by endonucleases and removal of damaged bases via exonucleases, followed by gap filling and ligation via DNA synthesis proteins (Kimura et al., 2006). RPA is involved in both the pre-incision
and post-incision steps of NER, and is required for incision, gap filling and DNA synthesis reactions (Overmeer et al., 2011; Li et al., 1995). During NER, RPA interacts with XPA and XPB, proteins essential for UV damage recognition in the pre-initiation complex, as well as PCNA and DNA polymerase δ during DNA synthesis after incision (Li et al., 1995; Overmeer et al., 2011). Studies that inhibit the correct binding of RPA with NER proteins are deficient in the ability to perform NER in UV-irradiated cells, highlighting the importance of RPA in DNA repair (Li et al., 1995). RPA is also important in homologous recombination of double-stranded DNA breaks and during meiosis. After double-stranded DNA breaks, RPA binds single-stranded DNA and interacts with key enzymes such as Rad51, Rad52, and Rad54, which facilitate repair (Binz et al., 2004; Golub et al., 1998; West et al., 2004).

The three subunits of RPA have individual functions that contribute to the overall role of RPA. RPA70, the largest subunit, has been extensively studied because of its DNA binding and protein interaction abilities. RPA70 has multiple functional domains, including the central DNA binding domain and C-terminal domain required for binding the other two subunits, RPA32 and RPA14 (Gomes et al., 1996). The C-terminal domain also contains a zinc-finger binding domain, which has been shown to bind zinc (Bochkareva et al., 2000). This interaction with zinc is required for single-stranded DNA binding and stability of the heterotrimer (Bochkareva et al., 2000). Inactivation of RPA70 is lethal, emphasizing its requirement for proper DNA replication (Brill et al., 1991).
RPA32 is the middle subunit, and plays a role in regulating the heterotrimer as well as interacting with other important enzymes during DNA metabolism. Phosphorylation of RPA32 is thought to be the major component in regulating the activity of RPA during the cell cycle. RPA32 is phosphorylated during late G1 and S phase of mitosis, during DNA synthesis (Din et al., 1990). It is then dephosphorylated late in mitosis, at the G2 or M phase (Din et al., 1990). Phosphorylation also regulates DNA repair activity of RPA during times of cell stress. RPA32 is hyperphosphorylated after genotoxic stress, which prevents its DNA replication activity while leaving its DNA repair activity intact (Binz et al., 2004). Phosphorylation of RPA32 appears to occur only when found in the heterotrimer form and bound to DNA, linking this regulation as part of the DNA metabolism process (Din et al., 1990).

RPA14 is the smallest subunit, but is important for RPA function. RPA14 must bind RPA32 before RPA70 can join the complex, and a knock out of RPA14 in yeast is lethal (Henricksen et al., 1994; Brill et al., 1991). Most studies have found that although RPA14 appears to have a DNA binding OB-fold, it does not bind single-stranded DNA (Bochkarev et al., 1999). In contrast, a recent study found that RPA14 does in fact bind single-stranded DNA (Salas et al., 2009). Its role or function in DNA binding, therefore, is still unknown. RPA14 may have regulatory functions as well, considering RPA14 binds nucleolin in a complex that inhibits DNA replication under heat shock and genotoxic stress (Daniely et al., 2000; Kim et al., 2005).

RPA has been studied most extensively in humans and yeast, but more recently orthologs of RPA have been discovered in plants such as rice and Arabidopsis thaliana.
(van der Knaap et al., 1997; Ishibashi et al., 2001). The first discovery of RPA in plants was in deepwater rice, when gene expression of a putative \textit{RPA1} gene was induced by gibberellin, a growth hormone (van der Knaap et al., 1997). Since then, it has been studied in both rice and \textit{Arabidopsis}, and plant RPAs appear to have many of the same functions as human RPAs.

\textbf{Replication Protein A in Plants}

In higher plants, there are multiple RPA complexes that appear to have different functions. Rice has three homologs of RPA70 (RPA70a, RPA70b, and RPA70c), three homologs of RPA32 (RPA32-1, RPA32-2, and RPA32-3), and one RPA14 (RPA14) (Ishibashi et al., 2006). These homologs create three unique RPA complexes, with only RPA14 in common. The RPA70a—RPA32-2—RPA14 complex localizes to the chloroplast, while the RPA70b—RPA32-1—RPA14 and RPA70c—RPA32-3—RPA14 complexes localize to the nucleus (Ishibashi et al., 2006).

The mRNA expression of RPA subunits changes during cellular processes of plants in which RPA plays a pivotal role. During the cell cycle, genes for all three RPA subunits are up-regulated at the same time, slightly before S phase, and expression stays high through replication before dropping back to a basal level (Marwedel et al., 2003). Protein levels are relatively stable during the cell cycle, perhaps suggesting that there is significant turnover of protein product during the replication part of the cell cycle (Marwedel et al., 2003). RPA mRNA expression is closely tied to replication, as expression is high in proliferating tissues of plants, and low in mature, non-replicating
tissues (Ishibashi et al., 2001; Chang et al., 2009). RPA mRNA expression is also up-regulated after DNA damage from chemical mutagens (Takashi et al., 2009). Based on the microarray study by O’Rourke et al. (2009), RPA3 mRNA expression also changes under iron stress, though its role in iron homeostasis, if there is one, is still unclear.

Arabidopsis contains the same three homologs of RPA70 as rice. Knock out and RNAi studies of genes for RPA70, known as RPA1, have alluded to the function of each unique complex in rice and Arabidopsis thaliana. A T-DNA insertion mutant of RPA1-70a (AtRPA70a) in Arabidopsis thaliana was lethal, as well as an RNAi line, suggesting an essential role in DNA replication (Ishibashi et al., 2005). However, a separate study found the knock out of RPA1-70a in Arabidopsis thaliana to be viable, with only an increased sensitivity to DNA mutagens and increased telomere length (Takashi et al., 2009). This discrepancy may be indicative of the complexity of the RPA subunits’ functions in the cell. A T-DNA insertion mutant of RPA1-70b (AtRPA70b) as well as an anti-RPA1-70b RNAi line in Arabidopsis thaliana was also viable and showed greater sensitivity to DNA damaging agents, suggesting a role in DNA repair (Ishibashi et al., 2005). Knock out of RPA1-70c (AtRPA70c) in Arabidopsis thaliana is lethal (Ishibashi et al., 2006). Based on these studies, it is believed the RPA70a and RPA70c complexes are involved in essential processes such as replication, while the RPA70b complex is involved in repair (Sakaguchi et al., 2009).

No functional studies have been done on any subunit of RPA in soybean. Though transformation is possible in soybean, efficiencies can be as low as 3% with a wait time of nine months or more for transformed seed (Paz et al., 2006). A new strategy using
virus-induced gene silencing (VIGS) allows for more efficient and high-throughput functional characterization in soybean (Zhang et al., 2006).

**Virus-Induced Gene Silencing**

As more plant genomes are sequenced and annotated, reverse genetics becomes a useful approach in characterizing gene function. Reverse genetics involves manipulating known DNA sequences in vivo (e.g., transgenes or gene knock out) and observing resultant phenotypes in order to determine gene function. Virus-induced gene silencing (VIGS) has been developed as a rapid, high-throughput method to study gene function in soybean (Zhang et al., 2006).

VIGS uses the plant’s natural defense against RNA-based viruses to silence an endogenous gene. Aberrant double-stranded RNA (dsRNA) molecules produced during viral replication trigger Post-Transcriptional Gene Silencing (PTGS) of viral genes (Zhang et al., 2006). PTGS was first discovered in petunia, when a transgene encoding the chalcone synthase gene (CHS) suppressed endogenous CHS gene expression, resulting in white flowers instead of the expected deep purple (Napoli et al., 1990). The same PTGS pathway that silences viral genes after infection is used to silence a targeted host gene in VIGS.

After virus infection, plants defend themselves by degrading viral RNA before it produces a protein, via PTGS (Brodersen et al., 2006; Lu, 2003). During PTGS, aberrant dsRNA is cut to lengths of 21-24 nucleotides by the RNaseIII enzyme Dicer (Bernstein et al., 2001; Hamilton et al., 1999). These short oligonucleotides are known as primary
small-interfering RNAs (siRNAs) because of their use in interfering with gene expression. Once created, siRNAs are loaded into the RNA-induced silencing complex (RISC), which includes an RNase, Argonaute (Hammond, 2005). Argonaute cleaves any mRNA molecules that pair perfectly with the bound siRNA, preventing translation and effectively silencing the gene (Hammond, 2005). mRNA molecules can still bind the siRNA-RISC complex without perfect base-pairing, but translation is inhibited by RISC competition for translation machinery rather than mRNA cleavage (Hammond, 2005). RNA-Dependent RNA Polymerase (RDRP) creates secondary siRNA molecules by amplifying primary siRNA molecules from virus dsRNA cleavage (Dalmay et al., 2000). Secondary siRNA molecules can target non-viral mRNA molecules with sequence homology, and they can act as a silencing signal in nearby cells to prevent the spread of virus (Peele et al., 2001). It is with this system that a plant can quickly recognize and destroy foreign virus RNA as it replicates, preventing or slowing infection.

VIGS takes advantage of this system by manipulating virus RNA to contain a small subsection of target host gene sequence from the host’s genome (Lu, 2003; Voinnet, 2001). As the virus replicates, it creates dsRNA from its entire genome, including the foreign host sequence (Purkayastha et al., 2009). As Dicer creates siRNAs, some will contain sequence that has perfect homology to the targeted host gene (Purkayastha et al., 2009). Host gene silencing is achieved when endogenous mRNA is cleaved, or translation inhibited by RISC (Purkayastha et al., 2009).

The first use of VIGS in plants silenced *phytoene desaturase (PDS)*, an enzyme important for the biosynthesis of carotenoids that protect chlorophyll from
photobleaching, in *Nicotiana benthamiana* (Kumagai et al., 1995). Since then, VIGS has been used in a wide variety of plant species, such as tomato, *Arabidopsis thaliana*, tobacco, barley, and pea (Purkayastha et al., 2009). The VIGS system is ideal for many plant species because it does not require transformation, vectors are simple to create, and results can be achieved in as little as 3 weeks post inoculation (Burch-Smith et al., 2004). Given that soybean has a duplicated genome, another advantage of VIGS is the ability to silence multiple genes with sequence homology, circumventing the hindrance of functional redundancy in knock out studies (Lawrence et al., 2003; Meyer et al., 2009). A significant drawback of VIGS is the likelihood that viral symptoms may mask mild phenotypes after gene silencing (Burch-Smith et al., 2004).

A successful VIGS system for soybean has been developed using *Bean pod mottle virus* (BPMV) (Zhang et al., 2006). BPMV is an RNA virus, which contains two RNA molecules, RNA1 and RNA2. RNA1 codes for five genes required for virus replication, while RNA2 codes for the cell-to-cell movement protein and coat proteins (Zhang et al., 2006). The first version of the BPMV VIGS vector inserted foreign sequence between the movement protein and coat protein genes of the RNA2 molecule, successfully silencing endogenous genes such as *PDS* (Zhang et al., 2006). However, the first version was RNA-based and did not easily translate to high-throughput functional studies (Zhang et al., 2009).

A DNA based BPMV VIGS vector was introduced in 2009, and the newest versions insert target gene sequences after the translation stop codon in RNA2 (Zhang et al., 2010). The insertion of target sequences after the stop codon allows the use of
non-coding and antisense sequences that are shown to be most effective for silencing of endogenous genes (Zhang et al., 2009; 2010). This version of BPMV vector for soybean has been successfully used to identify genes involved in Rpp2 and Rpp4 mediated resistance against Asian soybean rust (Meyer et al., 2009; Pandey et al., 2011). Though the BPMV VIGS system has been proven in identifying defense related genes, genes involved in DNA metabolism or abiotic stress have not yet been studied.

**VIGS and RPA**

VIGS is an ideal system to study RPA in soybean. Previous knock out studies in *Arabidopsis thaliana* and yeast have shown that eliminating RPA subunit gene expression can be lethal (Brill et al., 1991; Ishibashi et al., 2005). However, silencing with VIGS maintains gene expression in some tissues, allowing for functional characterization of genes that would be lethal when knocked out. The ability to silence multiple genes with VIGS is also ideal for the duplicated soybean genome (Shoemaker et al., 1996; Schmutz et al., 2010). With the BPMV-VIGS system, the functional characterization of RPA subunits in soybean is promising.

**Role of RPA in Iron Homeostasis**

Iron deficiency chlorosis in soybeans is a complicated problem that is yet to be fully understood. Many studies have focused on differential gene expression under iron stress to explain the causes of IDC, and genes important in the uptake, storage, and transport of iron have been identified (Buckhout et al., 2009; Colangelo et al., 2004;
DNA metabolism genes have been implicated in the iron stress response, though the role of these genes is unclear (O’Rourke et al., 2009). Replication protein A plays an essential role in DNA replication, repair, and recombination, and was implicated as differentially expressed between iron efficient and iron inefficient genotypes in the O’Rourke et al. (2009) study.

The purpose of this study is to determine if replication protein A has a role in iron homeostasis. The first step is to determine if gene expression patterns of all three RPA subunits differ under iron insufficiency between Clark and Isoclark, two near-isogenic lines (NILs) of soybean varying in their iron efficiency. The second step is to use VIGS to silence expression of RPA3, a gene already found to respond differently to iron insufficiency in these same lines (O’Rourke et al., 2009). We will evaluate, in the context of a possible role for RPA in iron homeostasis, gene expression patterns and phenotypic results from silencing.

References


Proceedings of the National Academy of Sciences of the United States of America 92, 1679-83.


CHAPTER 2. Replication protein A gene expression in two near-isogenic lines (NILs) of soybean during the early iron stress response

Atwood, Sarah E.\textsuperscript{1}, O’Rourke, Jamie A.\textsuperscript{2}, Peiffer, Gregory A.\textsuperscript{1}, Whitham, Steven A.\textsuperscript{4}, Graham, Michelle A.\textsuperscript{3}\textsuperscript{*}, Shoemaker, Randy C.\textsuperscript{3}\textsuperscript{*}

* Corresponding author; e-mails: michelle.graham@ars.usda.gov, randy.shoemaker@ars.usda.gov

\textsuperscript{1} Interdepartmental Genetics Program, Iowa State University, Ames, Iowa 50011

\textsuperscript{2} USDA-Agricultural Research Service, Plant Science Research Unit, Saint Paul, Minnesota 55108

\textsuperscript{3} USDA-Agricultural Research Service, Corn Insects and Crop Genetics Research Unit, Ames, Iowa 50011

\textsuperscript{4} Department of Plant Pathology, Iowa State University, Ames, Iowa 50011

Abstract

Iron is a micronutrient required for proper growth and development of plants. Iron deficiency results in interveinal chlorosis and an overall decrease in photosynthetic capacity, leading to stunting and yield losses. The effect of iron stress on iron uptake and transport genes is well documented in plants. However, the effect of iron stress on DNA replication and repair genes is not well known. Replication protein A (RPA) is a heterotrimeric protein essential for DNA replication, repair, and recombination. In this study, the gene expression of replication protein A subunits RPA1, RPA2, and RPA3 were studied at 1, 6, and 24 hours post iron stress in two near-isogenic lines (NILs) of soybean, each varying in their iron efficiency. Overall, RPA
genes were found to be down-regulated after 24 hours of iron stress in iron-efficient line Clark, while gene expression was up-regulated in iron-inefficient line Isoclark at the same time point. The opposite expression patterns in Clark and Isoclark, which differ only in iron homeostasis efficiency, suggests RPA proteins may play a role in iron homeostasis in soybean.

**Introduction**

Iron is an essential micronutrient required for proper photosynthesis, respiration, and other essential metabolic processes in plants. However, an overabundance of iron is toxic to cells, as free iron can cause reactions that damage DNA, proteins, and lipids (Winterbourn, 1995). Iron homeostasis is therefore very important for proper growth, and iron uptake, transport, and storage is tightly regulated (Guerinot et al., 1994).

Iron content in soybean has both nutritional and agricultural importance. It is estimated that nearly 25% of the global population is anemic, with the highest percentage in pregnant women and young children in developing countries (de Benoist et al., 2008). Biofortification of crops is considered to be the best solution for solving iron deficiency in the developing world, where diets are mainly plant-based (Mayer et al., 2008). Agriculturally, yield losses from iron deficiency in soybean can be quite large. In 2004, the estimated loss from iron deficiency chlorosis (IDC) of soybeans in the United States was over $120 million (Hansen et al., 2004). Furthering our knowledge of
the uptake, transport, or regulation of iron in plants is essential to improving both human nutrition and preventing detrimental yield losses for farmers.

Iron deficiency occurs in plants when iron is unavailable, either by a lack of iron or a lack of iron in the ferrous (Fe$^{2+}$) form. Iron deficiency chlorosis (IDC) is a problem for soybeans in the upper Midwest where fields may have alkaline, calcareous soils. Though iron is usually abundant in soil, the plants' ability to uptake the ferrous form of iron is hindered by various soil properties, such as high moisture content, high pH, and an abundance of soluble salts (Hansen et al., 2003). Iron stress decreases chlorophyll production and photosynthetic rate, leading to yellow leaves with green veins (interveinal chlorosis) (Terry, 1980; Spiller et al., 1980). Other symptoms of IDC include stunting and significant yield losses.

In addition to photosynthesis, iron depletion affects the DNA replication and repair machinery in soybean. A microarray study performed under iron stress conditions found that differentially expressed genes involved in DNA replication and repair were over-represented in iron-efficient line Clark (O'Rourke et al., 2009). A probe corresponding to a replication factor (GmaAffx.36066.1.S1_at) was also found to be differentially expressed between the two NILs during iron stress (O'Rourke et al., 2009). Upon further investigation, it was found that GmaAffx36066.1.S1_at is representative of gene expression for a homolog of *replication protein A subunit 3* (*RPA3*). In addition to differential expression between the two NILs, this homolog of *RPA3* was located within a known iron efficiency QTL on soybean chromosome 20 thought to be involved in chlorophyll content variation (Lin et al., 1997). In this study,
the expression of all RPA genes is investigated in an effort to further understand the response of RPA during iron stress.

Replication protein A (RPA) is the eukaryotic single-stranded binding protein, involved in both DNA replication and repair (Wold, 1997). RPA is a heterotrimeric protein made of three subunits: RPA70 (70 kDa), RPA32 (32 kDa), and RPA14 (14 kDa). The genes RPA1, RPA2, and RPA3 code for RPA70, RPA32, and RPA14, respectively. RPA gene expression is closely tied to replication, as expression is high in proliferating tissues of plants, and low in mature, non-replicating tissues (Ishibashi et al., 2001; Chang et al., 2009). RPA gene expression is also up-regulated after DNA damage from chemical mutagens (Takashi et al., 2009). The study by O'Rourke et al. (2009) was the first connection made between RPA gene expression and abiotic stress in plants.

In this study, we identified the homologs of RPA genes in Glycine max (L.) Merr. and determined their gene expression during the early stages of the iron stress response. Gene expression of all homologs of RPA1, RPA2, and RPA3 was assessed in Clark (iron-efficient) and Isoclark (iron-inefficient) at 1, 6, and 24 hours post iron stress. Homologs for all three genes were down-regulated in Clark (iron-efficient) at 24 hours post iron stress. Isoclark, however, had the opposite gene expression pattern, with an up-regulation of homologs at the same time point. Though the exact role of RPA in iron homeostasis remains unknown, it is clear RPA responds to iron stress differently in iron-efficient verses iron-inefficient lines.
Materials and Methods

Homolog Identification

All the RPA genes have been identified in *Arabidopsis thaliana* (*RPA1*: At2g06510, At5g08020, At5g45400, At5g61000, At4g19130, *RPA2*: At2g24490, At3g02920, and *RPA3*: At2g24490, At3g02920) (Shultz et al., 2007). *Arabidopsis* RPA coding sequences were aligned using ClustalW (Thompson et al., 1994) and HMMER (Durbin et al., 1998) was used to build a hidden Markov model (HMM) for each RPA subunit. The HMM was then searched against all predicted coding sequences in the soybean genome (www.phytozome.net), which were translated into all six reading frames. Any soybean gene above the default e-value cut-off (E-value = -1) was considered in our analysis. This comparison identified nine homologs of the *RPA1* gene, five homologs of the *RPA2* gene, and four homologs of the *RPA3* gene in soybean.

For our study, each homolog was given a name corresponding to the *Arabidopsis* homolog for which it had the greatest homology (Table 1). *GmRPA1A*, *GmRPA1B*, *GmRPA2*, and *GmRPA3* names correspond to genes matching *AtRPA1A* (*AtRPA70a*), *AtRPA1B* (*AtRPA70b*), *AtRPA2* and *AtRPA3*.

Phylogenetic Analysis

Amino acid sequences for RPA homologs in the species *Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*, *Ricinus communis*, and *Glycine max* were assessed for sequence conservation. Peptide sequences were obtained for RPA homologs already identified in *Arabidopsis thaliana* and *Oryza sativa* (Shultz et al., 2007). Peptide
sequences corresponding to RPA proteins in *Arabidopsis thaliana* were BLASTed against the proteomes of *Medicago truncatula* and *Ricinus communis*, and those with protein homology $>50\%$ were used in the alignment analyses. Peptide sequences were aligned with Pileup in the Accelrys GCG software (Accelrys Inc., San Diego, CA). The sequence alignment was visually inspected and trimmed to eliminate gaps and unconserved regions. Sequence alignments for all three RPA subunits were visualized with Multiple Align Show (http://www.bioinformatics.org/SMS/multi_align.html).

The unrooted phylogenetic tree for the RPA1 subunit was created in MEGA5 (Tamura et al., 2011) from trimmed amino acid sequence alignments. The evolutionary history was inferred using the neighbor-joining method (Saitou et al., 1987). A percentage of replicate trees in which the associated taxa clustered together was calculated from a bootstrap test (500 replicates) (Felsenstein et al., 1985). Evolutionary distances were computed using the p-distance method (Nei et al., 2000) and are in the units of the number of amino acid differences per site.

**Germplasm**

Clark (PI 548533), Isoclark (PI 547430) and T203 (PI 54619) were used to study the role of RPA in iron homeostasis. Clark is iron-efficient, while Isoclark and T203 are iron-inefficient. Isoclark is a near-isogenic line (NIL) of Clark, containing an introgressed region from parent T203 that is hypothesized to cause iron inefficiency.
Plant Growth Conditions

Clark, Isoclark, and T203 seed were germinated for 5 to 7 days in a growth chamber at 27°C until unifoliates had emerged but were not fully expanded. After removing from germination paper, seedlings were placed in iron sufficient hydroponic conditions (100 µM Fe(NO$_3$)$_3$•9H$_2$O) in the greenhouse. When the first trifoliate was fully expanded (13 days after placing in hydroponics), plant roots from each bucket were rinsed 6 times in fresh double distilled water, each for 15 seconds minimum and returned to a new hydroponic bucket. Six buckets were returned to iron sufficient conditions (100 µM Fe(NO$_3$)$_3$•9H$_2$O), while the other six were placed in iron insufficient conditions (50 µM Fe(NO$_3$)$_3$•9H$_2$O). Nutrient solutions were based on growth conditions described in Chaney et al. (1992), with solutions adjusted for 10 L buckets.

Tissue from two Clark and two Isoclark plants was pooled from each bucket at the time points 1 hour, 6 hours, and 24 hours after introduction into new iron conditions, for a total of six biological replicates at each time point and iron condition. First trifoliates were harvested and flash frozen in liquid nitrogen and stored at -80°C for later RNA extraction.

Primer Design and Quality Assessment

Reverse transcription polymerase chain reaction (RT-PCR) primers were designed for all homologues using the program Primer 3 (Rozen et al., 2000). Primers were designed using Primer 3 defaults, except for amplicon size (125-175 bp). Primers were designed based on coding sequences (CDS) of predicted genes
RPA CDS sequences were compared using BLASTN (Altschul et al., 1990), and only unique sequences were used in order to distinguish between homeologs located in duplicated genomic regions.

Primers were tested on Clark and Isoclark total RNA harvested from an iron-insufficient bucket at 14 days post iron stress. Primers were amplified using the RT-PCR kit from Stratagene with the same cycle conditions as later qRT-PCR experiments (described below). Reactions were then run on poly-acrylamide gels and were chosen for future qRT-PCR if a single band and few primer-dimers were found. Primers that had no amplification or had multiple bands were not used for later qRT-PCR studies, and were subsequently redesigned and retested. If redesigned primers did not amplify, the gene of interest was not studied in later qRT-PCR experiments. Seven of the original 18 homologs were found to not amplify, and this coincided with low expression levels in a separate RNA-seq study of Clark (unpublished data) as well as low expression levels in an RNA-seq study of Williams82 (Severin et al., 2010). Final primer designs used for qRT-PCR are found in Table 2.

RNA Isolation and Quality Assessment

Flash frozen leaf tissue from the 1st trifoliate was ground in liquid nitrogen and RNA was extracted using a Qiagen® RNeasy® Plant Mini Kit. This protocol was used with the following specifications or changes: ~200 mg tissue was lysed with RLT buffer, tubes were incubated at 56°C Celsius for 2 minutes with 800 rpm shaking to aid in tissue disruption, and columns were incubated at room temperature for 10 minutes during
elution. RNA was then DNased with an Ambion\textsuperscript{®} TURBO DNA-free\textsuperscript{™} kit to remove all DNA.

After isolation, RNA was assessed for quality using a Thermo Fisher Scientific\textsuperscript{®} NanoDrop\textsuperscript{™} ND-1000 Spectrophotometer. RNA was considered to be of good quality for qRT-PCR if the 260/280 ratio was above 2.0 and the 260/230 ratio was above 1.7.

RNA for the short-term study was also analyzed for quality using an Agilent\textsuperscript{®} 2100 Bioanalyzer\textsuperscript{™}. RNA was considered to be of good quality if the RNA was not degraded or was only marginally degraded. This assessment was based upon the amount of degradation indicated by the spectrum.

Reverse Transcription and qRT-PCR

RNA templates were amplified using the Invitrogen\textsuperscript{™} SuperScript\textsuperscript{™} III Platinum\textsuperscript{®} SYBR\textsuperscript{®} Green One-Step qRT-PCR Kit. Reactions were carried out according to the SYBR\textsuperscript{®} Green protocol with the following specifications: total starting RNA was 100 ng, and reactions were a final volume of 25 µl instead of 50 µl. RNA was diluted to 9.52 ng/µl for greater pipetting accuracy. All experiments included a standard curve of 600, 400, 100, 50 and 10 ng concentration as well as No Reverse-Transcriptase (NRT) and No Template Control (NTC) wells for each primer. NRT wells replaced Superscript III with Invitrogen\textsuperscript{®} Platinum\textsuperscript{®} Taq DNA polymerase at the same volume. NTC wells replaced RNA with double distilled, nuclease-free H\textsubscript{2}O at the same volume.

qRT-PCR was carried out on a Stratagene Mx3000P\textsuperscript{™} Real-Time PCR System. After amplification, a dissociation reaction was performed for later analysis of reaction
quality. Reaction conditions were from the Invitrogen\textsuperscript{TM} SuperScript\textsuperscript{TM} III Platinum\textsuperscript{®} SYBR\textsuperscript{®} Green One-Step qRT-PCR Kit with the following modifications: cDNA synthesis was carried out at 60°C, initial denaturing time was 10 minutes, and a total of 45 cycles were carried out with an additional extension time of 15 seconds at 72°C. A measurement of fluorescence was taken after each cycle. The default SYBR\textsuperscript{®} Green dissociation reaction conditions were used from the Stratagene Mx3000P\textsuperscript{TM} Real-Time PCR software.

Each 96 well plate contained either the Clark or Isoclark genotype at all three time points and both iron conditions. Three biological replicates were chosen at random (see “Early Response Study” above) for qRT-PCR analysis in order to maintain all time points and iron conditions on one plate, allowing for direct comparison of expression level (Rieu et al., 2009). Each plate amplified one gene-of-interest primer and one reference gene primer, CYP2 (cyclophilin 2). CYP2 was chosen as a reference gene based upon previous qRT-PCR reference gene studies in plants as well as an in-house study (Phillips et al., 2009; Wang et al., 2011).

Data Analysis

Reactions were considered of good quality for further data analysis if the NRT and NTC cycle thresholds (Cts) were greater than 5 cycles away from the lowest data point and the dissociation curve showed only one peak for reaction wells.

Each sample was amplified with gene-of-interest primers and reference gene primers on the same plate, allowing for direct comparison of quantity values (Rieu et
al., 2009). RNA quantities were determined for each well by aligning to the standard curve for that primer set. A normalized value for RNA quantity was calculated as a ratio of gene-of-interest RNA quantity to reference gene RNA quantity for each sample. Averages of the normalized data were then calculated over technical replicates. Relative expression is a ratio of normalized values in insufficient conditions over normalized values in sufficient conditions at each time point. This ratio is then log base 2 transformed. Log-transformed data was analyzed for standard deviation and standard error.

Differences in relative quantity were analyzed with an Analysis of Variance test (ANOVA) (Chambers et al., 1992) and then Tukey's Honestly Significant Difference test (HSD) (Yandell, 1997) for pair-wise comparisons, with a significance cut-off of 0.05.

Results

Homolog Identification and Phylogenetic Analysis

Using our search criteria, homologs of RPA1, RPA2, and RPA3 were identified in Glycine max (L.) Merr. using known RPA genes in Arabidopsis thaliana. Soybean was found to have seven homologs of RPA1A, two homologs of RPA1B, five homologs of RPA2, and four homologs of RPA3 (Table 1).

In order to determine which soybean homologs were most closely related to the characterized genes in rice and Arabidopsis, multiple sequence alignment and evolutionary history was inferred from RPA amino acid sequences in several plant species. Multiple sequence alignment based on amino acid similarity between RPA
homologs in *Arabidopsis thaliana, Oryza sativa, Medicago truncatula, Ricinus communis,* and *Glycine max* were created with Pileup in GCG (Accelrys Inc., San Diego, CA). Based on the protein alignment, a phylogenetic tree was created for the RPA1 homologs in MEGA5 (Tamura et al., 2011) to determine which of the multiple versions of RPA1 are most related to the homologs in *Glycine max* (Figure 1). Of the nine *GmRPA1* genes identified, three were not included in the alignment analysis: *GmRPA1Ab, GmRPA1Ad,* and *GmRPA1Ae* (Table 1). These three genes were low confidence predictions, and found to be truncated versions of full-length RPA1 genes.

Three RPA1 genes have been characterized in *Arabidopsis thaliana* and *Oryza sativa* (rice): *RPA1A, RPA1B,* and *RPA1C.* RPA1A and RPA1C are thought to be involved in DNA replication in the chloroplast and nucleus, respectively, while RPA1B may be involved in DNA repair in the nucleus (Ishibashi et al., 2005; 2006). Phylogenetic analysis of the multiple sequence alignment (Figure 2) identified two related soybean proteins for each of the three characterized RPA1 proteins in rice and *Arabidopsis* (Figure 1).

RPA1A and RPA1C are most alike in protein sequence, according to the phylogenetic analysis and sequence alignments (Figures 1 and 2). RPA1B, although similar in the oligonucleotide-binding fold (OB-fold, red box, Figure 2) to RPA1A and RPA1C, appears to have a more unique C-terminal region (yellow box, Figure 2).

Figures 3 and 4 show the multiple amino acid sequence alignments of subunits RPA2 and RPA3, respectively. It is clear the oligonucleotide-binding fold (OB-fold) is highly conserved among homologs in all species for both subunits. Glyma08g18770,
named \textit{GmRPA2a} for this study, has the most divergent protein sequence among the RPA2 homologs analyzed. The expression of \textit{GmRPA2a} also differs from the others during iron stress in soybean (see below).

Primer Design and Testing

Primers were designed in Primer3 (Rozen et al., 2000) to amplify one individual homolog. Primers were used for qRT-PCR analysis if they produced a single band of the predicted size. Out of 18 RPA homologs tested, 11 met this criterion and were used for further gene expression analysis (Table 2). \textit{GmRPA1Ab}, \textit{GmRPA1Ad}, and \textit{GmRPA1Ae} are low-confidence gene predictions (www.phytozome.net). These genes may be pseudogenes or may not be expressed under our experimental conditions. These were the same genes found to be truncated and excluded from phylogenetic analysis. \textit{GmRPA1Ac}, \textit{GmRPA1Ag}, \textit{GmRPA1Ba}, and \textit{GmRPA3a} are high-confidence gene predictions, but did not express under our conditions. These genes were also found to have little or no expression in an expression analysis of soybean tissues (Severin et al., 2010). Interestingly, the homologs with the highest homology to \textit{AtRPA1} genes (\textit{GmRPA1Aa}, \textit{GmRPA1Af}, and \textit{GmRPA1Bb}) were the only \textit{GmRPA1} homologs that express during iron stress.

After primer testing, two homologs of \textit{RPA1A} (\textit{GmRPA1Aa} and \textit{GmRPA1Af}), one homolog of \textit{RPA1B} (\textit{GmRPA1Bb}), five homologs of \textit{RPA2} (\textit{GmRPA2a}, \textit{GmRPA2b}, \textit{GmRPA2c}, \textit{GmRPA2d}, and \textit{GmRPA2e}), and three homologs of \textit{RPA3} (\textit{GmRPA3b}, \textit{GmRPA3c}, \textit{GmRPA3d}).
and GmRPA3d) were analyzed by qRT-PCR. Final primer designs used for qRT-PCR are found in Table 2.

RPA Expression Analysis During Iron Stress

GmRPA3c (Glyma20g24590) was previously identified as differentially expressed after 10 days of iron stress (O’Rourke et al., 2009). Iron homeostasis related genes, such as FRO and IRT, can change expression levels in as little as 6 hours after iron insufficiency (Buckhout et al., 2009). In order to determine if other RPA subunits were responding to iron stress earlier, we looked at the expression level of RPA homologs in soybean 1st trifoliate leaf tissue after 1, 6, and 24 hours of iron insufficiency (50 μM Fe(NO₃)₃•9H₂O) and compared their expression to that of non-stressed 1st trifoliate leaf tissue.

Relative gene expression was determined with quantitative reverse transcription polymerase chain reaction (qRT-PCR). Relative gene expression value is a ratio of RNA quantity in iron insufficient conditions to the RNA quantity at the same time point in iron sufficient conditions, averaged over three biological replicates, then log transformed (base 2). The log transformation creates a more normal distribution of data, allowing for more accurate statistical analysis (Ramsey et al., 2002).

Analysis of Variance (ANOVA) (Chambers et al., 1992) was performed on normalized gene expression values in iron sufficient conditions at 1, 6, and 24 hours post stress (hps) to determine if RPA gene expression is stable over the time points in iron sufficient conditions. Normalized gene expression values in iron sufficient
conditions were also log transformed (base 2) to ensure a normal distribution. Genes were considered to be stable if the ANOVA analysis was insignificant, indicating no change in expression among time points.

The expression of all 11 RPA genes did not change significantly over time in Isoclark while 8 of the 11 RPA genes did not change significantly in Clark under iron sufficient conditions (data not shown). *GmRPA1Bb* and *GmRPA2c* had significant p-values (<0.05) in Clark, suggesting gene expression was not stable across time points for these two genes in iron sufficient conditions. However, pair-wise comparisons between time points found a significant change in expression between only 6 hps and 24 hps for *GmRPA1Bb* and *GmRPA2c* in Clark. This result suggests if these genes are truly changing under iron sufficient conditions, it is for a short period around 6 hps before returning back to the same basal level. Where gene expression values are stable across time points in iron sufficient conditions, any change in relative gene expression is dependent on iron insufficiency.

Figure 5 shows the change in relative expression level of RPA genes over time in Clark leaves under iron stress. To determine how relative expression changes over time, the relative expression at 6 and 24 hps was compared to relative expression at 1 hps. Values above zero indicate greater expression in iron insufficient conditions, while values below zero indicate lesser expression in iron insufficient conditions. Figure 6 shows the same comparisons in Isoclark over time under iron stress. Statistical significance data can be found in supplementary tables S1 and S2.
RPA expression decreases in iron-efficient Clark under iron stress (Figure 5). Of the three GmRPA1 genes, GmRPA1Bb is down-regulated at 24 hps while GmRPA1Af and GmRPA1Aa remain unchanged under iron stress. Four of the five GmRPA2 genes are down-regulated, while the fifth, GmRPA2a, is up-regulated at 24 hps. Two of the three GmRPA3 genes are down-regulated at 24 hps, while GmRPA3b remains unchanged. In all cases, the change in expression from 1 hps to 6 hps is insignificant, suggesting changes in gene expression occur between 6 hps and 24 hps in Clark.

Interestingly, RPA genes have the opposite expression pattern in iron-inefficient Isoclark under iron stress (Figure 6). In Isoclark, GmRPA1Af and GmRPA1Bb are up-regulated at 24 hps, while GmRPA1Aa again remains unchanged. The opposite expression pattern happens again in the GmRPA2 genes, except for GmRPA2e, which remains unchanged. For subunit RPA3, the same genes that were down-regulated by Clark are up-regulated at 24 hps in Isoclark (GmRPA3c and GmRPA3d). GmRPA3b is down-regulated slightly in Isoclark. In all cases, the change in expression from 1 hps to 6 hps is insignificant, suggesting response to iron occurs between 6 hps and 24 hps in Isoclark as well.

For the RPA1 subunits, GmRPA1Bb appears to be responding most drastically to iron stress in both Clark and Isoclark, with fold changes of -18 and +11, respectively, between 1 and 24 hps. Though most of the GmRPA2 genes respond drastically to iron stress, GmRPA2b and GmRPA2c have the greatest change in the RPA2 family, with fold changes of -104 and -25, respectively, in Clark, and +13 for both genes in Isoclark. GmRPA3c and GmRPA3d also respond to iron stress in both lines, with fold changes of
-28 and -50, respectively, in Clark. Isoclark fold changes were +7 and +18 for GmRPA3c and GmRPA3d, respectively.

**Discussion**

Replication protein A is the eukaryotic single-stranded DNA binding protein consisting of three subunits: RPA1, RPA2, and RPA3 (Iftode et al., 1999). RPA binds and stabilizes single-stranded DNA during replication, repair, and recombination (Wold, 1997). RPA gene expression is tied to replication, the cell cycle, and response to DNA damage (Ishibashi et al., 2001; Chang et al., 2009; Marwedel et al., 2003; Takashi et al., 2009).

Glyma20g24590, named GmRPA3c for this study, was found to have greater expression in iron-inefficient line Isoclark (PI 547430) than in iron-efficient line Clark (PI 548533) after 10 days of iron stress (O’Rourke et al., 2009). In our early response study, we found that Glyma20g24590 (GmRPA3c) expression is up-regulated in Isoclark at 24 hps, but down-regulated in Clark at the same time point. This confirms the differential expression seen in the microarray study by O’Rourke et al. (2009). In order to determine if the replication protein A complex responds to iron stress, homologs of RPA genes were identified in soybean, and gene expression was determined at 1, 6, and 24 hours post iron stress.

RPA homologs were identified in Glycine max and protein sequences aligned with RPA protein sequences in *Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*, and *Ricinus communis* (Figures 2, 3, and 4). Protein sequence of the OB-fold was most
conserved across species for both the RPA1 and RPA2 subunits, and the entirety of RPA3 appears to be well conserved. Though the sequence is well conserved across species, the C-terminus of RPA1 varies among types of RPA1 proteins. RPA1A and RPA1C appear to be very similar in the C-terminal region, while RPA1B differs. The function of different RPA1 proteins has been studied in *Arabidopsis thaliana* and *Oryza sativa*, finding that RPA1A and RPA1C are involved in DNA replication, while RPA1B may be involved in DNA repair (Ishibashi et. al, 2005; 2006). The variation in the C-terminal region may be indicative of the difference in cellular function.

Based on cellular localization and knock out studies in rice and *Arabidopsis thaliana*, three distinct RPA complexes have been described in plants (Sakaguchi et al., 2009; Ishibashi et al., 2006). RPA1A localizes to the chloroplast, and its knock out in *Arabidopsis* is lethal (Ishibashi et al., 2005; 2006). *GmRPA1Aa* in soybean has the greatest homology to *AtRPA1A*, suggesting it may be involved in DNA replication in the chloroplast as well. RPA1B localizes to the nucleus, and a knock out in *Arabidopsis* is viable, but with sensitivity to UV and MMS (Ishibashi et al., 2005; 2006). *GmRPA1Bb* in soybean has the greatest homology to *AtRPA1B*, suggesting it may also be involved in DNA repair in the nucleus. Preliminary results suggest knock out of AtRPA1C may also be lethal, and localization studies find it in the nucleus (Ishibashi et al., 2005; 2006). *GmRPA1Af* has the highest homology to *AtRPA1C*.

Only three RPA1 soybean genes, *GmRPA1Aa*, *GmRPA1Af*, and *GmRPA1Bb*, were found to be expressed. Interestingly, each of the soybean RPA1 genes studied corresponds to one of the three functionally characterized genes in *Arabidopsis thaliana*.
and rice (RPA1A, RPA1B, and RPA1C). The other six RPA1 homologs identified (see Table 1) did not express under our conditions (data not shown). Three were truncated, low-confidence genes that are likely pseudogenes with no expression. The other three were those found in the homeologous regions of GmRPA1Aa, GmRPA1Af, and GmRPA1Bb. These may be pseudogenes with no expression, or they may not be expressed under the conditions used in this study.

GmRPA1Aa and GmRPA1Af expression did not change in the iron-efficient line Clark, but GmRPA1Bb is significantly down-regulated at 24 hps. Though the Arabidopsis homolog of GmRPA1Bb is indicated in DNA repair, it has not been functionally characterized in soybean. It is possible GmRPA1Bb is involved in DNA replication, DNA repair, or both. Two separate knock out studies of AtRPA1A found contradictory results: one knock out mutant was lethal, while the other was viable with sensitivity to mutagens (Takashi et al., 2009; Ishibashi et al., 2005). Their results suggest that RPA1 proteins may have multiple roles in the plant cell. Our gene expression results suggest Clark, the iron-efficient genotype, may be decreasing its needs for DNA replication or repair under iron stress.

Four of the five RPA2 soybean genes were down-regulated in Clark at 24 hps. RPA2 is important because of its regulatory function in the RPA complex (Binz et al., 2004). Functional studies have tied RPA2 to both DNA replication and repair. Three independent studies found that T-DNA insertion mutants of the AtRPA2 gene AtRPA2A (At2g24490) resulted in stunted plants, earlier flowering, and an increased sensitivity to DNA damaging agent MMS (Elmayan et al., 2005; Kapoor et al., 2005; Xia et al., 2006).
Ishibashi et al. (2006) hypothesized that RPA2A is involved in DNA repair based on its interaction with RPA1B, while RPA2B may be involved in DNA replication based on its interaction with RPA1A. Again, our gene expression results suggest Clark may be decreasing its needs for DNA replication, repair, or both, under iron stress.

GmRPA2a is the only subunit that does not follow the same trend as the rest of the RPA genes studied. GmRPA2a has less homology to RPA2 genes across species, and has a much different gene structure than the other RPA2 homologs in soybean. GmRPA2a may be an RPA gene, but it may have different functions than the rest of the gene family, explaining its difference in gene expression.

Functional studies of RPA3 have not been done in plants, though a knock out of RPA3 in yeast was lethal (Brill et al., 1991). However, yeast only contains one RPA3 gene, while soybean has four homologs. Given the differential functions of RPA1 proteins in plants, multiple RPA3 genes may suggest multiple RPA3 proteins with different functions. Three of the four soybean RPA3 homologs express during iron stress. GmRPA3b remains unchanged in Clark under iron stress, while GmRPA3c and GmRPA3d are significantly down-regulated at 24 hps in Clark.

Clark (iron-efficient) down-regulates homologs of all three RPA subunits under iron stress, suggesting a decrease in DNA replication or repair. A decrease in DNA replication may slow growth and decrease the need for nutrients, such as iron. In an environment where iron is not readily available, such as calcareous soils, this may be an advantage. Often, iron is unavailable to the plant early in the growing season, when soils are cool and moist (Hansen et al., 2003). If the plant is able to adapt to this
environment by decreasing its overall need for iron, it may be able to better recover and grow when soils dry, and iron is more available.

The near-isogenic line of Clark, Isoclark, is iron-inefficient. Interestingly, the \textit{RPA} genes have the opposite expression response to iron stress in Isoclark. Isoclark increases its expression of \textit{GmRPA1Af} and \textit{GmRPA1Bb} at 24 hours post iron insufficiency, while \textit{GmRPA1Aa} remains unchanged. Three of the four \textit{RPA2} genes homologous to \textit{AtRPA2} increase gene expression at 24 hps, while the fourth does not change. \textit{GmRPA3c} and \textit{GmRPA3d} also increase expression at 24 hps. These results suggest Isoclark is increasing its DNA replication or repair mechanisms under iron stress.

It is unclear why Isoclark has the opposite expression pattern of Clark. It may be increasing its need for DNA repair, if iron stress causes genotoxic stress as well. Though it’s well known that an overabundance of iron causes genotoxic stress, no studies have been published indicating a lack of iron can cause the same stress in plants. The microarray study published by O’Rourke et al. (2009) found that Clark and Isoclark respond to the iron stress response differently, with most differentially expressed genes outside of known iron efficiency QTL. The O’Rourke et al. (2009) study suggested transcription factors within iron efficiency QTL might be differentially regulating genes. The opposite expression pattern of RPA genes found in this study supports the hypothesis that Isoclark cannot respond correctly to iron stress.

Our study indicates that replication protein A genes respond to iron stress in soybean. Iron-efficient line Clark may be down-regulating \textit{RPA} genes in order to slow
DNA replication and decrease the need for iron. Isoclark has the opposite expression of Clark, which may partly explain its iron inefficiency. Functional studies, such as gene knock outs or silencing, could elucidate the exact role of RPA in the iron stress response.

References


<table>
<thead>
<tr>
<th>Gene of Interest</th>
<th>Gene ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmRPA1Aa</td>
<td>Glyma15g19090</td>
</tr>
<tr>
<td>GmRPA1Ab</td>
<td>Glyma20g19560</td>
</tr>
<tr>
<td>GmRPA1Ac</td>
<td>Glyma09g07850</td>
</tr>
<tr>
<td>GmRPA1Ad</td>
<td>Glyma14g12480</td>
</tr>
<tr>
<td>GmRPA1Ae</td>
<td>Glyma06g32800</td>
</tr>
<tr>
<td>GmRPA1Af</td>
<td>Glyma09g34670</td>
</tr>
<tr>
<td>GmRPA1Ag</td>
<td>Glyma04g34970</td>
</tr>
<tr>
<td>GmRPA1Ba</td>
<td>Glyma17g08660</td>
</tr>
<tr>
<td>GmRPA1Bb</td>
<td>Glyma05g00370</td>
</tr>
<tr>
<td>GmRPA2a</td>
<td>Glyma08g18770</td>
</tr>
<tr>
<td>GmRPA2b</td>
<td>Glyma14g17270</td>
</tr>
<tr>
<td>GmRPA2c</td>
<td>Glyma17g29730</td>
</tr>
<tr>
<td>GmRPA2d</td>
<td>Glyma17g07020</td>
</tr>
<tr>
<td>GmRPA2e</td>
<td>Glyma13g00960</td>
</tr>
<tr>
<td>GmRPA3a</td>
<td>Glyma19g38100</td>
</tr>
<tr>
<td>GmRPA3b</td>
<td>Glyma03g35460</td>
</tr>
<tr>
<td>GmRPA3c</td>
<td>Glyma20g24590</td>
</tr>
<tr>
<td>GmRPA3d</td>
<td>Glyma10g42530</td>
</tr>
</tbody>
</table>

**Table 1.** Homologs to *Arabidopsis thaliana* RPA genes in *Glycine max*. See Materials and Methods for homolog identification explanation. Genes of interest were named for use throughout the study.
<table>
<thead>
<tr>
<th>Gene of Interest</th>
<th>Gene ID</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Amplicon Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPA1Aa</td>
<td>Glyma15g19090</td>
<td>CAGTGAGAGCTTCAGATTCAGC</td>
<td>GGTGGTTGGGTATGGAGGTTG</td>
<td>151</td>
</tr>
<tr>
<td>RPA1Af</td>
<td>Glyma09g34670</td>
<td>AAGTTACAGAAAGGTTCGGTG</td>
<td>ATCTTTTGGTGCAAGGACAG</td>
<td>132</td>
</tr>
<tr>
<td>RPA1Bb</td>
<td>Glyma05g00370</td>
<td>ATTCTACGCGAACAGCATGG</td>
<td>AGGTACGCATGTTCCCTTTG</td>
<td>145</td>
</tr>
<tr>
<td>RPA2a</td>
<td>Glyma08g18770</td>
<td>ACACCAACACCAAGCAGT</td>
<td>GAGATCGGGATTCTCTGC</td>
<td>136</td>
</tr>
<tr>
<td>RPA2b</td>
<td>Glyma14g17270</td>
<td>GGGTTTCGAAGATTTCTCG</td>
<td>GCCTCGCTTTATCTGCTAC</td>
<td>169</td>
</tr>
<tr>
<td>RPA2c</td>
<td>Glyma17g29730</td>
<td>TCTCAGGCGACCATCATCAA</td>
<td>AAGCTTGGACTCCGTGATA</td>
<td>156</td>
</tr>
<tr>
<td>RPA2d</td>
<td>Glyma17g07020</td>
<td>TCAAGCGATGATAGATCAATT</td>
<td>CCAGTACCATCATCAAGCACA</td>
<td>128</td>
</tr>
<tr>
<td>RPA2e</td>
<td>Glyma13g00960</td>
<td>GGAATGAGGGGTCCATC</td>
<td>CAACCATTGAGTTGACTTTATAGT</td>
<td>145</td>
</tr>
<tr>
<td>RPA3b</td>
<td>Glyma03g35460</td>
<td>TGGTCTTGTAAACAGTGATAAGTCAA</td>
<td>GGAACAAGTGTTAAAATTCTCCA</td>
<td>125</td>
</tr>
<tr>
<td>RPA3c</td>
<td>Glyma20g24590</td>
<td>CCAATGTGATGTTGGGGTTA</td>
<td>AAGGTGCTGCAAAGTCAGT</td>
<td>165</td>
</tr>
<tr>
<td>RPA3d</td>
<td>Glyma10g42530</td>
<td>AACGCAAAGGTTTCTTTCC</td>
<td>AAGTTGGGAATCAACTGAGCA</td>
<td>147</td>
</tr>
</tbody>
</table>

Table 2. Primer sequences for GmRPA homologs. Primers were designed with Primer3 v. 0.4.0 (Rozen et al., 2000).
Figure 1. Evolutionary history of the RPA1 subunit in Arabidopsis thaliana, Oryza sativa, Medicago truncatula, Ricinus communis, and Glycine max. The evolutionary history was inferred from the sequence alignment in Figure 2, using the Neighbor-Joining method (Saitou et al., 1987). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein et al., 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distance method (Nei et al., 2000) and are in the units of the number of amino acid differences per site. The analysis involved 19 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 265 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).
Figure 2. Multiple Amino Acid Sequence Alignment of RPA1 subunits in *Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*, *Ricinus communis*, and *Glycine max*. Sequences were aligned using Pileup in GCG (Accelrys Inc., San Diego, CA). Extraneous sequence was trimmed to include the most conserved regions. Sequence in the oligonucleotide-binding fold (OB-fold) and C-terminus is underlined in red and yellow, respectively.
Figure 2. (continued)
Figure 2. (continued)
<table>
<thead>
<tr>
<th>Strain</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gm15g19090</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Gm09g07850</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Mcg061820</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>A12g06510</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>RV:987</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>C02532686</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Gm19g34670</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>C04g34970</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Mcg027080</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>A16g19130</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>A9g45480</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Re22612.T000002.29</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>C09g107040</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Gm05g08370</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Gm17g08660</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>A12g08020</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>A15g61080</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>Re30183.T000143.20</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
<tr>
<td>C06g11540</td>
<td>KEELELYGDEOKVKITTV</td>
</tr>
</tbody>
</table>

Figure 2. (continued)
**Figure 3. Multiple Amino Acid Sequence Alignment of RPA2 subunits in** *Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*, *Ricinus communis*, and *Glycine max*.  Sequences were aligned using Pileup in GCG (Accelrys Inc., San Diego, CA). Extraneous sequence was trimmed to include the most conserved regions. Sequence in the oligonucleotide-binding fold (OB-fold) underlined in red.
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gm10g42530</td>
<td>MDTIIPVNAQMPINPIGKVRITYVNSQCGGVATAKSKIDDCQITKGL - QVLLMNVEVAGAE 68</td>
</tr>
<tr>
<td>Gm20g24550</td>
<td>MDTIIPVNAQMPINPIGKVRITYVNSQCGGVTTAKSKIDDCQITKGL - QVLLMNVEVAGAE 68</td>
</tr>
<tr>
<td>Gm39g35460</td>
<td>MDIIPVNAELHHFVGRRRRTIVQQVRSDDGGVIIKSDDEKVLVKGSPPAELTTFEVEVQTVN 70</td>
</tr>
<tr>
<td>Gm19g38100</td>
<td>MDIIPVNAELHHFVGRRRRTIVQQVRSDDGGVIIKSDDEKVLVKGSPPAELTTFEVEVQTVN 70</td>
</tr>
<tr>
<td>Mt7g127630</td>
<td>MDPILSPAVINELPPYVVGRVRTIVQQVRSDDVLAVIKSDDEKLYVKGSPPAELTTFEVEVQTVN 68</td>
</tr>
<tr>
<td>Mt3g52630</td>
<td>MDPILSPAVINELPPYVVGRVRTIVQQVRSDDVLAVIKSDDEKLYVKGSPPAELTTFEVEVQTVN 68</td>
</tr>
<tr>
<td>Rc28s69.T000007_28</td>
<td>MDPILSPAVINELPPYVVGRVRTIVQQVRSDDVLAVIKSDDEKLYVKGSPPAELTTFEVEVQTVN 68</td>
</tr>
<tr>
<td>Os01g254980</td>
<td>MDPILSPAVINELPPYVVGRVRTIVQQVRSDDVLAVIKSDDEKLYVKGSPPAELTTFEVEVQTVN 68</td>
</tr>
</tbody>
</table>

Figure 4. Multiple Amino Acid Sequence Alignment of RPA3 subunits in *Arabidopsis thaliana*, *Oryza sativa*, *Medicago truncatula*, *Ricinus communis*, and *Glycine max*. Sequences were aligned using Pileup in GCG (Accelrys Inc., San Diego, CA). Extraneous sequence was trimmed to include the most conserved regions.
Figure 5. Relative gene expression values of RPA homologs over three time points in iron-efficient line Clark. Relative gene expression was determined with quantitative Reverse Transcription PCR (qRT-PCR). Values are the change in relative expression values from 1 hour post stress (hps). Relative gene expression is a ratio to the value at the same time point in iron sufficient conditions, averaged over three biological replicates, then log transformed (base 2).
Figure 6. Relative gene expression values of RPA homologs over three time points in iron-inefficient line *Isoclar*. Relative gene expression was determined with quantitative Reverse Transcription PCR (qRT-PCR). Values are the change in relative expression values from 1 hour post stress (hps). Relative gene expression is a ratio to the value at the same time point in iron sufficient conditions, averaged over three biological replicates, then log transformed (base 2).
CHAPTER 3. Silencing replication protein A subunit 3 suggests a role in the iron stress response

Atwood, Sarah E.¹, O’Rourke, Jamie A.², Zhang, Chunquan³, Hill, John H.³, Whitham, Steven A.³, Graham, Michelle A.⁴*, Shoemaker, Randy C.⁴*

* Corresponding author; e-mails: michelle.graham@ars.usda.gov, randy.shoemaker@ars.usda.gov

¹ Interdepartmental Genetics Program, Iowa State University, Ames, Iowa 50011

² USDA-Agricultural Research Service, Plant Science Research Unit, Saint Paul, Minnesota 55108

³ Department of Plant Pathology, Iowa State University, Ames, Iowa 50011

⁴ USDA-Agricultural Research Service, Corn Insects and Crop Genetics Research Unit, Ames, Iowa 50011

Abstract

Replication protein A (RPA) binds and stabilizes single-stranded DNA during DNA replication, repair, and recombination. RPA is composed of three subunits: RPA1 (70 kDa), RPA2 (32 kDa), and RPA3 (14 kDa). RPA1 and RPA2 have been well studied in yeast, humans, and plants, but so far the role of RPA3 is not well known. Using virus-induced gene silencing (VIGS), we show RPA3 is essential for proper growth of soybean as RPA3 silenced plants show a stunted growth phenotype. The silencing of RPA3 also produced plants with reduced iron deficiency chlorosis during iron stress, suggesting a link between DNA replication and the iron stress response.
Introduction

Replication protein A (RPA) is an essential protein that binds and stabilizes single-stranded DNA (ssDNA) during DNA replication, repair, and recombination (Wold, 1997). RPA is essential for proper growth and development in yeast and plants (Brill et al., 1991; Ishibashi et al., 2005). RPA is composed of three subunits: RPA1 (70 kDa), RPA2 (32 kDa), and RPA3 (14 kDa) (Wold, 1997). RPA1 has multiple OB-folds (oligonucleotide-binding folds) and acts as the ssDNA-binding subunit (Pfuetzner et al., 1997). RPA2 has a phosphorylation domain important in regulating RPA activity during the cell cycle and stress response (Binz et al., 2004; Din et al., 1990). The role of RPA3, however, is still unclear.

Replication protein A was first identified in humans as a protein required for the replication of simian virus 40 DNA in vitro (Wold et al., 1988). Initial studies identified human RPA (hRPA) to be a single-stranded DNA binding protein (SSB), binding single-stranded DNA and promoting the rapid unwinding of SV40 DNA at the origin of replication (Wold et al., 1988; Kenny et al., 1989). RPA has been studied most extensively in humans and yeast, but more recently orthologs of RPA have been discovered in plants such as rice and Arabidopsis (van der Knaap et al., 1997; Ishibashi et al., 2001). The first discovery of RPA in plants was in deepwater rice, when gene expression of a putative RPA1 gene was induced by gibberellin (van der Knaap et al., 1997).

Unlike humans and yeast, rice and Arabidopsis contain multiple copies of the RPA genes whose proteins form complexes with unique roles in the cell (Ishibashi et al.,
2006). All three RPA complexes found in rice and Arabidopsis share the same RPA3 subunit, but its role has not been tested in plants (Sakaguchi et al., 2009). Soybean (Glycine max (L.) Merr.) contains at least four copies of the RPA3 gene, and the exact role of the different protein products is unknown.

RPA3 is the smallest subunit, but is important in RPA complex formation and RPA binding to ssDNA. Knock out of RPA3 in yeast is lethal, and RPA3 is required for correct formation of the RPA complex (Henricksen et al., 1994; Brill et al., 1991). Although RPA1 is considered to be the DNA-binding subunit, RPA3 in combination with RPA2 improves the ssDNA-binding ability of RPA1 (Kim et al., 1994). RPA3 has been implicated in the response to stress as well. RPA3 binds nucleolin under heat shock and genotoxic stress, inhibiting DNA replication (Daniely et al., 2000; Kim et al., 2005).

RPA3 may also play a role in the iron stress response of soybean. In a recent microarray study, DNA replication and repair genes were over-represented among differentially expressed genes in Clark (iron-efficient) during iron stress (O'Rourke et al., 2009). A probe corresponding to a replication factor (GmaAffx.36066.1.S1_at) was also found to be differentially expressed between Clark and Isoclark, two near-isogenic lines (NILs) differing in iron efficiency, during iron stress (O'Rourke et al., 2009). Upon further investigation, it was found that GmaAffx36066.1.S1_at is representative of gene expression for a homolog of replication protein A subunit 3 (RPA3). In addition to differential expression between the two NILs, this homolog of RPA3 was located within a known iron efficiency QTL on soybean chromosome 20 thought to be involved in chlorophyll content variation in iron deficiency chlorosis mapping populations (Lin et
In this work, we used virus-induced gene silencing (VIGS) to study the role of \textit{RPA3} in the iron stress response in soybean.

VIGS begins by inserting a small fragment of target gene sequence from the host into the virus genome (Lu, 2003; Voinnet, 2001). After infection with the recombinant virus, aberrant double-stranded RNA (dsRNA) molecules produced during viral replication trigger the creation of small-interfering RNAs (siRNAs) specific to not only the virus genome, but the foreign host gene fragment as well (Purkayastha et al., 2009). Small-interfering RNAs specific to the host sequence bind to endogenous mRNA, triggering mRNA degradation by the RNA-induced silencing complex (RISC) (Purkayastha et al., 2009). The VIGS system is ideal for many plant species because it does not require transformation, vectors are simple to create, and results can be achieved in as little as 3 weeks post inoculation (Burch-Smith et al., 2004). Another advantage of VIGS is the ability to silence multiple genes with sequence homology, circumventing the hindrance of functional redundancy in knock out studies (Lawrence et al., 2003; Meyer et al., 2009).

A successful VIGS system for soybean has been developed using \textit{Bean pod mottle virus} (BPMV) (Zhang et al., 2010). BPMV is an RNA virus, containing two RNA molecules: RNA1, which includes five genes required for virus replication, and RNA2, which includes the cell-to-cell movement protein and coat proteins (Zhang et al., 2006). The host gene sequence is inserted after the translation stop codon in RNA2, allowing for insertion of antisense sequences, shown to be most effective for silencing endogenous genes (Zhang et al., 2009; 2010). The BPMV vector for soybean has been
successfully used to identify genes related to *Rpp2* and *Rpp4* resistance against Asian soybean rust (Meyer et al., 2009; Pandey et al., 2011).

In this study, two homeologous *RPA3* genes with high nucleotide identity (97%), Glyma20g24590 and Glyma10g42530, were silenced with one BPMV VIGS construct in soybean. Silenced plants showed vegetative growth stunting as well as an improved tolerance to iron stress. We conclude that RPA3 is essential for proper growth of soybean, and hypothesize a link between DNA replication and growth and the iron stress response.

**Materials and Methods**

**Germplasm**

Clark (PI 548533), Isoclark (PI 547430) and T203 (PI 54619) were used to study the role of RPA in growth and iron homeostasis. Clark is iron-efficient, while Isoclark and T203 are iron-inefficient. Isoclark is a near-isogenic line (NIL) of Clark, containing an introgression from donor parent T203 that is hypothesized to cause iron inefficiency.

**Vector Creation**

A 302 base pair segment of Glyma20g24590 (*RPA3*) was amplified with forward (5’-ATGCGGATTCCTCTTCTGTATTTGTAAATGCTCAG-3’) and reverse (5’-ATGCGGATTCGCAGACCCCTTTAAATTCAACA-3’) primers containing BamHI sites using Invitrogen™ Platinum® *Taq* DNA Polymerase High Fidelity and Clark cDNA as
template. The resulting amplicon was cloned into the BamHI cloning site in pBPMV-IA-V1 RNA2 (Zhang et al., 2010). Vectors with RPA3 target sequence in the sense (BPMV-RPA3-S) and antisense (BPMV-RPA3-AS) direction were confirmed via sequencing. BPMV-E (empty vector), BPMV-RPA3-S, or BPMV-RPA3-AS constructs were bombarded into Clark and Isoclark seedlings (~10 days old). After 42 days, tissue was flash frozen with liquid nitrogen and stored at -80°C. BPMV infection was confirmed with ELISA (see below). Stored tissue was used for rub inoculation in subsequent experiments.

Iron Stress Experiment

48 Isoclark seed, 12 Clark seed, and 6 T203 seed were germinated on paper for 5 to 7 days at 27°C before moving into hydroponic buckets. 8 Isoclark, 2 Clark, and 1 T203 seedlings were placed in 10 L buckets with nutrient solutions that were either iron sufficient (100 µM Fe(NO₃)₃•9H₂O; 3 buckets) or iron insufficient (50 µM Fe(NO₃)₃•9H₂O; 3 buckets). Nutrient solutions were based on growth conditions described in Chaney et al. (1992), with solutions adjusted for 10 L buckets. Fully expanded unifoliates of six Isoclark seedlings in each treatment (two Isoclark seedlings in each bucket) were rub-inoculated with one of four treatments: Mock (buffer control), BPMV-E, BPMV-RPA3-S, or BPMV-RPA3-AS. Clark (iron-efficient) and T203 (iron-inefficient) served as bucket controls to ensure proper conditions for developing an iron deficiency chlorosis (IDC) phenotype.
Vegetative growth, IDC score and chlorophyll content were assessed at 21 days post inoculation. IDC score and chlorophyll content were assessed on the expanding trifoliate. Vegetative growth was measured as height in centimeters. IDC score was rated on a scale of 1 to 5, with 1 being green and healthy leaves and 5 being yellow and necrotic leaves. Chlorophyll content was assessed with a Minolta SPAD-520 Chlorophyll Meter. Phenotypic differences were analyzed with a student’s t-test (Ramsey et al., 2002) with a significance cut-off of 0.05. Equal variance was assumed among data sets.

Scored leaf tissue was harvested at 21 days post inoculation, flash frozen with liquid nitrogen, and stored at -80°C. Stored tissue was used for RNA extraction and confirmation of BPMV infection by ELISA.

Soil Experiment

Six Clark and six Isoclar seed were germinated on paper for 5-7 days at 27°C before transplanting into pots (two plants per pot) with sterile soil. Fully expanded unifoliates were dusted with carborundum and rubbed with virus infected tissue ground in phosphate buffer. Two Clark and two Isoclar seedlings were rub-inoculated with each treatment: BPMV-E, BPMV-RPA3-S, or BPMV-RPA3-AS. Vegetative growth, measured as height in centimeters, was taken at four time points: 21, 30, 36, and 42 days post inoculation. Plants were grown in a growth chamber at 20°C with 16 hours light. The experiment was repeated with an additional inter-node length phenotype
measured at 42 days post inoculation. Inter-node lengths were measured in centimeters between each node beginning at the unifoliate node.

RNA Isolation

Flash frozen tissue was ground in liquid nitrogen and RNA extracted using a Qiagen® RNeasy® Plant Mini Kit. The RNeasy® Mini Handbook protocol was used with the following specifications or changes: ~200 mg tissue was lysed with RLT buffer, tubes were incubated at 56°C Celsius for 2 minutes with 800 rpm shaking to aid in tissue disruption, and columns were incubated at room temperature for 10 minutes during elution. RNA was then DNased with an Ambion® TURBO DNA-free™ kit to remove all DNA.

After isolation, RNA was assessed for quality using a Thermo Fisher Scientific® NanoDrop™ ND-1000 Spectrophotometer. RNA was considered to be of good quality for qRT-PCR if the 260/280 ratio was above 2.0 and the 260/230 ratio was above 1.7.

Reverse Transcription and qRT-PCR

RNA templates were amplified using the Invitrogen™ SuperScript™ III Platinum® SYBR® Green One-Step qRT-PCR Kit. Reactions were carried out according to the SYBR® Green protocol with the following specifications: total starting RNA was 50 ng, and reactions were a final volume of 25 µl instead of 50 µl. RNA was diluted to 4.76 ng/µl for greater pipetting accuracy. All experiments included a standard curve of 200, 50, 25 and 5 ng concentration as well as No Reverse-Transcriptase (NRT) and No
Template Control (NTC) wells for each primer. NRT wells replaced Superscript III with Invitrogen® Platinum® Taq DNA polymerase at the same volume. NTC wells replaced RNA with double distilled, nuclease-free H₂O at the same volume.

qRT-PCR was carried out on a Stratagene Mx3000P™ Real-Time PCR System. After amplification, a dissociation reaction was performed for later analysis of reaction quality. Reaction conditions were modified from the Invitrogen™ SuperScript™ III Platinum® SYBR® Green One-Step qRT-PCR Kit: cDNA synthesis was carried out at 60°C, initial denaturing time was 10 minutes, and a total of 45 cycles were carried out with an additional extension time of 15 seconds at 72°C. A measurement of fluorescence was taken after each cycle. The default SYBR® Green dissociation reaction conditions were used from the Stratagene Mx3000P™ Real-Time PCR software.

Each 96 well plate contained all biological replicates in either the iron sufficient or iron insufficient buckets in two technical replicates. Each plate amplified an RPA3 specific primer and one reference gene primer, CYP2 (cyclophilin 2). CYP2 was chosen as a reference gene based upon previous qRT-PCR reference gene studies in plants as well as an in-house study (Phillips et al., 2009; Wang et al., 2011).

302 of the 321 base pairs of the Glyma20g24590 coding sequence was used in the vector creation. Genomic sequence for Glyma20g24590 and the target in the VIGS vector overlap by 95%, making it impossible to make an amplicon specific for the plant cDNA and not the target gene fragment in the virus. Amplification of the virus would mask effects of silencing. Therefore, given the high nucleotide identity between Glyma20g24590 and its homeolog (Glyma10g42530), we designed gene specific
primers for Glyma10g42530, assuming silencing trends observed for it would reflect silencing of Glyma20g24590 (forward: 5’-CCAGCAATTAGATGGGGTTT-3’ and reverse: 5’-GCTAACATCAGAGATAATGGAACA-3’).

qRT-PCR Data Analysis

Reactions were considered of good quality for further data analysis if the No Reverse Transcriptase (NRT) and No Template Control (NTC) cycle thresholds were greater than 5 cycles away from the lowest data point and the dissociation curve showed only one peak for reaction wells. Samples were discarded if duplicate sample thresholds had a difference greater than 1 cycle.

Each sample was amplified with an *RPA3* specific primer pair and reference gene primer pair on the same plate, allowing for direct comparison of values (Rieu et al., 2009). RNA quantities were determined for each well by aligning to the standard curve for that primer set. A normalized value for RNA quantity was calculated as a ratio of *RPA3* RNA quantity to reference gene RNA quantity for each sample. Averages of the normalized data were then calculated across technical replicates. This average was then log-transformed with base 2. Log-transformed data was analyzed for standard deviation and standard error.

Differences in relative quantity were analyzed with an Analysis of Variance test (ANOVA) (Chambers et al., 1992) and then Tukey’s Honestly Significant Difference test (HSD) (Yandell, 1997) for pair-wise comparisons, with a significance cut-off of 0.05.
ELISA

Infection was confirmed with a Double Antibody Sandwich (DAS) Enzyme-linked immunosorbent assay (ELISA) PathoScreen® BPMV Kit from Agdia®. All samples were tested in duplicate.

Results

Soybean has four homologs of replication protein A subunit 3 (RPA3), the smallest subunit of the single-stranded DNA binding protein replication protein A (RPA). One of these homologs, Glyma20g24590 (RPA3), and other cell cycle genes were identified as differentially expressed in soybean under iron stress (O’Rourke et al., 2009). Glyma20g24590 was targeted for silencing with virus-induced gene silencing (VIGS) in order to determine its role in plant development and iron homeostasis.

Glyma20g24590 has two exons, a coding sequence of 321 nucleotides (http://www.phytozome.net/soybean), and high nucleotide identity (97%) to its homeolog, Glyma10g42530. Two BPMV VIGS vectors were created that target both Glyma20g24590 and Glyma10g42530 for silencing: BPMV-RPA3-S has the target sequence inserted in the sense orientation, while BPMV-RPA3-AS has the target sequence inserted in the antisense orientation. Sense and antisense constructs have been previously shown to have different strengths of silencing, with antisense having a greater silencing activity (Zhang et al., 2010).

RPA3 was targeted for silencing in Isoclark (iron-inefficient). Microarray data has shown that Isoclark (iron-inefficient) has more RPA3 expression than Clark (iron-
efficient) under iron stress (O’Rourke et al., 2009). In another study, Isoclark was found to up-regulate *RPA3* after 24 hours of iron stress, while Clark down-regulated *RPA3* at the same time point (unpublished results). These expression results suggest a down-regulation of *RPA3* may be important for resistance to iron stress.

If a down-regulation of *RPA3* confers iron stress resistance, silencing *RPA3* in Isoclark during iron stress should improve IDC symptoms. In fact, twenty-one days post inoculation, roughly 28 days after iron stress, plants inoculated with *RPA3* antisense constructs (BPMV-*RPA3*-AS) were found to be smaller and with reduced chlorosis compared to those inoculated with the BPMV virus alone (BPMV-E). BPMV-*RPA3*-AS plants were ~30% shorter on average than plants inoculated with BPMV-E (Figure 1). IDC visual scores improved by 1 point, from an average of 3 (interveinal chlorosis) in plants inoculated with BPMV-E to an average of 2 (slight yellowing) in plants inoculated with BPMV-*RPA3*-AS (Figure 2). SPAD readings demonstrated greater chlorophyll content in BPMV-*RPA3*-AS plants ($\bar{x} = 25$) than BPMV-E plants ($\bar{x} = 15$) (Figure 3). Though height was affected in iron sufficient conditions, chlorophyll content was not significantly different among any of the four treatments (Figures 1 and 3).

Tissue in iron stress conditions was harvested at 21 dpi in order to observe silencing at its strongest. Glyma10g42530 silencing was confirmed in BPMV-*RPA3*-AS plants at an average level of 3-fold less expression than BPMV-E plants (Figure 4). Glyma10g42530 was silenced in BPMV-*RPA3*-S plants, but to a lesser extent; expression was an average of 2-fold less than BPMV-E plants (Figure 4).
RPA3 silencing in soil confirms the vegetative stunting phenotype. Two plants of each genotype (Clark and Isoclark) were inoculated via rub inoculation for the treatments BPMV-E, BPMV-RPA3-S and BPMV-RPA3-AS, for a total of four plants per treatment. All plants had virus symptoms, and Bean pod mottle virus (BPMV) infection was confirmed with an enzyme-linked immunosorbent assay (ELISA) (data not shown). Heights were measured at 21, 30, 36, and 42 dpi to track the effect of silencing over time. BPMV-RPA3-AS plants were ~25% shorter on average than BPMV-E plants at all time points (Figure 5). Inter-node length was decreased nearly 70% on average at the third and fourth trifoliate stage, when silencing is at its strongest (Figures 6 and 7). The stunting and inter-node length phenotypes suggest RPA3 plays an important role in plant growth.

Discussion

Replication protein A (RPA) is a single-stranded DNA binding protein essential for DNA replication, repair, and recombination (Wold, 1997). RPA has three subunits: RPA1 (70 kDa), RPA2 (32 kDa), and RPA3 (14 kDa) (Wold, 1997). The two largest subunits, RPA1 and RPA2, have been well studied in yeast, humans, and plants such as Arabidopsis and rice (Sakaguchi et al., 2009). However, the molecular role of RPA3, the smallest subunit, is not clear. Recently, a microarray study found Glyma20g24590 (RPA3), as well as other DNA replication and repair genes, to be differentially expressed under iron stress (O’Rourke et al., 2009), but it is unclear what role those genes may play in the iron stress response. Iron is an important micronutrient in many processes,
but especially in photosynthesis. Under iron deficiency, chlorophyll is depleted and photosynthetic capacity is compromised (Spiller et al., 1980). To determine if RPA3 could be involved in the response to iron deficiency, RPA3 was silenced in soybean under iron stress conditions. In addition, we attempted to learn more about the role of RPA3 in soybean growth and development.

Virus-induced gene silencing (VIGS) was used to silence gene expression of Glyma20g24590 and Glyma10g42530, homologs of RPA3 in soybean. A target sequence of 302 base pairs from Glyma20g24590 was used in the VIGS vector, which had 97% homology with sequence from its homeolog Glyma10g42530. The high homology in the target region allows the silencing of both RPA3 homeologs, avoiding the issue of functional redundancy masking a phenotype.

RPA3 was silenced in Isoclark, an iron-inefficient line that was shown to have greater RPA3 expression than Clark, its near-isogenic iron-efficient line under iron stress (O’Rourke et al., 2009). RPA3 silencing caused vegetative growth defects such as stunting and a decrease in inter-node lengths between the 2nd and 4th trifoliate nodes, suggesting a role in DNA replication. Silencing RPA3 in ‘Isoclark’ under iron stress resulted in plants with greater chlorophyll content and improved visual iron deficiency chlorosis (IDC) scores, suggesting a link between DNA replication and the response to iron stress.

Silencing RPA3 may have slowed growth by inhibiting DNA replication, and in doing so reduced the requirement for iron, resulting in healthier plants under iron deficiency. Clark, the iron-efficient near-isogenic line of Isoclark, down-regulates all
RPA subunits under iron stress (unpublished results). Slowing growth and decreasing the need for iron may be advantageous, as often iron deficiency is a problem early in the season, when soils are cool and wet, but iron becomes available as soils dry (Hansen et al., 2003).

Previous studies have also found DNA replication to be affected by iron depletion. In other species, iron depletion arrests the cell cycle at G₁ and leads to apoptosis (Yu et al., 2007). O’Rourke et al. (2009) found iron stress in an iron-efficient soybean line affected not only the expression of iron acquisition genes, but genes involved in DNA replication as well. These studies support that hypothesis that Clark may slow DNA replication during iron stress in an effort to prevent cellular damage.

The vegetative growth stunting caused by silencing RPA3 in soybean suggests an important molecular role in DNA replication. RPA1, the subunit responsible for single-stranded DNA binding, can only join the RPA complex when RPA2 and RPA3 are already bound together (Henricksen et al., 1994). It is likely that silencing RPA3 disrupted the creation of one or more RPA complexes, therefore inhibiting DNA replication controlled by those complexes. Disruption of gene expression of RPA1 and RPA2 in plants produces results that are consistent with DNA replication inhibition. Knock out of RPA1 results in lethality in some cases, but only mutagen sensitivity in others (Ishibashi et al., 2005; Takashi et al., 2009). Knock out of RPA2, the second subunit, results in stunted plants (Elmayan et al., 2005; Kapoor et al., 2005; Xia et al., 2006). In this study we demonstrate that silencing of RPA3 also causes stunting.
This study is the first instance of silencing \textit{RPA3} in plants. Silencing of \textit{RPA3} has shown it to be important not only in plant growth, but also in the response to iron stress. The improvement in iron deficiency chlorosis symptoms upon \textit{RPA3} silencing provides intriguing details to our understanding of the mechanisms of iron efficiency in soybean.

\textbf{References}


Figure 1. Silencing RPA3 alters vegetative growth under iron stress. Vegetative growth (measured as height in cm) of soybean inoculated via rub inoculation with treatments Mock (buffer control), BPMV-E (empty vector), BPMV-RPA3-S (sense vector), and BPMV-RPA3-AS (antisense vector). Plants inoculated with BPMV-RPA3-AS were ~30% shorter on average than BPMV-E plants 21 dpi in both iron conditions. Each data point is the average of 6 plants ± standard error. * BPMV-RPA3-AS is significantly different from BPMV-E in iron sufficient and insufficient conditions (p-values <0.05).
Figure 2. Silencing RPA3 alters chlorophyll content under iron stress. Expanding trifoliate leaves of plants inoculated with BPMV-RPA3-AS constructs had an improved IDC visual score when compared to plants inoculated with BPMV-E. Each row represents plants from an individual bucket. BPMV infection has been shown to produce mild chlorotic symptoms, thus, the mock infected leaves remain more green, even under iron stress.
Figure 3. Silencing RPA3 alters chlorophyll content under iron stress. 3rd trifoliate leaves of plants inoculated with BPMV-RPA3-AS constructs had an improved SPAD reading when compared to plants inoculated with BPMV-E. Each data point is the average of 6 plants ± standard error. * BPMV-E is significantly different from BPMV-RPA3-AS and BPMV-RPA3-S under iron stress (p-values <0.01).
Figure 4. Silencing confirmed in BPMV-RPA3 inoculated plants under iron stress. Expression of Glyma10g42530 is an average ~2-fold less in BPMV-RPA3-S plants when compared to BPMV-E plants. Expression of Glyma10g42530 is an average of ~3-fold less in BPMV-RPA3-AS plants than BPMV-E plants. Each data point is the average difference in normalized expression compared to BPMV-E ± standard error. * BPMV-RPA3-AS is significantly different from BPMV-E (p-value <0.001).
Figure 5. Silencing RPA3 affects vegetative growth. Vegetative growth (measured as height in cm) of soybean inoculated via rub inoculation with BPMV-E (empty vector), BPMV-RPA3-S (sense vector), and BPMV-RPA3-AS (antisense vector). Plants inoculated with BPMV-RPA3-AS averaged 25% shorter than BPMV-E plants across all four time points. Each data point is the average of 8 plants ± standard error. * BPMV-RPA3-AS is significantly different from BPMV-E at 21, 30, 36, and 42 dpi (p-value <0.01 in all cases).
Figure 6. Silencing RPA3 alters inter-node length. Inter-node length (measured in cm) of soybean inoculated via rub inoculation with BPMV-E (empty vector), BPMV-RPA3-S (sense vector), and BPMV-RPA3-AS (antisense vector). Inter-node lengths were measured between the 2\textsuperscript{nd} and 4\textsuperscript{th} trifoliate nodes at 42 dpi. Inter-node length of plants inoculated with BPMV-RPA3-AS was nearly 70\% shorter on average than BPMV-E plants. Each data point is the average of 4 plants ± standard error. * BPMV-RPA3-AS and BPMV-RPA3-S are significantly different from BPMV-E (p-values <0.001).
Figure 7. Silencing RPA3 alters inter-node length. Arrows show areas of decreased inter-node length. Inter-node length was decreased as much as 70% between the 2nd and 4th trifoliate nodes. Four biological replicates were assessed for each treatment.
CHAPTER 4. GENERAL CONCLUSIONS

Conclusions

Understanding the genetic basis behind iron efficiency in soybean is the first step in breeding for improved resistance to iron stress. Soybean lines tolerant to iron stress will increase yields and profits for many farmers in the upper Midwest, and may be useful in increasing iron content of soybean seed later used for human nutrition. A recent microarray study found DNA replication and repair genes to be over-represented among those differentially expressed in soybean during iron stress (O’Rourke et al., 2009). The role of DNA metabolism genes in the iron stress response has not been studied. This work aimed to study the role of replication protein A, a single-stranded DNA binding protein, in the iron stress response.

The first study found gene expression of replication protein A homologs had opposite expression patterns in two near-isogenic lines (NILs) of soybean, differing only in their iron efficiency, after 24 hours of iron stress. The second study found silencing of replication protein A subunit 3 (RPA3) improved iron deficiency chlorosis (IDC) symptoms in Isoclark (iron-inefficient) during iron stress. Silencing of RPA3 also resulted in stunted plants, which is consistent with a role in DNA replication.

Clark (iron-efficient) down-regulates the RPA genes during iron stress, and is healthier under iron stress than its NIL Isoclark (iron-inefficient). Clark also down-regulates its RPA genes during iron stress, while Isoclark does the opposite. Silencing RPA3 in Isoclark improved its IDC symptoms, suggesting a low expression level may play a role in tolerance to iron stress. We hypothesize that Clark may down-regulate
RPA genes in order to decrease DNA replication and slow growth, therefore allowing it to better utilize the limited iron available during iron stress.

**Future Research**

This research provides an interesting insight into our understanding of iron efficiency in soybean. Future research is needed to determine if reduced DNA replication and repair is a phenomenon seen in other iron efficient lines of soybean, and if our hypothesis tying together DNA replication and iron efficiency is appropriate.

The gene expression of the RPA subunits could easily be distinguished in multiple soybean lines with varying iron efficiencies to determine if the patterns we observed are universal or specific. Gene expression of other genes involved in DNA replication and/or repair could also be informative in determining if the change in RPA expression is more related to DNA replication or DNA repair.

Silencing RPA3 in iron-efficient line Clark under iron stress could enlighten us to any potential interaction between BPMV and IDC. It could also confirm the improvement of IDC symptoms obtained in Isoclark as specific to an improvement in the iron stress response.

The other subunits of RPA (RPA1 and RPA2) could also be silenced with the same VIGS system during iron stress to determine if improvement in iron efficiency was specific to the RPA3 subunit. Observing iron efficiency after silencing additional genes required for DNA replication, but not DNA repair, could determine if the effect is specific to DNA replication.
References

APPENDIX A. SUPPLEMENTARY TABLES

Table S1. Relative gene expression values of GmRPA homologs over three time points in iron efficient line Clark. Relative gene expression was determined with quantitative Reverse Transcription PCR (qRT-PCR). Relative gene expression values are presented as a ratio to the value at the same time point in iron sufficient conditions, averaged over three biological replicates, then log transformed (base 2).

<table>
<thead>
<tr>
<th>Gene of Interest ID</th>
<th>Time Point (Hours Post Stress)</th>
<th>Relative Gene Expression</th>
<th>Standard Error</th>
<th>p-value (ANOVA)</th>
<th>Pair-wise Comparison</th>
<th>p-value (Tukey’s Honestly Significant Difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmRPA1Aa</td>
<td>1 hps</td>
<td>-0.33</td>
<td>0.27</td>
<td>0.13</td>
<td>1-6</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>0.18</td>
<td>0.07</td>
<td></td>
<td>1-24</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.18</td>
<td>0.13</td>
<td></td>
<td>6-24</td>
<td>0.33</td>
</tr>
<tr>
<td>GmRPA1Af</td>
<td>1 hps</td>
<td>0.06</td>
<td>0.12</td>
<td><strong>0.02</strong></td>
<td>1-6</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>0.37</td>
<td>0.18</td>
<td></td>
<td>1-24</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.28</td>
<td>0.16</td>
<td></td>
<td>6-24</td>
<td><strong>0.02</strong></td>
</tr>
<tr>
<td>GmRPA1Bb</td>
<td>1 hps</td>
<td>0.10</td>
<td>0.28</td>
<td><strong>&lt;0.001</strong></td>
<td>1-6</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>1.47</td>
<td>0.56</td>
<td></td>
<td>1-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-4.09</td>
<td>0.58</td>
<td></td>
<td>6-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td>GmRPA2a</td>
<td>1 hps</td>
<td>-0.26</td>
<td>0.65</td>
<td><strong>&lt;0.001</strong></td>
<td>1-6</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-0.89</td>
<td>0.22</td>
<td></td>
<td>1-24</td>
<td><strong>0.01</strong></td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>1.78</td>
<td>0.29</td>
<td></td>
<td>6-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td>GmRPA2b</td>
<td>1 hps</td>
<td>1.43</td>
<td>0.28</td>
<td><strong>&lt;0.001</strong></td>
<td>1-6</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>2.09</td>
<td>1.11</td>
<td></td>
<td>1-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-5.28</td>
<td>0.49</td>
<td></td>
<td>6-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td>GmRPA2c</td>
<td>1 hps</td>
<td>0.54</td>
<td>0.32</td>
<td><strong>&lt;0.001</strong></td>
<td>1-6</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>1.81</td>
<td>0.72</td>
<td></td>
<td>1-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-4.12</td>
<td>0.53</td>
<td></td>
<td>6-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td>GmRPA2d</td>
<td>1 hps</td>
<td>-0.20</td>
<td>0.22</td>
<td><strong>&lt;0.001</strong></td>
<td>1-6</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>0.57</td>
<td>0.49</td>
<td></td>
<td>1-24</td>
<td><strong>0.02</strong></td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-1.52</td>
<td>0.17</td>
<td></td>
<td>6-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td>GmRPA2e</td>
<td>1 hps</td>
<td>0.23</td>
<td>0.27</td>
<td><strong>&lt;0.001</strong></td>
<td>1-6</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>1.10</td>
<td>0.27</td>
<td></td>
<td>1-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-2.36</td>
<td>0.53</td>
<td></td>
<td>6-24</td>
<td><strong>&lt;0.001</strong></td>
</tr>
</tbody>
</table>
Table S1. (continued)

<table>
<thead>
<tr>
<th>Gene of Interest ID</th>
<th>Time Point (Hours Post Stress)</th>
<th>Relative Gene Expression</th>
<th>Standard Error</th>
<th>p-value (ANOVA)</th>
<th>Pairwise Comparison</th>
<th>p-value (Tukey’s Honestly Significant Difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmRPA3b</td>
<td>1 hps</td>
<td>-0.37</td>
<td>0.08</td>
<td>0.31</td>
<td>1-6</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-0.60</td>
<td>0.30</td>
<td></td>
<td>1-24</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.00</td>
<td>0.36</td>
<td></td>
<td>6-24</td>
<td>0.29</td>
</tr>
<tr>
<td>GmRPA3c</td>
<td>1 hps</td>
<td>0.63</td>
<td>0.27</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>1.72</td>
<td>0.63</td>
<td></td>
<td>1-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-4.19</td>
<td>0.51</td>
<td></td>
<td>6-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GmRPA3d</td>
<td>1 hps</td>
<td>0.78</td>
<td>0.19</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>1.92</td>
<td>0.66</td>
<td></td>
<td>1-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-4.87</td>
<td>0.50</td>
<td></td>
<td>6-24</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table S2. Relative gene expression values of GmRPA homologs over three time points in iron inefficient line Isoclark. Relative gene expression was determined with quantitative Reverse Transcription PCR (qRT-PCR). Relative gene expression values are presented as a ratio to the value at the same time point in iron sufficient conditions, averaged over three biological replicates, then log transformed (base 2).

<table>
<thead>
<tr>
<th>Gene of Interest ID</th>
<th>Time Point (Hours Post Stress)</th>
<th>Relative Gene Expression</th>
<th>Standard Error</th>
<th>p-value (ANOVA)</th>
<th>Pairwise Comparison</th>
<th>p-value (Tukey's Honestly Significant Difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmRPA1Aa</td>
<td>1 hps</td>
<td>-0.05</td>
<td>0.13</td>
<td>0.95</td>
<td>1-6</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-0.12</td>
<td>0.20</td>
<td></td>
<td>1-24</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.10</td>
<td>0.15</td>
<td></td>
<td>6-24</td>
<td>0.99</td>
</tr>
<tr>
<td>GmRPA1Af</td>
<td>1 hps</td>
<td>-1.58</td>
<td>0.27</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-1.08</td>
<td>0.32</td>
<td></td>
<td>1-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.14</td>
<td>0.17</td>
<td></td>
<td>6-24</td>
<td>0.008</td>
</tr>
<tr>
<td>GmRPA1Bb</td>
<td>1 hps</td>
<td>-2.78</td>
<td>0.40</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-2.61</td>
<td>0.81</td>
<td></td>
<td>1-24</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>0.68</td>
<td>0.39</td>
<td></td>
<td>6-24</td>
<td>0.001</td>
</tr>
<tr>
<td>GmRPA2a</td>
<td>1 hps</td>
<td>0.46</td>
<td>0.07</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-0.13</td>
<td>0.11</td>
<td></td>
<td>1-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.67</td>
<td>0.27</td>
<td></td>
<td>6-24</td>
<td>0.09</td>
</tr>
<tr>
<td>GmRPA2b</td>
<td>1 hps</td>
<td>-2.86</td>
<td>0.47</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-2.54</td>
<td>0.87</td>
<td></td>
<td>1-24</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>0.85</td>
<td>0.49</td>
<td></td>
<td>6-24</td>
<td>0.002</td>
</tr>
<tr>
<td>GmRPA2c</td>
<td>1 hps</td>
<td>-2.65</td>
<td>0.35</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-2.15</td>
<td>0.80</td>
<td></td>
<td>1-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>1.05</td>
<td>0.31</td>
<td></td>
<td>6-24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GmRPA2d</td>
<td>1 hps</td>
<td>-1.34</td>
<td>0.22</td>
<td>0.03</td>
<td>1-6</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-0.78</td>
<td>0.48</td>
<td></td>
<td>1-24</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.06</td>
<td>0.20</td>
<td></td>
<td>6-24</td>
<td>0.27</td>
</tr>
<tr>
<td>GmRPA2e</td>
<td>1 hps</td>
<td>-0.34</td>
<td>0.29</td>
<td>0.11</td>
<td>1-6</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-0.19</td>
<td>0.28</td>
<td></td>
<td>1-24</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>0.43</td>
<td>0.20</td>
<td></td>
<td>6-24</td>
<td>0.23</td>
</tr>
</tbody>
</table>
### Table S2. (continued)

<table>
<thead>
<tr>
<th>Gene of Interest ID</th>
<th>Time Point (Hours Post Stress)</th>
<th>Relative Gene Expression</th>
<th>Standard Error</th>
<th>p-value (ANOVA)</th>
<th>Pairwise Comparison</th>
<th>p-value (Tukey's Honestly Significant Difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GmRPA3b</td>
<td>1 hps</td>
<td>0.24</td>
<td>0.14</td>
<td>0.04</td>
<td>1-6</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>0.03</td>
<td>0.17</td>
<td></td>
<td>1-24</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>-0.50</td>
<td>0.26</td>
<td></td>
<td>6-24</td>
<td>0.16</td>
</tr>
<tr>
<td>GmRPA3c</td>
<td>1 hps</td>
<td>-2.17</td>
<td>0.40</td>
<td>0.002</td>
<td>1-6</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-2.40</td>
<td>0.89</td>
<td></td>
<td>1-24</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>0.69</td>
<td>0.39</td>
<td></td>
<td>6-24</td>
<td>0.004</td>
</tr>
<tr>
<td>GmRPA3d</td>
<td>1 hps</td>
<td>-2.92</td>
<td>0.40</td>
<td>&lt;0.001</td>
<td>1-6</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>6 hps</td>
<td>-3.31</td>
<td>1.11</td>
<td></td>
<td>1-24</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>24 hps</td>
<td>1.26</td>
<td>0.39</td>
<td></td>
<td>6-24</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>