RAONI GWINNER

CONTRIBUTION TO THE DIVERSITY OF SOYBEAN GERMPLASM WITH INSIGHTS ON PATHOGENICITY FACTORS DURING THE INTERACTION WITH *Botrytis cinerea*

LAVRAS – MG
2017

RAONI GWINNER

CONTRIBUTION TO THE DIVERSITY OF SOYBEAN GERMPLASM WITH INSIGHTS ON PATHOGENICITY FACTORS DURING THE INTERACTION WITH *Botrytis cinerea*

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agronomia/Fitotecnia, área de concentração em Produção Vegetal, para a obtenção do título de Doutor.

Prof. Dr. Moacir Pasqual
Orientador

LAVRAS - MG
2017

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Gwinner, Raoni .
Contribution to the diversity of soybean germplasm withinsights on pathogenicity factors during the interaction with Botrytis cinerea / Raoni Gwinner. - 2017.
56 p.

Orientador(a): Moacir Pasqual.
.
Tese (doutorado) - Universidade Federal de Lavras, 2017.
Bibliografia.

1. soybean. 2. plant pathogen interaction. 3. genetic diversity. I. Pasqual, Moacir . . II. Título.

RAONI GWINNER

CONTRIBUTION TO THE DIVERSITY OF SOYBEAN GERMPLASM WITH INSIGHTS ON PATHOGENICITY FACTORS DURING THE INTERACTION WITH *Botrytis cinerea*

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agronomia/Fitotecnia, área de concentração em Produção Vegetal, para a obtenção do título de Doutor.

APROVADA em 21 de fevereiro de 2017.

Dra. Ester Alice Ferreira	EPAMIG
Dr. Ludwing Heinrich Pfenning	UFLA
Dr. Samuel Pereira de Carvalho	UFLA
Dr. Welison Andrade Pereira	UFLA

Prof. Dr. Moacir Pasqual
Orientador

LAVRAS - MG
2017

A Deus, por me fazer superar todas as dificuldades e persistir.

AGRADEÇO

Ao meu orientador, Moacir Pasqual, pela liberdade e confiança depositada. Ao amigo pesquisador, Tesfahun, pelo exemplo profissional, e a todos os colaboradores que com conduta ética e profissional fizeram do ambiente de trabalho um local propício ao crescimento e desenvolvimento pessoal.

OFEREÇO

À minha mãe Lorna, ao meu irmão Pedro, à minha avó Nellie, à tia Vera e aos meus padrinhos, Silvio, Dora, Dirceu e Maria José, pelo apoio e incentivo.

DEDICO

AGRADECIMENTOS

Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), à Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e à FAPEMIG (Fundação de Amparo a Pesquisa do Estado de Minas Gerais), pelo apoio financeiro que possibilitou a realização deste trabalho.

Ao DAG (Departamento de Agricultura) e à Universidade Federal de Lavras, pela oportunidade de cursar uma pós-graduação.

À minha mãe Lorna, ao meu irmão Pedro, à minha avó Nellie, aos tios Vera, Stanley e Kris, aos primos Joey e Becca, e a todos os membros da família Gwinner, que apesar de longe sempre torcem pelo meu sucesso.

Aos amigos da UC, Davis, Baohua, Gongjun, Nicole, Josh, Rachel, Chris, Roberta e Rose.

Ao meu Orientador Moacir Pasqual pela confiança depositada.

Ao meu orientador no exterior, Daniel Kliebenstein.

Aos amigos do Laboratório de Genética Molecular, Lamartine, Gisele, Renato, Antonio, Luciana, Fernanda e Dayane.

Ao Professor João Bosco, pelo exemplo de conduta no meio acadêmico.

Aos grandes amigos, Tesfahun Setotaw e Filipe “Batata”, pelo enorme apoio.

Aos laboratoristas Vantuil e Claret, pela amizade.

À Marli, pela competência e rapidez nas muitas vezes em que me ajudou e me instruiu.

Aos membros do setor de grandes culturas, em especial ao Antônio, Everton e Alan, pelo auxílio nas atividades.

Aos meus padrinhos Silvio, Dora, Dirceu e Maria José.

A todos aqueles que de alguma forma ajudaram na realização deste trabalho

A todos os companheiros do curso de pós graduação em Agronomia/Fitotecnia.

MUITO OBRIGADO!

RESUMO GERAL

A soja (*Glycine max*) é uma espécie de fundamental importância para a economia brasileira, e seu cultivo vem se expandindo expressivamente ao longo das ultimas décadas. A importanção de tal *commodity* representa uma parcela expressiva da balança comercial, e seu uso está disseminado por diversos segmentos industriais. No presente trabalho, foram abordados dois aspectos distintos, diretamente relacionados a esta espécie, a interação patógeno hospedeiro utilizando um fungo necrotrofico (*Botrytis cinerea*), e a diversidade genética utilizando um grupo de genótipos de uso comum. O estudo da interação patógeno hospedeiro foi realizado com o objetivo de se obter uma melhor compreensão do efeito da domesticação da soja sobre a virulência do patógeno, enquanto a avaliação da diversidade genética teve como objetivo principal, determinar a amplitude da base genética de materiais utilizados para hibridação. No estudo de interação patógeno hospedeiro, foram utilizados 12 genótipos de soja derivados do 'NAM Population Project' e 98 isolados de *B. cinérea*, em um ensaio de folhas destacadas. Nesse trabalho, houve variabilidade significativa para tamanho de lesão entre grupos de diferentes níveis de domesticação. Alem disso, foram detectadas regiões no genoma do patógeno, associadas com a virulência neste hospedeiro. A diversidade genética foi estudada com a utilização de 35 marcadores SSR ('Simple sequence repeat'), em 77 genótipos de soja. De forma geral, o germoplasma utilizado apresentou baixa diversidade genética, o que representa um impecilho à progressão dos ganhos de seleção com o melhoramento.

Palavras-chave: Soja. *Botrytis cinerea.* Marcadores moleculares. Diversidade genética.

GENERAL ABSTRACT

Soybean (*Glycine max*) is a specie that has fundamental importance to the Brazilian economy and the production area has been largely increased during the past few decades. The importation of such commodity represents a large share of the trade balance, and its uses are spreaded for several industrial segments. In this study, two distincts aspects were approached related to this species, the plant-pathogen interaction using a necrotrophic fungi (*Botrytis cinerea*) and the genetic diversity in a group of commonly used soybean genotypes. The plant pathogen interaction study was done to reach a better understanding of the effect of soybean domestication on the pathogen virulence. The genetic diversity evaluation had as a main goal determine the range of the genetic base in a soybean breeding set. The plant pathogen interaction study was performed using a detached leaves assay with12 genotypes from the NAM Population Project and 98 isolates of *B. cinerea*. In this assay, significant variability for lesion size was found between domestication groups. Futhermore, pathogen genome regions associated with virulence were detected. The genetic diversity was done with 35 SSR markers (Simple sequence repeat) with 77 soybean genotypes. Overall, the germplasm showed low genetic diversity what represents a obstacle to the progression of selection gains through breeding.

Keywords: Soybean. *Botrytis cinerea*. Molecular marker. Genetic diversity.

SUMÁRIO

1 THEORETICAL FRAMEWORK

Plant pathogens are the reason for significant yield and financial losses in agriculture however, aside from gene-for-gene related resistance mechanisms; there is still a considerable lack of information about plant/pathogen interactions. What we measure as the disease outcome of a plant/pathogen interaction is the complex interaction of the biological mechanisms encoded by the two interacting organism's genomes. For the plant, this can include several layers of defense mechanisms, ranging from innate immune responses, through defense responses activated by classes or species of pathogens, to defense strategies effective against single pathogen isolates. Correspondingly, these disease mechanisms can be present across multiple plant species or be specific even for conserved topics such as non-host resistance (CLAY et al. 2009; BEDNAREK et al. 2009; FAN et al. 2011) As both the plant defense responses and the pathogens virulence mechanisms frequently show significant genetic variation that is under directional and/or balancing selection (BAKKER et al. 2008; BAKKER et al. 2006; KORVES; BERGELSEN, 2004; TIAN et al. 2003; MAURÍCIO et al. 2003), it is more appropriate to describe the plant/pathogen interaction as an interaction of two species genomes and all the genetic variation contained within these two species when discussing specialist pathogens. This problem grows in scale when discussing generalists such as *Botrytis cinerea* that can infect a wide range of different plants. In this instance, it may be more appropriate to view the evolution of plant/generalist pathogen systems as an interaction of the pathogen with numerous different plant genomes. In spite of the observed importance of genetic variation in both host and pathogen there have been few studies attempting to simultaneously assess how the variation across plant families interacts with a pathogens genetic variation to control the virulence outcome.

The impact of genetic variation in host and pathogen has been most often studied using the gene-for-gene system whereby direct or indirect recognition of pathogen avr gene products (effectors) by plant proteins activates a salicylic acid-dependent signaling cascade within the plant leading to localized programmed cell death. However, these gene-for-gene systems are frequently unique to individual specialist pathogens and they typically involve variation in single genes that produce qualitative resistance. Further, these gene-for-gene systems are frequently limited to biotrophic pathogens and as such do not provide information about necrotrophic pathogens which may not utilize the same variable plant genes to gain entry into the plant (KLIEBENSTEIN; ROWE 2008).

In contrast to qualitative resistance genes in specialized plant/biotroph interactions, no naturally occurring qualitative resistance genes have been described for generalist necrotrophic plant pathogens such as *Botrytis cinerea*, or *Sclerotinia sclerotorium* that can infect nearly every plant that has been tested including all eudicots. For *Botrytis cinerea*, this virulence capacity extends into the gymnosperms and bryophytes (WILLIAMSON et al. 2007). Frequent studies on the genetic variation of *Botrytis cinerea* have shown that there is no specialization within the fungus leading to a situation where isolates obtained from tomato are often better at infecting other eudicots such as Brassica than isolates obtained from other Brassicas (STAATS et al. 2007; ROWE; KLIEBENSTEIN, 2007; MYRESIOSTIS et al. 2007). Additionally, *Botrytis cinerea* while never being an epidemic disease of plants leading to absolute crop loss, it is a frequent endemic disease on nearly all fruits and vegetables causing crop loss both pre and post-harvest. This ability of a single pathogen to infect nearly all eudicots, both wild and domestic, presents a unique ability to conduct comparative quantitative studies across the eudicots of defense mechanisms to understand a novel aspect of plant/pathogen interactions that is frequently overlooked.

One possible manner in which generalist necrotrophic pathogen diversity may differ from specialist biotrophic pathogens is that the genomic variation in the pathogen may not focus on gene-for-gene interactions but instead may rely upon network-for-network interactions that shift from eudicot to eudicot or even genotype to genotype within a species.
Combining whole genomes and phenotypes: Next generation sequencing has enabled a new level of whole genome diversity analyses. However, these studies often solely focus on describing the genomic variation when investigating plant/pathogen interactions without measuring the underlying biology in a quantitative fashion.

GWA identifies association between phenotypes and genotypes, at a genome-wide level, using ‘unrelated’ individuals that have been simultaneously genotyped and phenotyped (HIRSCHHORN; DALY, 2005; WEIGEL; NORDBORG, 2005; NORDBORG; WEIGEL, 2008). Genetic recombination across generations leads to a decay of linkage disequilibrium (LD), or apparent genetic linkage, between neighboring polymorphisms such that polymorphisms separated by hundreds to thousands of bases are effectively inherited independently (NORDBORG; WEIGEL, 2008; KIM et al., 2007). GWA-mapping aims to identify polymorphisms associated with the quantitative traits of interest and its potential has been demonstrated in genome wide analyses (CHAN et al., 2010; ZHAO et al., 2007; EASTON et al., 2007; GHAZALPOUR et al., 2008).

REFERENCES

BAKKER, E. G.; TOOMAJIAN, C.; KREITMAN, M.; BERGELSON, J. A genome-wide survey of R gene polymorphisms in Arabidopsis. **The Plant Cell**, v. 18, n. 8, p. 1803-1818, 2006.

BAKKER, E. G.; TRAW, M. B.; TOOMAJIAN, C.; KREITMAN, M.; BERGELSON, J. Low levels of polymorphism in genes that control the activation of defense response in Arabidopsis thaliana. **Genetics**, v. 178, n. 4, p. 2031-2043, 2008.

BEDNAREK, P.; PIŚLEWSKA-BEDNAREK, M.; SVATOŠ, A.; SCHNEIDER, B.; DOUBSKÝ, J.; MANSUROVA, M.; HUMPHRY, M.; CONSONNI, C.; PANSTRUGA, R.; SANCHEZ-VALLET, A. A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. **Science**, v. 323, n. 5910, p. 101-106, 2009.

CHAN, E. K.; ROWE, H. C.; KLIEBENSTEIN, D. J. Understanding the evolution of defense metabolites in Arabidopsis thaliana using genome-wide association mapping. **Genetics**, v. 185, n. 3, p. 991-1007, 2010.

CLAY, N. K.; ADIO, A. M.; DENOUX, C.; JANDER, G.; AUSUBEL, F. M. Glucosinolate metabolites required for an Arabidopsis innate immune response. **Science**, v. 323, n. 5910, p. 95-101, 2009.

EASTON, D. F.; POOLEY, K. A.; DUNNING, A. M.; PHAROAH, P. D.; THOMPSON, D.; BALLINGER, D. G.; STRUEWING, J. P.; MORRISON, J.; FIELD, H.; LUBEN, R. Genome-wide association study identifies novel breast cancer susceptibility loci. **Nature**, v. 447, n. 7148, p. 1087-1093, 2007.

FAN, J.; CROOKS, C.; CREISSEN, G.; HILL, L.; FAIRHURST, S.; DOERNER, P.; LAMB, C. Pseudomonas sax genes overcome aliphatic isothiocyanate–mediated non-host resistance in arabidopsis. **Science**, v. 331, n. 6021, p. 1185-1188, 2011.

GHAZALPOUR, A.; DOSS, S.; KANG, H.; FARBER, C.; WEN, P.-Z.; BROZELL, A.; CASTELLANOS, R.; ESKIN, E.; SMITH, D. J.; DRAKE, T. A. High-resolution mapping of gene expression using association in an outbred mouse stock. **PLoS Genet**, v. 4, n. 8, p. e1000149, 2008.

HIRSCHHORN, J. N.; DALY, M. J. Genome-wide association studies for common diseases and complex traits. **Nature Reviews Genetics**, v. 6, n. 2, p. 95-108, 2005.

KIM, S.; PLAGNOL, V.; HU, T. T.; TOOMAJIAN, C.; CLARK, R. M.; OSSOWSKI, S.; ECKER, J. R.; WEIGEL, D.; NORDBORG, M. Recombination and linkage disequilibrium in Arabidopsis thaliana. **Nature genetics**, v. 39, n. 9, p. 1151-1155, 2007.

KLIEBENSTEIN, D. J.; ROWE, H. C. Ecological costs of biotrophic versus necrotrophic pathogen resistance, the hypersensitive response and signal transduction. **Plant Science**, v. 174, n. 6, p. 551-556, 2008.

KORVES, T.; BERGELSON, J. A novel cost of R gene resistance in the presence of disease. **The American Naturalist**, v. 163, n. 4, p. 489-504, 2004.

MAURICIO, R.; STAHL, E. A.; KORVES, T.; TIAN, D.; KREITMAN, M.; BERGELSON, J. Natural selection for polymorphism in the disease resistance gene Rps2 of Arabidopsis thaliana. **Genetics**, v. 163, n. 2, p. 735-746, 2003.

MYRESIOTIS, C.; KARAOGLANIDIS, G.; TZAVELLA-KLONARI, K. Resistance of Botrytis cinerea isolates from vegetable crops to anilinopyrimidine, phenylpyrrole, hydroxyanilide, benzimidazole, and dicarboximide fungicides. **Plant Disease**, v. 91, n. 4, p. 407-413, 2007.

NORDBORG, M.; WEIGEL, D. Next-generation genetics in plants. **Nature**, v. 456, n. 7223, p. 720-723, 2008.

ROWE, H. C.; KLIEBENSTEIN, D. J. Elevated genetic variation within virulence-associated Botrytis cinerea polygalacturonase loci. **Molecular Plant-Microbe Interactions**, v. 20, n. 9, p. 1126-1137, 2007.

STAATS, M.; VAN BAARLEN, P.; SCHOUTEN, A.; VAN KAN, J. A.; BAKKER, F. T. Positive selection in phytotoxic protein-encoding genes of Botrytis species. **Fungal Genetics and Biology**, v. 44, n. 1, p. 52-63, 2007.

TIAN, D.; TRAW, M.; CHEN, J.; KREITMAN, M.; BERGELSON, J. Fitness costs of R-gene-mediated resistance in Arabidopsis thaliana. **Nature**, v. 423, n. 6935, p. 74-77, 2003.

WEIGEL, D.; NORDBORG, M. Natural variation in Arabidopsis. How do we find the causal genes? **Plant physiology**, v. 138, n. 2, p. 567-568, 2005.

WILLIAMSON, B.; TUDZYNSKI, B.; TUDZYNSKI, P.; VAN KAN, J. A. Botrytis cinerea: the cause of grey mould disease. **Molecular plant pathology**, v. 8, n. 5, p. 561-580, 2007.

ZHAO, K.; ARANZANA, M. J.; KIM, S.; LISTER, C.; SHINDO, C.; TANG, C.; TOOMAJIAN, C.; ZHENG, H.; DEAN, C.; MARJORAM, P. An Arabidopsis example of association mapping in structured samples. **PLoS Genet**, v. 3, n. 1, p. 4, 2007.

CHAPTER 1 GENETIC VARIATION IN THE Botrytis cinerea-SOYBEAN PATHOSYSTEM HIGHLIGHTS NEW PATHOGEN VIRULENCE GENES AND AN ALTERNATIVE ROLE FOR OXALATE IN VIRULENCE

RESUMO

O patossistema soja-*Botrytis cinerea* é um importante modelo de interação para estudo e elucidação de novas fontes de resistência, assim como genes associados à agressividade do patógeno. Além disso, analisar o papel de compostos associados com a patogenicidade pode auxiliar melhoristas no desenvolvimento de novas estratégias, que visam superação de problemas proporcionados por doenças em plantas. No presente trabalho, apresenta-se a variação na patogenicidade e produção de ácido oxálico em um conjunto de isolados de *B. cinerea.* O patossistema foi criado pela combinação de dois grupos de genótipos de soja, sendo um, domesticado, e outro "landrace", com o grupo de isolados de *B. cinerea,* e o efeito da domesticação foi analisado no patossistema soja-*B. cinerea.* Foi realizado um estudo de associação genômica ampla (GWAS), com objetivo de detectar regiões do genoma do patógeno associado à intensidade de infecção, e para entendimento de como tais regiões variam de acordo com o grupo de domesticação do hospedeiro. As principais conclusões foram de que a domesticação do hospedeiro afeta a permormance de infecção do patógeno, e os dados desse estudo, sugerem uma correlação negativa entre produção de acido oxálico e tamanho da lesão promovida pelo patógeno. Além disso, destacam-se regiões do genoma do patógeno, que possivelmente estão associadas com a produção de ácido oxálico e capacidade de infecção.

Palavras-chave: Interação planta-patógeno. Soja. Botrytis cinerea. Ácido oxálico.

ABSTRACT

The *Botrytis cinerea*- soybean pathosystem is an important plant pathogen interaction model for general studies and to reveal novel sources of resistance and genes associated with pathogen aggressiveness as well. Moreover, analyze the role of compounds associated with pathogenicity can be helpful for plant breeders to develop new strategies aiming to overcome common plant diseases. In this study, we present the variation of pathogenicity and oxalic acid (OA) production in a *B. cinerea* isolate set. The pathosystem was created combining domesticated and landrace soybean genotypes with the *B. cinerea* set, and analyze the effect of host domestication in the *B. cinerea*- soybean pathosystem. A GWAS (Genome Wild Association Study) was performed aiming to detect regions into the pathogen genome associated with the infection performance and to understand how these regions change under a different host domestication group. Our main conclusions were that the host domestication affect pathogen performance and our data suggest a negative correlation between OA and pathogen lesion size. Despite that, we highlight candidate genome regions associated with oxalic acid production and pathogen aggressiveness.

Keywords: Plant-pathogen interaction. Soybean. *Botrytis cinerea*. Oxalic acid.

1 INTRODUCTION

Soybean production have been expanded largely worldwide during the last few decades (BOEREMA et al., 2016) mostly due to its role as protein and oil source (CLEMENTE; CAHOON, 2009). The substantial economic value of this crop favored the intensive domestication process during the last decades, promoting a human-made bottleneck (ZHOU et al., 2015). The genetic diversity is one factor driving plant-pathogen co-evolution, and the effect of domestication on plant-generalist pathogen interaction have not been completely clarified. SALEH et al. (2014) and more recently, DE GRACIA et al. (2015) proposed the term 'pestification' that refers to evolutionary changes in the pathogen according to its adaptation to the host domestication process, leading to the improvement of pathogen virulence. Although landrace soybean genotypes had pass through a domestication process, they are genetically closer to soybean wild relative *Glycine soja,* presenting higher diversity index and unique alleles when compared to modern elite lines (HYTEN et al., 2006).

Plant pathogens are responsible for significant yield reduction in agriculture. In soybean crops, plant pathogens can lead to losses of billions of dollars globally (WRATHER et al., 2001; KOENNING et al., 2010). In this scenario, generalist necrotrophic pathogens represent a considerable share for the yield losses, especially *Sclerotinia sclerotiorum*. Despite, *Botrytis cinerea* does not represent a major problem in soybean crops, it is a close relative of *S. sclerotiorum* and can be considered a good model for plant-necrotrophic pathogen interaction studies, since the infection process and genes associated with pathogenicity are similar in these necrotrophic fungi (AMSELEM et al., 2011). *Botrytis cinerea* is a broad-host necrotrophic fungus capable to infect more than 200 plant species, mostly broadleaf species, causing severe losses in agriculture. Its capability to infect several plant structures and stages of plant development, make it difficult to control (KUNZ et al., 2006). During the infection process several compounds are secreted, triggering the plant defense response. The induced programed cell death (PCD) stimulates the establishment of the fungus and its crucial for the host colonization (KARS; VAN KAN, 2007).

Botrytis cinerea secrete oxalic acid during the infection process, this compound can be framed as a key factor associated to the pathogenicity, acting by many ways as: lowering the pH facilitating the activity of endopolygalacturonases, chelating Ca2+ from Ca-pectin complexes contributing for activity of cell-wall-degrading enzymes, suppressing the host defense oxidative burst, impacting the stomatal functioning by stimulating K+ uptake into

guard cells and inducing programmed cell death (GUIMARAES; STOTZ, 2004; WILLIAMSON et al, 2007; DAVIDSON et al. 2016). LIANG et al. (2015) showed that oxalate deficient necrotrophic mutants produce limited lesion in soybean. HOWEVER; XU et al. (2015) demonstrated that oxalic acid is not a primary pathogenic factor, indicating that oxalate deficient mutants were able to cause disease under low pH and suggested that oxalic acid is not necessary for host infection since the pathogen can use other mechanism to lower the host tissue pH. In this case, *oah* mutants were able to acidify the host tissue by fumaric acid production and were less efficient acidifying the media (XU et al., 2015). Although the description of oxalic acid as a primary pathogenic factor is being questionable, it plays a key role in plant-pathogen interaction (KIM et al., 2008; WILLIAMS et al., 2011; NAKAJIMA; AKUTSU, 2014; LIANG et al., 2015). Transgenic soybean plants expressing OA degrading enzymes show more resistance to necrotrophic disease development (DONALDSON et al., 2001; COBER et al., 2003; Cunha et al., 2010; DAVIDSON et al., 2016) and enhance plant mechanism against OA production can be considered an alternative strategy for disease control management.

Understand the variability of virulence in generalist pathogens taking hosts with distinct level of domestication would help to clarify the evolution of virulence mechanisms according to the domestication process. Besides, the characterization of different levels of resistance can be helpful for germplasm management. Here, we investigated the pathogenicity of *B. cinerea* in two groups of soybean germplasm with distinct degrees of domestication and evaluate the production of oxalic acid in a collection of *B. cinerea* isolates making a link of their production with virulence.

Furthermore, by association analyses we demonstrate new genes related to pathogenicity. Conducting a GWAS we aim to detect *B. cinerea* genome regions associated with both pathogenicity and OA production. By analyzing the different patterns in the GWAS study according to the host we wish to find how this pathogen vary its infection mechanisms depending on the genetic constitution of the host.

2 MATERIAL AND METHODS

2.1 Botrytis growth

Oxalic acid and lesion development was measured in 98 isolates of *Botrytis cinerea*. Isolates were maintained as conidial suspensions in 30% glycerol for long term storage at -80°C. For regrowth, spore solutions were diluted to 10% in 50% filter-sterilized grape juice, then inoculated onto 39g/L potato dextrose agar (PDA) media. Isolates were grown at 25°C in 12h light, and propagated every two weeks.

2.2 Soybean genetic resources

We obtained seeds for 12 soybean genotypes selected from the soyNAM Project. These include a diverse sample of six landrace genotypes and six modern varieties of *Glycine max* as well (TABLE 1). All genotypes were planted in greenhouse trays in a completely randomized block design under controlled conditions at UC Davis. Plants were grown under day/night temperatures at 25°C/18°C in 4" pots filled with standard potting soil (Sunshine mix #1, Sun Gro Horticulture). Plants were watered regularly and staked upright. At week 2 the plants were watered every two days with added nutrient solution (0.5% N-P-K fertilizer in a 2-1- 2 ratio; Grow More 4-18-38). Plants were used for detached leaf virulence assays 4 weeks after sowing.

2.3 Detached leaf assay

To study how lesion formation is influenced by the genetic interaction of host (soybean) and pathogen (*B. cinerea*), we infected 12 diverse soybean varieties with 96 Botrytis isolates. We used a randomized block design with three replicates of each plant by isolate pair per experiment. This whole experiment was repeated once with new randomization leading to at least six measurements per soybean genotype x *B. cinerea* genotype for a total of ~ 6,900 lesion measurements. We randomly sampled 5 adult leaves per plant, and 2 leaflet pairs per leaf for inoculation. For the statistical model, we kept track of the

plant source for each leaf and noted the apical vs. basal leaflet pairs. Leaflets were placed on 1% phytoagar flats with humidity domes on top.

Spores were collected from mature (2 weeks old) Botrytis cultures, and diluted to 10 spores/ µl in 50% filter-sterilized grape juice. 4 µl droplets of spore suspensions were inoculated onto detached leaves at room temperature with 24h light. Control leaves were mock-inoculated with 4 µl of grape juice without spores.

We photographed all leaflets infections at 24, 48, and 72 hours post inoculation for downstream image analysis.

2.4 Automated image analysis

We measured lesion areas using the EBImage and CRImage packages (PAU et al., 2010; FAILMEZGER et al., 2010) in the R statistical environment (R DEVELOPMENT CORE TEAM, 2009). Leaflets were identified as objects with green hue, and lesions were identified as low-saturation objects within leaves. Images masks were generated for both the leaf and lesion and then manually refined by a technician to ensure proper object calling. The area of the leaflets and lesions were then automatically measured as pixels per lesion and converted to area using a 1 cm control object within each image.

2.5 Oxalic acid content

The oxalic acid content of each *B. cinerea* isolate was determined according to DUAN et al. (2014) with some modifications. Mycelial plugs (1 cm diameter) from the margins of 7 days old hyphal colonies were placed in 50 ml falcon tubes (two per tube) containing 20 ml of PDB (pH 6.0). After 7 days the tubes were centrifuged at 2500 rpm for 15 minutes. The oxalic acid content within the supernatant was measured using the color intensity at 510 nm after the reaction with Iron (III) - sulfosalicylic acid solution. Absolute oxalic acid concentration was determined in relation to the below standard curve. There were four replicates per *B. cinerea* isolate, and the experiment was performed twice. A standard curve for oxalate measurement was prepared using 2 mL of FeCl3 solution (0.5 mg/mL), 20 mL of HCl–KCl buffer solution (KCl 50 mM, pH = 2), and 1.2 mL of sulfosalicylic acid solution (5 mg/mL). Different volumes (0, 0.1, 0.2, 0.4, or 0.6 mL) of sodium oxalate solution (2 mg/mL) were added, and volume to was increased to 25 mL with double distilled water. The tubes were vortex-stirred

for 5 s and incubated for 30 min at 22 °C. Absorbance was measured at 510 nm with a spectrophotometer.

2.6 Data analysis

All statistical analysis was performed using R statistical environment (R DEVELOPMENT CORE TEAM, 2009). For he detached leaf assay an analyze of variance was performed (ANOVA) using a linear mixed model implemented in the "lme4" package (Bates et al. 2014), using as fixed effects: *B. cnerea* isolates, plant genotype group (Landrace or Elite Cultivars) and soybean genotypes nested within group. As a random effect we analyze: experiment, replicate, image, plant number and the greenhouse tray, each term nested according to the model below, where the term "/" represent nesting.

Lesion size = isolate + group + group/soybean genotype + isolate * group + isolate * group/soybean genotype + (1|experiment) + (1|experiment/replicate) + (1| experiment/replicate /Flat) + (1|experiment/replicate /plantnumber) + (1|experiment/GreenHouseTray)

ANOVA was performed individually for each isolate examining the fixed effects of group of domestication and soybean genotypes (nested in groups). For the oxalate production data, analyze of covariance was performed (ANCOVA) using a linear model where the following sources of variation were considered in the model: oxalate production, plant genotype group (Landrace or Elite Cultivars), soybean genotype and interaction of oxalate production with group and soybean genotypes. The linear model is exemplified below, the term "/" represent nesting and the term "*" represent interaction.

Lesion size = oxalate production + group + group/ soybean genotype + group* oxalate production + group/ soybean genotype * oxalate production

The Pearson's correlation coefficient was estimated between lesion size and oxalate production considering the mean lesion size value for each plant-pathogen combination and the mean oxalate production value per isolate.

2.7 Association mapping

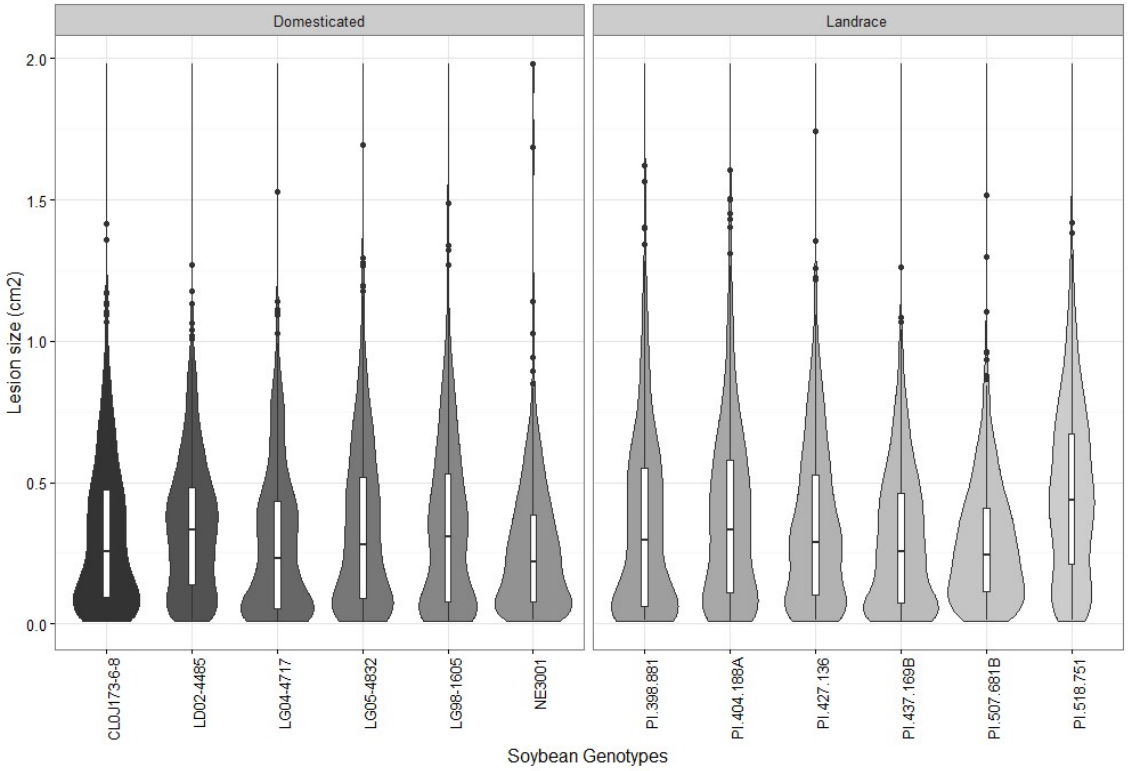
Association mapping was performed using the bigRR package (SHEN et al., 2013) in the R statistical environment. This method uses ridge regression to model the effects of all polymorphisms in a single model, treating each SNP as a random effect and introducing a bias to the regression estimates to reduce standard error. Thus, each polymorphism is assigned a heteroscedastic effect size (HEM), rather than a p-value, which is difficult to determine for random variables. Instead, a significant effect threshold value is delineated by permuting the phenotype data as it corresponds to the polymorphism data 1,000 times, and taking the 95 th and the 99 th quantiles. A gene was considered to be associated with a phenotype when two or more SNPs within the coding region have an effect size greater than the 95 th percentile threshold.

3 RESULTS

3.1 Detached leaves assay

A wide range of virulence were exhibited by the pathogen. Soybean genotypes presented different levels of susceptibility to *B. cinerea*. The overall mean lesion size by isolate varied from 0.045 cm^2 (KGB1) to 0.602 cm^2 (Fd1), and the overall mean lesion size in each soybean genotype varied from 0.257 cm^2 (NE3001) to 0.465 cm^2 (PI.518.751) (FIGURE 1).

Figure 1 – Violin plot of lesion size promoted by *B. cinerea* isolates in domesticated and landrace soybean genotypes



The random effects, experiment and replicate presented non-significant values (TABLE 2) suggesting that the environmental conditions were appropriate for the assay. The plant number also presented non-significant value showing the genetic integrity of the plant material.

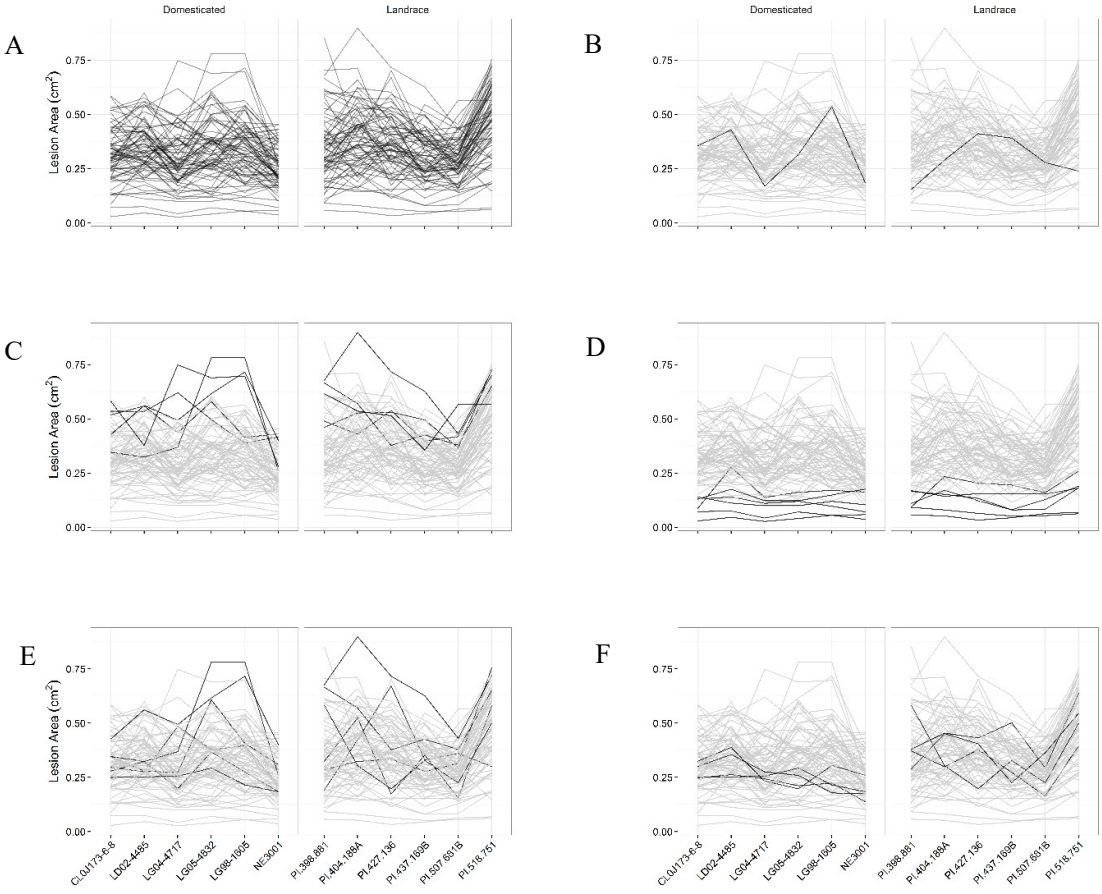
Table 2 – Analysis of Variance of Lesion Size

Fixed effects	Df	S.S.	F	p-value
Pathogen	95	78.163	17.3463	< 2.2e-16 ***
Group	1	3.108	65.5300	6.96e-16 ***
Plant	10	11.615	24.4886	< 2.2e-16 ***
Pathogen*Group	95	3.863	0.8572	0.8363
Plant *Pathogen	950	38.332	0.8507	0.9992
Random effects	Chi.DF	Chi.sq		p-value
Experiment	1	$3.81e^{-09}$		1
Replicate	2	$5.35e^{-11}$		1
Phyto-Agar Flat	3	$4.35e^{+02}$		<2e-16 ***
Plant Number	3	$5.50e^{-11}$		1
Greenhouse Tray	2	$7.56e^{+01}$		<2e-16 ***

Signif. codes: '***' 0.001

Among fixed effects, isolate, soybean genotype (nested in group) and soybean group, presented highly significant effect on lesion size. The isolate-soybean genotype (nested in group) interaction and isolate-group interaction were not significant (Table 2). Nevertheless, the isolate-genotype interaction present a considerable high Df, what could mask a significant result. However, it was possible to detect some variation in virulence patterns according to the host or host group (FIGURE 2)

Figure 2 – Lesion size pattern of each *B. cinerea* isolate across soybean genotypes where: A) All isolates; B) B05.10; C) 8 most aggressive isolates; D) 8 least aggressive isolates; E) Most responsive to host genetic variation and F) Most responsive to host domestication.



The non-significant interaction could possibly be a result from the pathogen generalist lifestyle leading to a lack of host specificity for different genotypes and levels of domestication. Through individual analysis, was possible to detect significant effect of host and host group in lesion size for a few isolates, indicating some host specificity (TABLE 3).

Table 3 - Isolates presenting significant p- values for Group and Plant Genotype

Source of Variation	Isolates	p-values
Group	94.4	0.01*
	1.04.21	0.009173**
	2.04.08	0.002752**
	Fd2	0.04578*
	Katie Tomato	0.04395*
Plant Genotype	94.4	0.02627*
	1.01.15	0.003513**
	1.01.6	0.011*
	1.02.04	0.03351*
	1.05.24	0.0365*
	Fd1	0.04205*

Signif. codes: '***' 0.001 ; '**' 0.01'*' ; 0.05'*'

The mean coefficient of variation (CV) of lesion size was similar between groups of domestication where the isolates presented a CV for lesion size of 75.24% in the domesticated group and 75.84 % for the Landrace group. Despite the similarity of CV values, many isolates presented divergent pattern of lesion variation, showing a lower CV in one group of domestication compared to the other one (Table 4), meaning that the isolates "behave" more as a generalist in one of the groups. Aiming to understand how some isolates change the generalist pattern in different groups we check the difference in lesion size variation (ΔCV). The isolates BPA1, 1.02.03, 1.03.23 and Gallo2 presented the most contrasting CV among domestication groups (TABLE 4). The isolates Geranium, "1.02.18" and "1.03.20" presented mean CV values above 100 % suggesting a non-generalist behavior (Table 4). In general, all isolates presented a considerable high CV.

Table 4 - Coeficient of Variation of Lesion size of each isolate according to each group of domestication. (continues…)

	CV (%)			
Isolates	Domesticated	Landrace	|Domesticated-Landrace|	Isolate mean
BPA1	81.54	49.07	32.47	65.31
1.02.03	66.34	97.70	31.36	82.02
1.03.23	100.39	70.95	29.44	85.67
Gallo2	83.82	112.98	29.16	98.40
2004	82.71	53.85	28.86	68.28
1.01.03	99.65	72.28	27.37	85.97
1.02.13	81.90	55.67	26.23	68.79
KernA2	100.38	75.56	24.81	87.97
1.03.02	87.69	111.20	23.51	99.45
1.03.16	72.74	96.20	23.46	84.47
1.04.02	78.46	101.23	22.76	89.85
Fd2	109.82	87.32	22.50	98.57
1.05.11	84.66	105.37	20.71	95.02
1.01.02	94.96	75.12	19.84	85.04
Navel	41.41	59.26	17.84	50.34
NobleRot	81.31	98.93	17.62	90.12
Triple7(T7)	74.39	90.75	16.36	82.57
Supersteak	68.77	84.53	15.75	76.65
1.01.15	59.36	74.31	14.95	66.84
1.04.03	74.86	60.65	14.21	67.76
2.04.08	69.73	55.64	14.09	62.69
1.05.04	63.13	76.95	13.82	70.04
1.01.01	63.53	77.34	13.81	70.44
94.4	56.64	68.85	12.21	62.75
KernB1	56.59	68.80	12.21	62.70
1.03.04	63.60	75.43	11.83	69.52
1.03.12	86.11	74.43	11.68	80.27
1.03.19	74.81	86.19	11.38	80.50
Pepper	65.68	76.90	11.22	71.29
1.04.21	62.31	51.34	10.97	56.83
1.05.24	86.55	75.60	10.95	81.08
1.03.18	79.51	69.20	10.31	74.36
1.04.19	85.95	75.90	10.05	80.93
1.04.17	71.31	81.25	9.93	76.28
1.02.17	70.97	61.17	9.81	66.07
FresaS.D.	75.78	66.53	9.24	71.16
1.02.05	74.56	65.51	9.05	70.04
2.04.04	65.67	74.69	9.01	70.18
Acacia	80.64	71.94	8.70	76.29
2.04.09	49.60	58.20	8.59	53.90
1.04.04	65.29	73.69	8.40	69.49

	CV (%)			
Isolates	Domesticated	Landrace	|Domesticated-Landrace|	Isolate mean
1.02.20	94.68	86.69	7.99	90.69
KGB1	71.61	63.69	7.92	67.65
Geranium	99.50	107.35	7.85	103.43
1.05.14	65.43	73.06	7.63	69.25
1.05.22	78.08	85.22	7.14	81.65
EsparatoFresa	83.42	76.30	7.12	79.86
MEAP6G	51.63	58.66	7.03	55.15
1.01.04	83.41	90.40	6.99	86.91
1.03.22	74.13	81.06	6.93	77.60
1.04.20	68.93	75.69	6.76	72.31
Grape	88.62	81.89	6.73	85.26
Gallo1	68.92	75.47	6.55	72.20
Ausubel	75.14	80.85	5.71	78.00
1.02.16	62.87	68.46	5.60	65.67
1.02.18	114.42	109.04	5.39	111.73
1.02.01	45.87	51.26	5.38	48.57
1.02.15	67.66	62.39	5.27	65.03
2.04.11	59.85	64.94	5.10	62.40
Apple404	65.92	70.99	5.07	68.46
2.04.20	62.93	58.10	4.83	60.52
1.04.12	85.29	80.62	4.67	82.96
UKrazz	79.39	84.04	4.65	81.72
2.04.18	50.83	55.34	4.51	53.09
1.02.02	87.57	83.31	4.26	85.44
PhiloMenlo	87.04	90.37	3.33	88.71
B05.10	84.94	88.17	3.23	86.56
2.04.12	50.76	47.57	3.20	49.17
BMM	59.27	62.35	3.08	60.81
94.1	74.62	71.62	3.00	73.12
KatieTomato	63.79	66.68	2.89	65.24
Triple3(T3)	84.41	81.55	2.86	82.98
2.04.21	90.79	88.21	2.57	89.50
1.02.04	82.27	80.03	2.24	81.15
1.03.20	109.31	111.12	1.81	110.22
2.04.17	61.98	60.41	1.57	61.20
1.05.17	79.84	78.35	1.49	79.10
Fresa525	61.87	60.41	1.47	61.14
2.04.03	73.49	72.16	1.32	72.83
1.04.01	81.84	83.01	1.17	82.43
Rose	50.75	51.84	1.10	51.30
KernB2	83.07	81.98	1.09	82.53
1.01.6	70.01	69.22	0.79	69.62
1.04.25	92.45	91.70	0.75	92.08
1.04.15	92.99	93.67	0.68	93.33
KGB2	75.09	74.47	0.62	74.78
Fd1	50.97	50.38	0.59	50.68

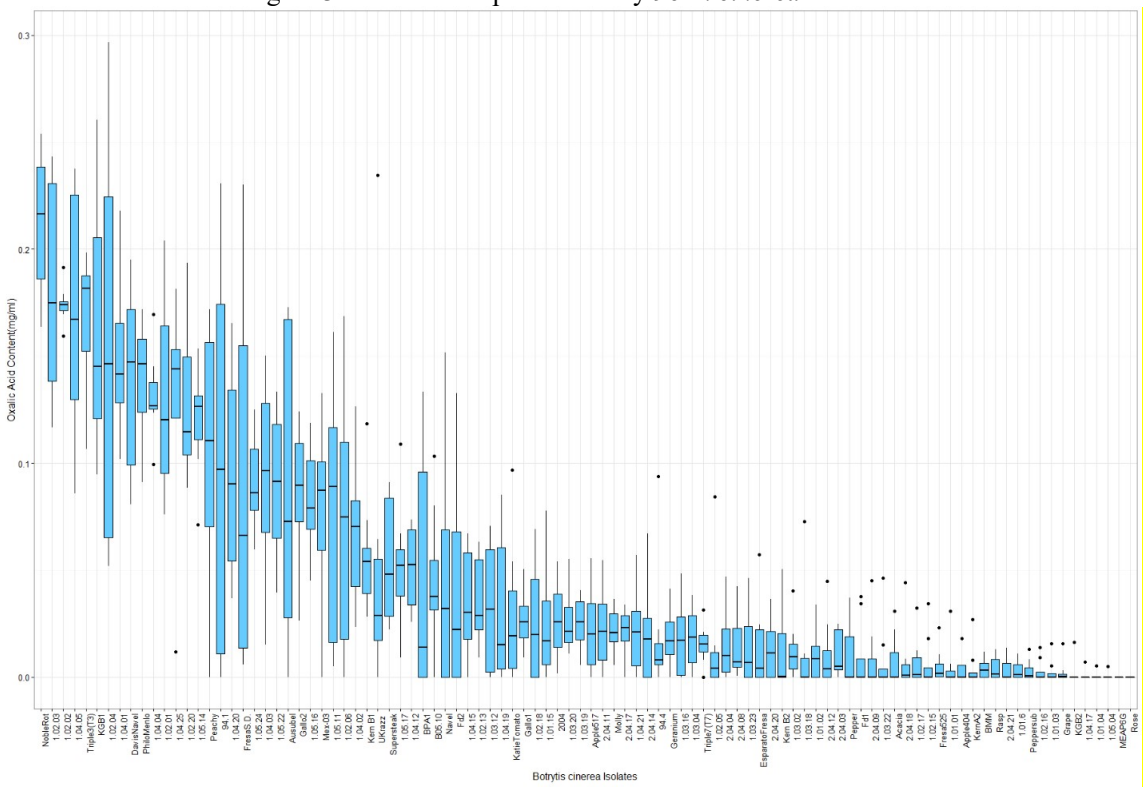
	CV (%)			
Isolates	Domesticated	Landrace	|Domesticated-Landrace|	Isolate mean
Rasp	70.49	69.94	0.54	70.22
1.02.06	90.73	91.23	0.50	90.98
Apple517	83.35	82.87	0.48	83.11
1.04.05	82.25	82.62	0.37	82.44
Molly	85.23	85.58	0.34	85.41
1.05.16	61.38	61.22	0.16	61.30
2.04.14	59.96	60.10	0.14	60.03
DavisNavel	65.67	65.78	0.11	65.73
Peppersub	83.10	83.07	0.03	83.09
Average	75.24	75.84	9.14	

(conclusion)

3.2 Oxalic acid production

Oxalic acid (OA) is considered a key compound during the infection process by necrotrophic generalist pathogens by measuring it, we aim to understand how the secretion capability is related to virulence measuring the correlation between the secretion capacity and lesion size. The studied isolates presented distinct levels of oxalic acid, and significant differences among them were detected. The OA production varied from 0 mg/ml (MEAP6G and Rose) to 0.21 mg/ml (NobleRot) (FIGURE 3).

Figure 3 - Oxalic acid production by 96 *B. cinerea* isolates



The production by each isolate is exhibited in Figure 3. The analyses of covariance demonstrate a significant effect of OA on lesion size (TABLE 5).

Table 5 – Analysis of Covariance of Lesion Size

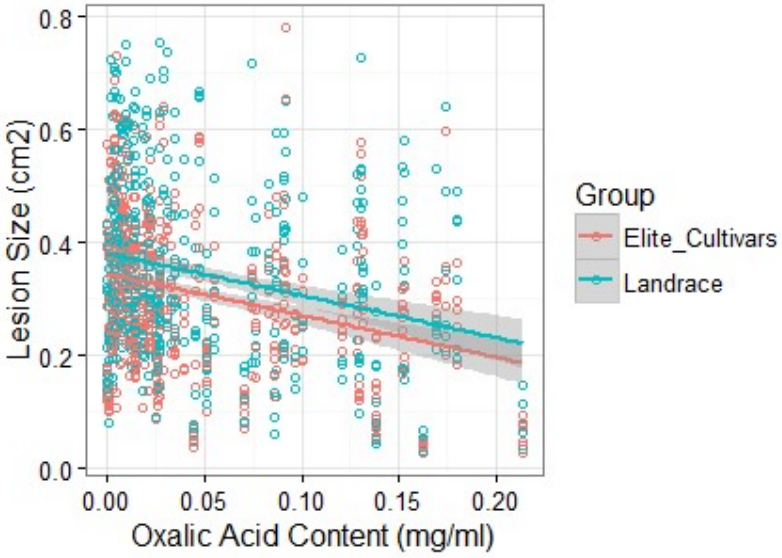
	Df	S.S.	F	p-value
Oxalic acid (OA)	1	1.6262	91.2929	<2.2e-16 ***
Group	1	0.3386	19.0074	1.433e-05 ***
Group/Plant Genotype	10	2.4979	14.0232	<2.2e-16 ***
OA:Group	1	0.0000	0.0020	0.9647
OA:Group/ Plant Genotype	10	0.0291	0.1636	0.9984
Residuals	1032	18.3829		

Signif. codes: '***' 0.001

By adding OA as a covariate, was observed similar results for the domestication group effect and genotype (nested in group), both having significant effect on lesion size. The trait OA production presented a broad-sense heritability of 69.5%. The high heritability value reveals that most of the variation is due to genetic differences in the pathogen set, rather than environmental influence.

Furthermore, our data revealed a negative correlation between lesion size and oxalic acid content (cor= -0.30 ; p-value= 0.00262). Both groups domestic and landrace presented similar correlation patterns (FIGURE 4).

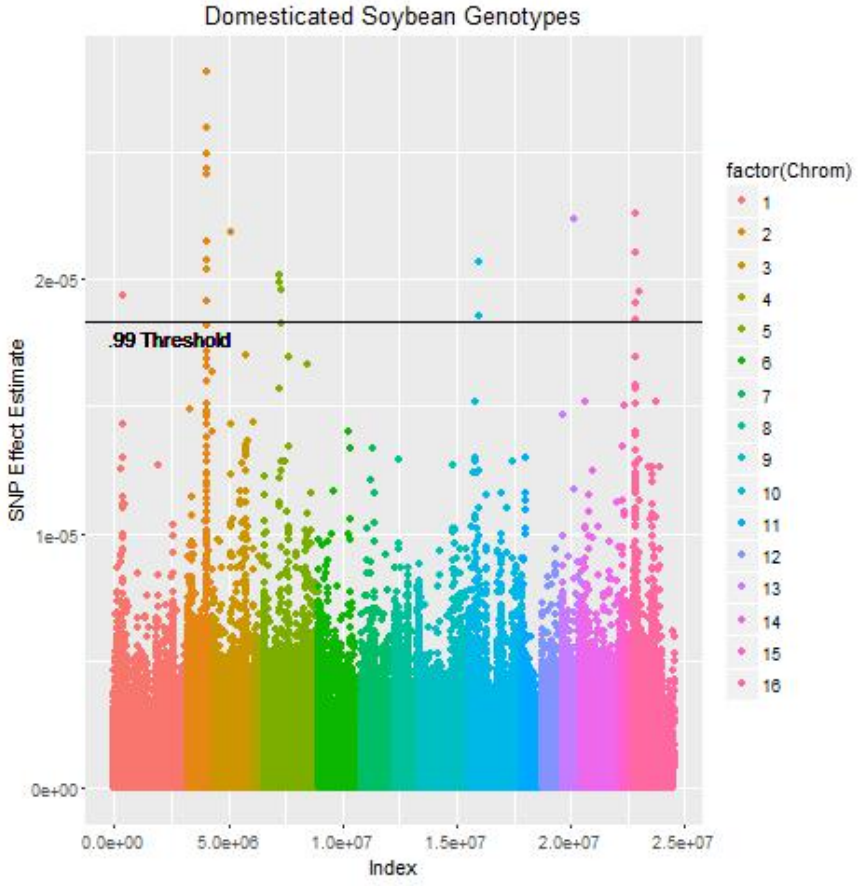
Figure 4 – Lesion size and OA production correlation in elite and landrace soybean genotypes.



3.3 Associative mapping

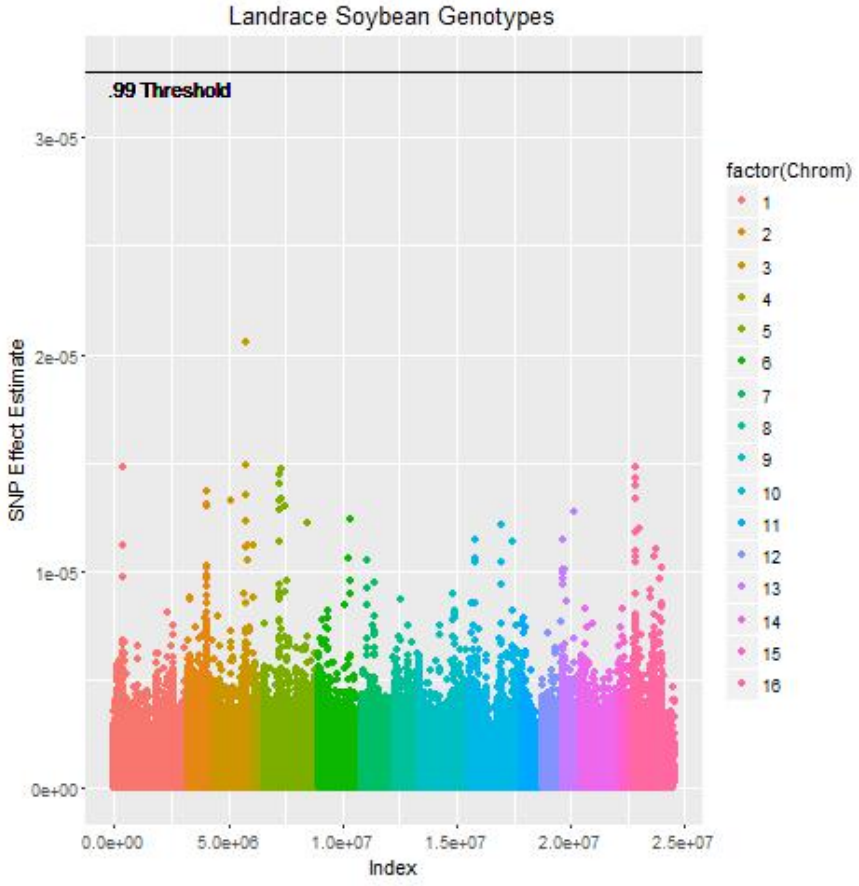
The presented study highlights chromossomes regions significantly associated with *B. cinerea* pathogenicity in domesticated and landrace soybean genotypes. Running the GWAS for domestication groups was possible to highlight chromosome regions significantly associated with B. cinerea pathogenicity in domesticated soybean genotypes, while there was no significant SNPs (single nucleotide polymorphisms) associated with virulence in landrace soybean genotypes (FIGURES 5 and 6).

Figure 5 – Manhattan plot of GWAS showing regions in *B. cinerea* genome associated with lesion size in domesticated soybean genotypes.



The GWAS in domestivated genotypes indicates significantly SNPs associated with lesion size in 7 different chromossomes of B. cinerea genome. Nevertheless, two main regions reveal a more expressive cluster of SNPs with a major effect (chromosome 2 and 15), associated with pathogenicity.

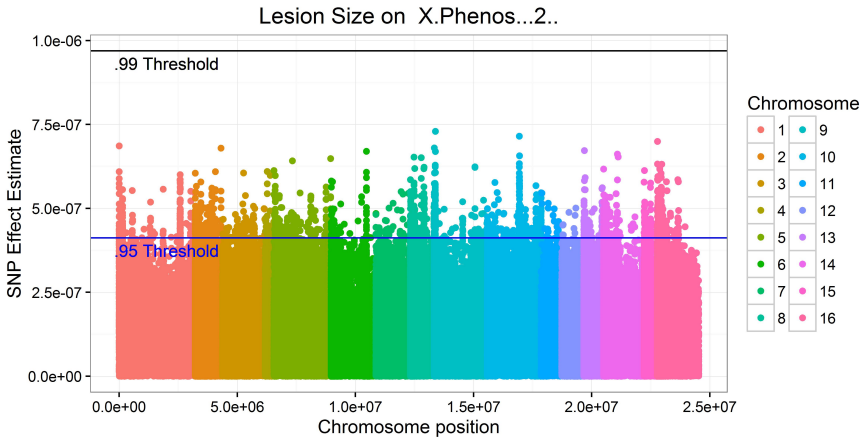
Figure 6 – Manhattan plot of GWAS showing regions in *B. cinerea* genome associated with lesion size in landrace soybean genotypes.



Considering only the landrace soybean genotypes, the GWAS was unable to detect any significant SNP.

The GWAS were also performed to analyse the association between *B. cinerea* genetic variation with OA production. There was no significant SNPs associated with oxalic acid production using a significant threshold of P= 0.01, but using a more relaxed significant threshold (P=0.05) we were able to detect several regions associated with the trait (FIGURE 7).

Figure 7 – Manhattan plot of GWAS showing regions in *B. cinerea* genome associated with oxalic acid production in *B. cinerea* isolates.



4 DISCUSSION

4.1 Detached leaves assay

This study demonstrated differences among isolates, groups of domestication and soybean genotypes within groups for the lesion size of *B. cinerea,* suggesting different levels of resistance to *B. cinerea* infection and distinct virulence capability by the pathogen. The non-significant interactions indicate the pathogen generalist profile. Generalist pathogens are under divergent select pressures due to multiple host species what can limit its capability to develop a more specialized host-pathogen interaction (KNISKERN et al., 2011; KARASOV et al., 2014). Nevertheless, some isolates have exhibited significant differences in virulence across groups and genotypes.

4.2 Oxalic acid

The distinct level of OA production among isolates was expected considering the diverse background of *B. cinerea* isolates. It's well know that this broad-host pathogen present high levels of genetic diversity (ATWELL et al., 2015) and some level of host specificity (MANTEAU et al., 2003) making the optimal condition for oxalic acid production variable and specific for some groups of isolates. Considering that oxalic acid acts shifting the pH in the infection zone and usually the host tissue pH can vary from host to host (MANTEAU et al., 2003) the OA production also relies on a complex sensing and signaling process. The results suggest that some isolates were more capable to produce OA under *in vitro* conditions.

B. cinerea is a versatile phytopatogen regarding virulence factors, and its capacity to adapt to hosts defense metabolites allows this pathogen to overcome strategies based on antifungal compounds (HAIN et al., 1993; QUIDDE et al., 1998). The development of transgenic plants expressing enzymes that can degrade OA secreted by the phytopatogen have been considered a successful strategy to achieve partial resistance in soybean against necrotrophic fungi (DONALDSON et al., 2001). Furthermore, understand how oxalic acid correlate with the infection intensity and its variation among different isolates provide us a better characterization of *B.cinerea* collection and can help with future studies aiming the

development of disease resistance in plants and a clear understanding about plant-pathogen interaction. Even though, the importance of OA in the virulence has been well defined (GODOY et al., 1990; DICKMAN; MITRA, 1992; ROLLINS, 2003; KUNZ et al., 2006) the correlation with lesion size has not been described. Oxalic acid can play distinct roles in the plant pathogen interaction. High levels of oxalic acid can trigger the plant cell death response in another hand a low amount of oxalic acid favor resistance mechanism by the plant. Lehner et al. (2008) described that a pre-treatment with 3mM OA did not have a negative effect on *A. thaliana* and increase the levels of defense-related genes transcripts.Moreover, the pre-treatment with 6mM induced programmed cell death in *A. thaliana* plants.

Among the studied isolates the highest oxalic acid mean (0.21 mg/ml) is below 3mM suggesting the possibility that oxalic acid in this set of isolates is considerably low and might be acting as a plant resistance elicitor causing the negative correlation, where the isolates with higher oxalic acid production were more efficient inducing the plant resistance response. Nevertheless, the specificity of oxalic acid production *in planta* by each isolate was not considered and several virulent factors as well (NAKAJIMA; AKUTSU, 2014), what makes the correlation estimative rough and superficial. Despite of that, we were able to characterize a large set of *B. cinerea* isolates *in vitro* for oxalic acid production. Identify isolates with low OA production may be helpful to develop new sources of biological control using the plant resistance induction mechanism.

4.3 GWAS

B. cinerea shares conserved virulence factors with other phytopatogens, but also present some specific features associates with virulence (CHOKER et al., 2007). Genes associated with specific infection stages have been described, as appressorium penetration (GOURGES et al., 2004), toxin biosynthesis (SIEWERS et al., 2005) H2O2 generation (ROLKE et al., 2004), plant cell wall degradation (KARS et al., 2005) among others.

Nevertheless, *B. cinerea* displays a significant variability depending on the strain genetic background, what raises the question about the plant host influence on the virulence evolution. We detect signicant differences in genes related to virulence according to the domestication degree of the plant host.

REFERENCES

AMSELEM, J.; CUOMO, C. A.; VAN KAN, J. A.; VIAUD, M.; BENITO, E. P.; COULOUX, A.; COUTINHO, P. M.; DE VRIES, R. P.; DYER, P. S.; FILLINGER, S. Genomic analysis of the necrotrophic fungal pathogens Sclerotinia sclerotiorum and Botrytis cinerea. **PLoS Genet**, v. 7, n. 8, p. 100-223, 2011.

ATWELL, S.; CORWIN, J.; SOLTIS, N.; SUBEDY, A.; DENBY, K.; KLIEBENSTEIN, D. J. Whole genome resequencing of Botrytis cinerea isolates identifies high levels of standing diversity. **Frontiers in microbiology**, v. 6, p. 996, 2015.

BATES, D.; MAECHLER, M.; BOLKER, B.; WALKER, S. lme4: Linear mixed-effects models using Eigen and S4. **R package version**, v. 1, n. 7, 2014.

BOEREMA, A.; PEETERS, A.; SWOLFS, S.; VANDEVENNE, F.; JACOBS, S.; STAES, J.; MEIRE, P. Soybean Trade: BalancingEnvironmental and Socio-Economic Impacts of an Intercontinental Market. **PloS one**, v. 11, n. 5, p.155-222, 2016.

CHOQUER, M.; FOURNIER, E.; KUNZ, C.; LEVIS, C.; PRADIER, J.-M.; SIMON, A.; VIAUD, M. Botrytis cinerea virulence factors: new insights into a necrotrophic and polyphageous pathogen. **FEMS Microbiology Letters**, v. 277, n. 1, p. 1-10, 2007.

CLEMENTE, T. E.; CAHOON, E. B. Soybean oil: genetic approaches for modification of functionality and total content. **Plant physiology**, v. 151, n. 3, p. 1030-1040, 2009.

COBER, E. R.; RIOUX, S.; RAJCAN, I.; DONALDSON, P. A.; SIMMONDS, D. H. Partial resistance to white mold in a transgenic soybean line. **Crop Science**, v. 43, n. 1, p. 92-95, 2003.

CUNHA, W.; TINOCO, M.; PANCOTI, H.; RIBEIRO, R.; ARAGÃO, F. High resistance to Sclerotinia sclerotiorum in transgenic soybean plants transformed to express an oxalate decarboxylase gene. **Plant Pathology**, v. 59, n. 4, p. 654-660, 2010.

DAVIDSON, A.; BLAHUT-BEATTY, L.; ITAYA, A.; ZHANG, Y.; ZHENG, S.; SIMMONDS, D. Histopathology of Sclerotinia sclerotiorum infection and oxalic acid function in susceptible and resistant soybean. **Plant Pathology**, v. 65, n. 6, p. 878-887, 2016.

DE GRACIA, M.; CASCALES, M.; EXPERT, P.; BELLANGER, M.-N.; LE CAM, B.; LEMAIRE, C. How did host domestication modify life history traits of its pathogens? **PloS one**, v. 10, n. 6, p. 122-129, 2015.

DICKMAN, M.; MITRA, A. Arabidopsis thaliana as a model for studying Sclerotinia sclerotiorum pathogenesis. **Physiological and Molecular Plant Pathology**, v. 41, n. 4, p. 255-263, 1992.

DONALDSON, P. A.; ANDERSON, T.; LANE, B. G.; DAVIDSON, A. L.; SIMMONDS, D. H. Soybean plants expressing an active oligomeric oxalate oxidase from the wheat gf-2.8 (germin) gene are resistant to the oxalate-secreting pathogen Sclerotina sclerotiorum. **Physiological and Molecular Plant Pathology**, v. 59, n. 6, p. 297-307, 2001.

DUAN, Y.-B.; GE, C.-Y.; ZHOU, M.-G. Molecular and biochemical characterization of Sclerotinia sclerotiorum laboratory mutants resistant to dicarboximide and phenylpyrrole fungicides. **Journal of pest science**, v. 87, n. 1, p. 221-230, 2014.

FAILMEZGER, H.; YUAN, Y.; RUEDA, O.; MARKOWETZ, F. **CRImage**: CRImage a package to classify cells and calculate tumour cellularity. R Packag version 1.4.0, 2010.

GODOY, G.; STEADMAN, J.; DICKMAN, M.; DAM, R. Use of mutants to demonstrate the role of oxalic acid in pathogenicity of Sclerotinia sclerotiorum on Phaseolus vulgaris. **Physiological and Molecular Plant Patholog**y, v. 37, n. 3, p. 179-191, 1990.

GOURGUES, M.; BRUNET-SIMON, A.; LEBRUN, M. H.; LEVIS, C. The tetraspanin BcPls1 is required for appressorium mediated penetration of Botrytis cinerea into host plant leaves. **Molecular microbiology**, v. 51, n. 3, p. 619-629, 2004.

GUIMARAES, R. L.; STOTZ, H. U. Oxalate production by Sclerotinia sclerotiorum deregulates guard cells during infection. **Plant physiology**, v. 136, n. 3, p. 3703-3711, 2004.

HAIN, R.; REIF, H.-J.; KRAUSE, E.; LANGEBARTELS, R.; KINDL, H.; VORNAM, B.; WIESE, W.; SCHMELZER, E.; SCHREIER, P. H. Disease resistance results from foreign phytoalexin expression in a novel plant. **Nature**, v. 361, n. 6408, p. 153, 1993.

HYTEN, D. L.; SONG, Q.; ZHU, Y.; CHOI, I.-Y.; NELSON, R. L.; COSTA, J. M.; SPECHT, J. E.; SHOEMAKER, R. C.; CREGAN, P. B. Impacts of genetic bottlenecks on soybean genome diversity. **Proceedings of the National Academy of Sciences**, v. 103, n. 45, p. 16666-16671, 2006.

KARASOV, T. L.; HORTON, M. W.; BERGELSON, J. Genomic variability as a driver of plant–pathogen coevolution? **Current opinion in plant biology**, v. 18, n., p. 24-30, 2014.

KARS, I.; MCCALMAN, M.; WAGEMAKERS, L.; VAN KAN, J. A. Functional analysis of Botrytis cinerea pectin methylesterase genes by PCR based targeted mutagenesis: Bcpme1 and Bcpme2 are dispensable for virulence of strain B05. 10. **Molecular plant pathology**, v. 6, n. 6, p. 641-652, 2005.

KARS, I.; VAN KAN, J. A. Extracellular enzymes and metabolites involved in pathogenesis of Botrytis. In: **Botrytis:** Biology, pathology and control: Springer, 2007. p. 99-118.

KIM, K. S.; MIN, J.-Y.; DICKMAN, M. B. Oxalic acid is an elicitor of plant programmed cell death during Sclerotinia sclerotiorum disease development. **Molecular Plant-Microbe Interactions**, v. 21, n. 5, p. 605-612, 2008.

KNISKERN, J. M.; BARRETT, L. G.; BERGELSON, J. Maladaptation in wild populations of the generalist plant pathogen Pseudomonas syringae. **Evolution**, v. 65, n. 3, p. 818-830, 2011.

KOENNING, S. R.; WRATHER, J. A. Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. **Plant Health Progress**, v. 10, 2010.

KUNZ, C.; VANDELLE, E.; ROLLAND, S.; POINSSOT, B.; BRUEL, C.; CIMERMAN, A.; ZOTTI, C.; MOREAU, E.; VEDEL, R.; PUGIN, A. Characterization of a new, nonpathogenic mutant of *Botrytis cinerea* with impaired plant colonization capacity. **New phytologist**, v. 170, n. 3, p. 537-550, 2006.

LEHNER, A.; MEIMOUN, P.; ERRAKHI, R.; MADIONA, K.; BARAKATE, M.; BOUTEAU, F. Toxic and signalling effects of oxalic acid: Oxalic acid-Natural born killer or natural born protector? **Plant signaling & behavior**, v. 3, n. 9, p. 746-748, 2008.

LIANG, X.; LIBERTI, D.; LI, M.; KIM, Y. T.; HUTCHENS, A.; WILSON, R.; ROLLINS, J. A. Oxaloacetate acetylhydrolase gene mutants of Sclerotinia sclerotiorum do not accumulate oxalic acid, but do produce limited lesions on host plants. **Molecular plant pathology**, v. 16, n. 6, p. 559-571, 2015.

MANTEAU, S.; ABOUNA, S.; LAMBERT, B.; LEGENDRE, L. Differential regulation by ambient pH of putative virulence factor secretion by the phytopathogenic fungus *Botrytis cinerea*. **FEMS Microbiology Ecology**, v. 43, n. 3, p. 359-366, 2003.

NAKAJIMA, M.; AKUTSU, K. Virulence factors of *Botrytis cinerea*. **Journal of General Plant Pathology**, v. 80, n. 1, p. 15-23, 2014.

PAU, G.; FUCHS, F.; SKLYAR, O.; BOUTROS, M.; HUBER, W. EBImage—an R package for image processing with applications to cellular phenotypes. **Bioinformatics**, v. 26, n. 7, p. 979-981, 2010.

QUIDDE, T.; OSBOURN, A.; TUDZYNSKI, P. Detoxification of α-tomatine by *Botrytis cinerea*. **Physiological and Molecular Plant Pathology**, v. 52, n. 3, p. 151-165, 1998.

R DEVELOPMENT CORE TEAM. Team RDC. **R**: A Language and Environment for Statistical Computing. R Found Stat Comput, 2009. Available: < http://www.r-project.org.>. Acesso em: 12 jan. 2017.

ROLKE, Y.; LIU, S.; QUIDDE, T.; WILLIAMSON, B.; SCHOUTEN, A.; WELTRING, K. M.; SIEWERS, V.; TENBERGE, K. B.; TUDZYNSKI, B.; TUDZYNSKI, P. Functional analysis of H2O2 generating systems in Botrytis cinerea: the major CuZn superoxide dismutase (BCSOD1) contributes to virulence on French bean, whereas a glucose oxidase (BCGOD1) is dispensable. **Molecular plant pathology**, v. 5, n. 1, p. 17-27, 2004.

ROLLINS, J. A. The Sclerotinia sclerotiorum pac1 gene is required for sclerotial development and virulence. **Molecular Plant-Microbe Interactions**, v. 16, n. 9, p. 785-795, 2003.

SALEH, D.; MILAZZO, J.; ADREIT, H.; FOURNIER, E.; THARREAU, D. South East Asia is the center of origin, diversity and dispersion of the rice blast fungus, Magnaporthe oryzae. **New phytologist**, v. 201, n. 4, p. 1440-1456, 2014.

SHEN, X.; ALAM, M.; FIKSE, F.; RÖNNEGÅRD, L. A novel generalized ridge regression method for quantitative genetics. **Genetics**, v. 193, n. 4, p. 1255-1268, 2013.

SIEWERS, V.; VIAUD, M.; JIMENEZ-TEJA, D.; COLLADO, I. G.; GRONOVER, C. S.; PRADIER, J.-M.; TUDZYNSK, B.; TUDZYNSKI, P. Functional analysis of the cytochrome

P450 monooxygenase gene bcbot1 of Botrytis cinerea indicates that botrydial is a strain-specific virulence factor. **Molecular Plant-Microbe Interactions**, v. 18, n. 6, p. 602-612, 2005.

WILLIAMS, B.; KABBAGE, M.; KIM, H.-J.; BRITT, R.; DICKMAN, M. B. Tipping the balance: Sclerotinia sclerotiorum secreted oxalic acid suppresses host defenses by manipulating the host redox environment. **PLoS Pathog**, v. 7, n. 6, p. 102-107, 2011.

WILLIAMSON, B.; TUDZYNSKI, B.; TUDZYNSKI, P.; VAN KAN, J. A. Botrytis cinerea: the cause of grey mould disease. **Molecular plant pathology**, v. 8, n. 5, p. 561-580, 2007.

WRATHER, J.; ANDERSON, T.; ARSYAD, D.; TAN, Y.; PLOPER, L.; PORTA-PUGLIA, A.; RAM, H.; YORINORI, J. Soybean disease loss estimates for the top ten soybean-producing counries in 1998. **Canadian Journal of Plant Pathology**, v. 23, n. 2, p. 115-121, 2001.

XU, L.; XIANG, M.; WHITE, D.; CHEN, W. pH dependency of sclerotial development and pathogenicity revealed by using genetically defined oxalate minus mutants of Sclerotinia sclerotiorum. **Environmental microbiology**, v. 17, n. 8, p. 2896-2909, 2015.

ZHOU, Z.; JIANG, Y.; WANG, Z.; GOU, Z.; LYU, J.; LI, W.; YU, Y.; SHU, L.; ZHAO, Y.; MA, Y. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. **Nature Biotechnology**, v. 33, n. 4, p. 408-414, 2015.

CHAPTER 2 - GENETIC DIVERSITY IN BRAZILIAN SOYBEAN GERMPLASM

RESUMO

A diversidade genética é um fator importante para o sucesso de qualquer programa de melhoramento de plantas, e deve ser considerada visando garantir ganhos genéticos por meio do melhoramento. No Brasil, a pesquisa sobre diversidade genética e estrutura populacional em soja e necessária, considerando que esta espécie e uma importante *commodity* para o país. O estudo em questão foi desenvolvido com o objetivo de avaliar a diversidade genética e a estrutura populacional de 77 genótipos de soja, por meio de 35 marcadores SSR. As estimativas de índices de diversidade revelaram baixos níveis de diversidade genética no grupo de genótipos de soja. Da mesma forma, o coeficiente de Jaccard e a análise de agrupamento baseada em modelo Bayesiano, confirmaram a baixa diversidade entre genótipos de soja, fornecendo evidências que corroboram a ocorrência do efeito *bottleck* nos genótipos de soja brasileiros. Com este trabalho, também destaca-se a importância de se adicionar recursos genéticos com bases genéticas mais contrastantes ao germoplasma de soja brasileiro, com o intuito de garantir ganhos genéticos no melhoramento de soja, a longo prazo.

Palavras-chave: *Glycine* max. Marcadores moleculares. SSR. Plantio.

ABSTRACT

The genetic diversity is an important factor for the success of any plant breeding program and should be considered to assure genetic gain through breeding. In Brazil, research about genetic diversity and population structure for soybean is necessary since the species is an important commodity in the country. This study was developed with the objective of studying the genetic diversity and population structures of 77 soybean genotypes using 35 SSR markers. The diversity index estimative showed low level of genetic diversity in the soybean collection. Similarly, the Jaccard coefficient and Bayesian model based clustering analysis confirmed the low diversity among soybean genotypes, providing an evidence that support the bottleneck effect in the Brazilian soybean genotypes. The study also highlight the importance of add genetic resources with a broader genetic base into the Brazilian soybean germplasm in order to assure the future genetic gain of the soybean breeding.

Keywords: *Glycine max*. Molecular markers. SSR. Plant breeding.

1 INTRODUCTION

The expansion of soybean (*Glycine max* [L.] Merr.) around the world is remarkable. During the past few decades the area cultivated with soybean increased drastically, being among the most cultivated crops in the world (PHALAN et al., 2013), mostly due to its important role as protein and oil source (CLEMENTE; CAHOON, 2009). In Brazil, the increase in area is associated with management innovations and made the country the second largest producer in the world (USDA 2014), with yield indexes increasing in the majority of cultivated areas (RAY et al., 2012).

The development of new crop management technologies and the release of more adapted genotypes can be considered the most important factor to the establishment of soybean as one of the major crops in Brazil (SEDIYAMA et al., 2012). Generally, plant breeding has contributed expressively to the gain of productivity in agriculture crops including soybean. In Iowa, 79% of the genetic gain was obtained from 1930 to 2011 due to the contribution of plant breeding (SMITH et al., 2014). However, the success of genetic improvement programs in any crop species is dependent on the access of a proper genetic diversity pool and its management.

The gradual reduction of genetic diversity lead inevitably to the reduction in the potential of genetic gains, increasing the susceptibility to biotic stress and adaptability to environmental changes (SMITH et al., 2015). Therefore, the breeder should be given attention in order to enrich the genetic pool of the breeding population to maintain the increase in genetic gain for a long period of time.

Genetic diversity studies have as a final goal the understanding of variation among genotypes or groups of genotypes (MOHAMMADI; PRASANNA, 2003). With this purpose, all sorts of data can be used, as an example, genealogy information, morphologic traits, biochemical profile, genotypic data and many others. SSR markers are commonly used in genetic studies due to its multiallelic nature and repeatability across different laboratories (KAGA et al., 2012). The characterization of plant genotypes by DNA molecular markers based on PCR provides a reliable data since it is not influenced by environmental factor as other methods.

Considering the importance of having better understanding on the genetic diversity and the population structure of the breeding program of soybean, we realized the research activity with the objective of evaluating the genetic diversity and population structure of 77

soybean cultivars that represent the soybean genetic resource in Brazil and are routinely used by the farmers and scientists.

2 MATERIALS AND METHODS

The seventy-seven soybean genotypes that represent different maturity group currently released in Brazil and used in the breeding program were included in this study. The list of the seventy-seven genotypes of soybean and its description is presented in Table 1.

Table 1 – The name, source, maturity group, type and year of release of the evaluated soybean genotypes in this study, *RMG – Relative Maturity Group. (continues…)

#	Genotypes	Source	Type	Year of Release	RMG*
1	SUPREMA	Nidera	Conventional	1998	III
2	TMG 801	FMT / TMG / UNISOJA	Conventional	2008	II
3	BRSMG 820 RR	Embrapa / Epamig	Transgenic	2013	II
4	BRSMG 250 (NOBREZA)	Embrapa	Conventional	2005	II
5	MG/BR 46 (CONQUISTA)	Embrapa / Epamig	Conventional	1998	II
6	NA 7337 RR	Nidera	Transgenic	2008	I
7	BRSGO 204 (GOIANIA)	Embrapa / AGDRF	Conventional	2009	II
8	BRSGO JATAÍ	Embrapa / EmaterGO	Conventional	1998	III
9	BRSMG 68 (Vencedora)	Embrapa / Epamig	Conventional	1998	II
10	SYN 1059 (V-TOP)	Syngenta	Transgenic	2010	I
11	TMG 123 RR	FMT / TMG / UNISOJA	Transgenic	2007	I
12	L8307 RR	Monsoy	Transgenic	2006	III
13	BRSMG GARANTIA	Embrapa / Epamig / APSEMG / COPAMIL	Conventional	2000	III
14	P 98Y11	Pioneer	Transgenic	2006	II
15	NK 7074 RR	Syngenta	Transgenic	2007	I
16	BRS SILVANIA RR	Embrapa	Transgenic	2003	III
17	M-SOY 6101	Monsoy / CTPA	Conventional	1998	I
18	ELITE	COOPADAP	Conventional	2001	III
19	M7211 RR	Monsoy	Transgenic	2007	I
20	BRSMG 850 GRR	Embrapa	Transgenic	2007	II
21	MONARCA	COOPADAP	Conventional	1999	III
22	BRSGO LUZIANIA RR	Embrapa	Transgenic	2007	III
23	TMG 1181 RR	FMT / TMG / UNISOJA	Transgenic	2009	II
24	P 98Y30	Pioneer	Transgenic	2009	III
25	M-SOY 8001	Monsoy	Conventional	1998	II
26	BRSMG 772	Embrapa / Epamig	Conventional	2013	II
27	BRS VALIOSA RR	Embrapa	Transgenic	2003	II

#	Genotypes	Source	Type	Year of Release	RMG*
28	NS 5106	Nidera	Transgenic	2012	I
29	NS 5151 IPRO	Nidera	Transgenic	2012	I
30	NA 5909 RG	Nidera	Transgenic	2008	I
31	5D 615 RR	TMG	Transgenic	2011	I
32	CD 215 RR	COODETEC	Transgenic	2011	I
33	BMX POTENCIA	Brasmax	Transgenic	2007	I
34	M-SOY 7908	Monsoy	Transgenic	2004	II
35	TMG 7161	TMG	Transgenic	2010	I
36	TMG 1176	FMT / TMG / UNISOJA	Transgenic	2009	II
37	BRS MG 760 SRR	Embrapa / Epamig	Transgenic	2010	II
38	TMG 132 RR	TMG	Transgenic	2008	III
39	5D 688 RR	COODETEC	Transgenic	2010	I
40	BRS 137	Embrapa	Conventional	2001	I
41	BRSMG 752 S	Embrapa	Conventional	2008	I
42	BRSMG 800 A	Embrapa / Epamig	Conventional	2010	II
43	CD 238 RR	COODETEC	Transgenic	2010	I
44	CD 250 RR	COODETEC	Transgenic	2010	I
45	CD 237	COODETEC	Transgenic	2009	II
46	BRS 284	Embrapa	Conventional	2007	I
47	BRSMG 790A	Embrapa / Epamig	Conventional	2009	II
48	BRSMG 780 RR	Embrapa / Epamig	Transgenic	2011	II
49	NA 7255 RR	Nidera	Transgenic	2008	I
50	NS 7100	Nidera	Transgenic	2009	I
51	CD 202 RR	COODETEC	Transgenic	2011	I
52	NS 7114	Nidera	Transgenic	2012	I
53	5D 711 RR	COODETEC	Transgenic	2010	I
54	NK 7059 (V-MAX RR)	Syngenta	Transgenic	2007	I
55	EMGOPA 316	EMATER-GO	Transgenic	2007	II
56	IAC 19	IAC	Conventional	1998	I
57	TMG 1174 RR	FMT / TMG / UNISOJA	Transgenic	2011	I
58	TMG 1179 RR	FMT / TMG / UNISOJA	Transgenic	2009	II
59	TMG 401	FMT / TMG / UNISOJA	Conventional	2007	II
60	BRSMG 810 C	Embrapa / Epamig	Conventional	2007	II
61	TMG 7262	TMG	Transgenic	2011	I
62	BRS 255 RR	Embrapa	Transgenic	2005	I
63	M 9144 RR	Monsoy	Transgenic	2004	III
64	DM 118	Pioneer	Conventional	1998	II
65	5D 690 RR	COODETEC	Transgenic	2010	I
66	5G 770 RR	COODETEC	Transgenic	2010	II
67	BRS FAVORITA RR	Embrapa	Transgenic	2005	II
68	BMX FORÇA RR	GDM	Transgenic	2008	I
69	NS 7200	Nidera	Transgenic	2011	I
70	TMG 127 RR	UNISOJA / FMT	Transgenic	2007	I
71	MONSOY 8000 RR	Monsoy	Transgenic	2004	II
#	Genotypes	Source	Type	Year of	RMG*

				Release	
72	CAC 1	COOPADAP	Conventional	1998	I
73	CD 2737 RR	COODETEC	Transgenic	2012	I
74	CD 2630 RR	COODETEC	Transgenic	2011	I
75	BRSMG 750 SRR	Embrapa / Epamig	Transgenic	2007	I
76	NA 7620 RR	Nidera	Transgenic	2008	II
77	ANTA 82	FMT / TMG / UNISOJA	Transgenic	2008	I

(conclusion)

The study was conducted in the Molecular Genetics Laboratory at the Department of Biology (DBI), Federal University of Lavras (UFLA). The DNA extraction followed the method described by Pereira et al. (2007). After the DNA samples of each genotype were extracted, the genomic DNA was quantified using spectrophotometer *NanoVue GE* and standardized to 30 ng/µL for SSR genotyping. A total of 35 SSR primers were used in this study. The SSR genotyping was realized using the PCR reaction mix composed of 5.7 µL de Milli-Q water, 3.0 µL of buffer 5X (Green GoTaq®Flexi Buffer - Promega), 2.0 µL of $MgCl_2$ (Promega), 100 µM of each deoxyribonucleotides (dATP, dGTP, dCTP e dTTP), 0.3 µL Taq (5U/µL) and 3.0 µL of each primer pair.

The amplification was performed in Eppendorf Mastercycler® Thermal Cyclers, with the following program: one step at 95 °C for two minutes for initial denaturation, followed by 32 cycles at 94 °C for 30 seconds each for DNA denaturation, 50 seconds for annealing primer at 55ºC, 40 seconds for extension at 72 °C, followed by a final extension at 72 °C for four minutes.

The reaction products were kept under 4°C and separated by electrophoresis in polyacrylamide gel 6% in TBE buffer (0.045 Mtris-borato and 0.001 M EDTA) at 240V for 1 hour and 15 minutes. The gel was submerged a silver nitrate solution (AgNO3) (0.2 %) for 10 minutes, and washed with water and slowly stirred in a NaOH solution (3 % NaOH, 0.5 % formaldehyde) until complete visualization and then photographed using a digital camera.

The DNA fragments were codified as presence and absence to form the allelic matrix for each SSR primer.

2.1 Data analyses

To study the genetic diversity and the population structure, the soybean genotypes were clustered according to its maturity group and method of varietal development (transgenic or conventional). The maturity group was characterized according to the Minas

Gerais macroregional classification (Macro Region 3) where: RMG I (RMG< 7.6); RMG II (7.6 < RMG < 8.2) and RMG III (RMG > 8.2).

The Slatkin genetic distance (SLATKIN, 1995) among the soybean genotypes was estimated using the software PowerMarker V3.25 (LIU et al., 2005). Furthermore, the Nei diversity index (NEI, 1973), Shannon-Weaver index and percentage of polymorphic loci (P%) were estimated using the software POPGENE V1.32 (YEH et al. 1999). The Jaccard similarity coefficient was estimated using R statistical environment (R Core Team, 2014), and used to plot a UPGMA dendrogram using the software MEGA (TAMURA et al., 2007). The AMOVA was done using the software GenAlEx 6.501 (PEAKALL; SMOUSE, 2006).

The Bayesian model based population structure analysis was done using the software Structure 2.3.4 (PRITCHARD et al., 2000) with 10000 burnin and 100000 MCMC. In order to determine the appropriate number of cluster the process was done from K = 1 to 12 with 21 interactions for each K. Finally, the ideal number of cluster (k), or true k, was determined by the ΔK statistics K (EVANNO et al., 2005) using the online software Structure Harvester (EARL, 2012).

3 RESULTS AND DISCUSSION

Among all analyzed primers, the primer Satt222 showed lowest value for both genetic diversity indexes whereas primer Satt251 showed the highest value for Nei genetic diversity index (Table 2). The primer Satt338 produced the highest value for Shannon index. The mean values for the genetic diversity indexes were 0.3863 for Nei index and 0.6294 for Shannon index. The loci Satt270 showedhighest number of polymorphic alleles. However, 21 loci hadonly 2 polymorphic alleles (TABLE 2).

Table 2 – Microsatellite loci, number of alleles (n), Nei genetic diversity index (H), Shannon-Weaver index (I) and Polymorphic Information Content (PIC). (continues….)

Primer name	n	H	I	PIC
Satt200	2.0	0.4979	0.6910	0.3739
Satt239	3.0	0.4876	0.8249	0.4241
Satt270	5.0	0.5943	1.1463	0.5508
Satt196	2.0	0.1214	0.2403	0.1141
Satt492	2.0	0.1847	0.3315	0.1676
Satt358	3.0	0.3771	0.6541	0.3298
Satt012	4.0	0.5158	0.9784	0.4749
Satt076	2.0	0.4252	0.6164	0.3348
Satt225	2.0	0.4898	0.6829	0.3698
Satt137	3.0	0.4192	0.7532	0.3810
Satt475	3.0	0.3287	0.6089	0.2995
Satt241	3.0	0.4557	0.6945	0.3631
Satt345	3.0	0.6044	0.9908	0.5213
Satt426	2.0	0.2088	0.3638	0.1870
Satt251	3.0	0.6610	1.0900	0.5869
Satt509	2.0	0.2922	0.4678	0.2495
Satt338	4.0	0.6606	1.2303	0.6137
Satt369	3.0	0.4459	0.7856	0.4005
Satt342	2.0	0.4581	0.6506	0.3532
Satt274	2.0	0.2449	0.4101	0.2149
Satt520	2.0	0.4554	0.6478	0.3517
Satt052	2.0	0.4932	0.6863	0.3716
Satt222	2.0	0.0256	0.0693	0.0253
Satt302	2.0	0.4992	0.6924	0.3746
Satt146	2.0	0.4369	0.6286	0.3415
Sct_034	2.0	0.2476	0.4135	0.2169
Satt553	3.0	0.3542	0.6589	0.3259
Satt486	2.0	0.0263	0.0708	0.0260
Satt386	2.0	0.2975	0.4741	0.2533
Satt513	4.0	0.4932	0.8421	0.4163
Satt476	3.0	0.4879	0.7285	0.3810
Satt471	2.0	0.1244	0.2449	0.1167
Satt150	2.0	0.4473	0.6395	0.3473

Primer name	n	H	I	PIC
Satt567	2.0	0.3682	0.5548	0.3004
Satt384	2.0	0.2896	0.4648	0.2477
mean	2.5429	0.3863	0.6294	0.3259

(conclusion)

The PIC (polymorphic information content) was determined for every marker (Table 2), and its mean value was 0.3259. The primer Satt338 showedthe highest PIC (PIC = 0.6137), while the lowest was observed for Satt222 (PIC = 0.0253).

The comparison among the soybean varieties obtained through transgenic and conventional in relation to the genetic diversity indexes showed slightly higher for the varieties obtained through conventional breeding method (TABLE 2). The transgenic cultivars had 100% of polymorphic loci while the conventional group showed 88.57% of polymorphic loci.

Except for the RMG III, all the subsets presented similar number of alleles. The RMG II presented higher diversity index and the RMG I the lowest for both genetic diversity indexes. Moreover, these two subsets show the highest PIC. The RMG III subset presented the smallest number of polymorphic alleles and Shannon diversity index (TABLE 3).

Table 3 – Evaluated soybean subsets (Subsets), number of genotypes by subset (#), average of observed alleles (n), Shannon-Waever diversity index (I), Nei diversity index (H), percent polymorphic (PC).

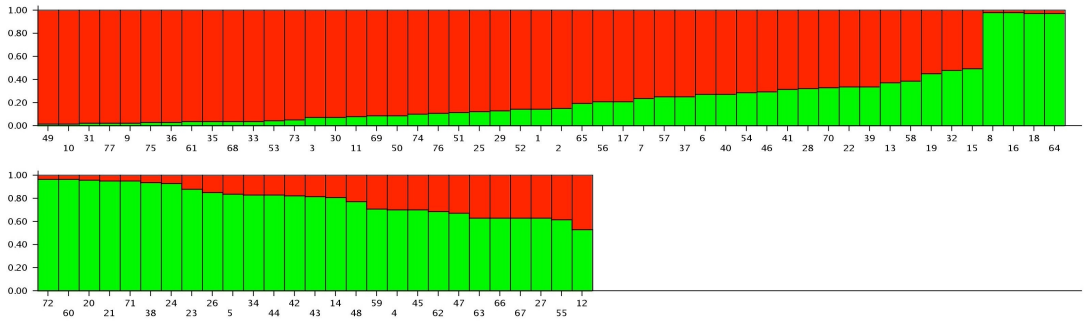
						AMOVA	
Subsets	#	n	I	H	PC(%)	Among	Within
Convencional	23	2.4	0.6198	0.3817	88.57	1%	99%
Transgênico	54	2.4857	0.6102	0.3761	100		
RMG I	38	2.4286	0.5611	0.3426	97.14	6%	94%
RMG II	28	2.4571	0.6230	0.3823	97.14		
RMG III	11	2.0571	0.5532	0.3583	80		
Total	77	2.5429	0.6294	0.3863	100		

Considering the subset of transgenic and conventional, the AMOVA showed that 99% of total variation was partitioned within the subset whereas the remaining 1% among subsets. Similarly, considering maturity groups, 94% the total variation was partitioned within groups and 1% among groups (TABLE 3).

The Bayesian model base clustering analysis clustered the 77 genotypes in to two principal groups (K=2). The first group (red) is composed mostly for short season cultivars, while the second group (green) mostly by long season cultivars. There is certain degree of admixture represented mostly by genotypes with medium cycle, but also by RMG I and III

(FIGURE 2). The formation of two groups showed low genetic diversity among the tested soybean genotypes.

Figure 2 – Two subgroups inferred from STRUCTURE analysis. The vertical coordinate indicates the membership coefficients for each individual, and the digits on the horizontal coordinate represent the accessions reference number corresponding to Table 1.



Low genetic diversity was observed among soybean genotypes tested in this study, and in all subsets. The genetic diversity indexes estimated in this study can be considered low when compared to results previously described for this species. DONG et al. (2014), analyzed the genetic diversity among 100 vegetable soybean cultivars and described an average Nei index of 0.6286. WANG et al. (2014) described a Shannon-Weaver index of 2.038 analyzing the genetic diversity of 10 landrace populations in China. The high Shannon-Weaver index observed by the authors may be explained by the type of genotypes used in both studies (vegetable and landrace). Vegetable soybean and had a smaller selection pressure when compared to the grain-type. Moreover, landrace genotypes are more close relatives to the wild genotypes (*G. soja*) where naturally have high level of genetic diversity. The present study, most of the genotypes included are improved soybean genotypes that passed through a more intensive selection pressure.

The low diversity in a germplasm is particularly alarming when we consider the agriculture vulnerability against the changes in the climate patterns. The development of genotypes with better performance in high temperatures, high CO_2 concentration, low water availability and salinity conditions become a challenge to be overcome with certain urge due to the climate change (CECCARELLI et al., 2010). But to face this challenges and develop soybean cultivars with the desirable traits requires diversifying the genetic background of the current breeding population through incorporating new genetic backgrounds from other countries.

The genetic base of Brazilian soybean germplasm is known to be narrow (WYSMIERSKI; VELLO, 2013). The frequent use of a small set of genotypes in the breeding process could be considered one key factor for the reduction of genetic diversity. The large majority of the Brazilian soybean germplasm are derived from 4 genotypes (CNS, S-100, Roanoke and Tokyo), and these soybean genotypes contribute more than half of the genetic base of the cultivars released in Brazil. If we consider the 14 soybean genotypes used in the breeding program, their contribution to the genetic base can reach up to 92.4% (WYSMIERSKI; VELLO, 2013). Their conclusion is supported by our result obtained in this study since 77 cultivars of soybean clustered only into two groups with high fixation index in each cluster. In summary, the soybean genetic bases used in the breeding program of Brazil were built from small numbers of genotypes. Our data agrees with that pattern and reinforce the results described by Wysmierski and Vello (2013). Besides, the soybean breeding methods contribute to the genetic bottleneck, once we assume that the backcrossing is a method routinely used to introduce qualitative traits in an elite cultivar especially transgenic lines. The subset of transgenic cultivars in this study showedlower genetic diversity indexes compared to traditional genotypes, despite the higher number of genotypes included in this group (TABLE 3).

The development of transgenic cultivars by backcrossing can lead to the reduction of the genetic richness since the process employed a reduced number of recurrent parents and a few donor parents carrying the transgenic segment. Furthermore, the small number of transformed lines reduce the range of genetic variation available to be selected by the plant breeder (SNELLER, 2003).

Considering the raising demand for transgenic and short season cultivars in the seed market, we can assume that man-made genetic bottleneck is still an ongoing process that should be mitigated. The concern about genetic diversity and the enlargement of soybean genetic pool can be helpful in the long run to preserve the increments in performance by means of plant breeding.

The result from AMOVA showed that the absence of clear differentiation among the groups of soybean genotypes indicates narrow genetic base within the soybean cultivars grown in Brazil (TABLE 3). Therefore, the fellow soybean breeder should follow strong strategy in order to broaden the genetic base of different groups through incorporating new genetic base, especially the landraces.

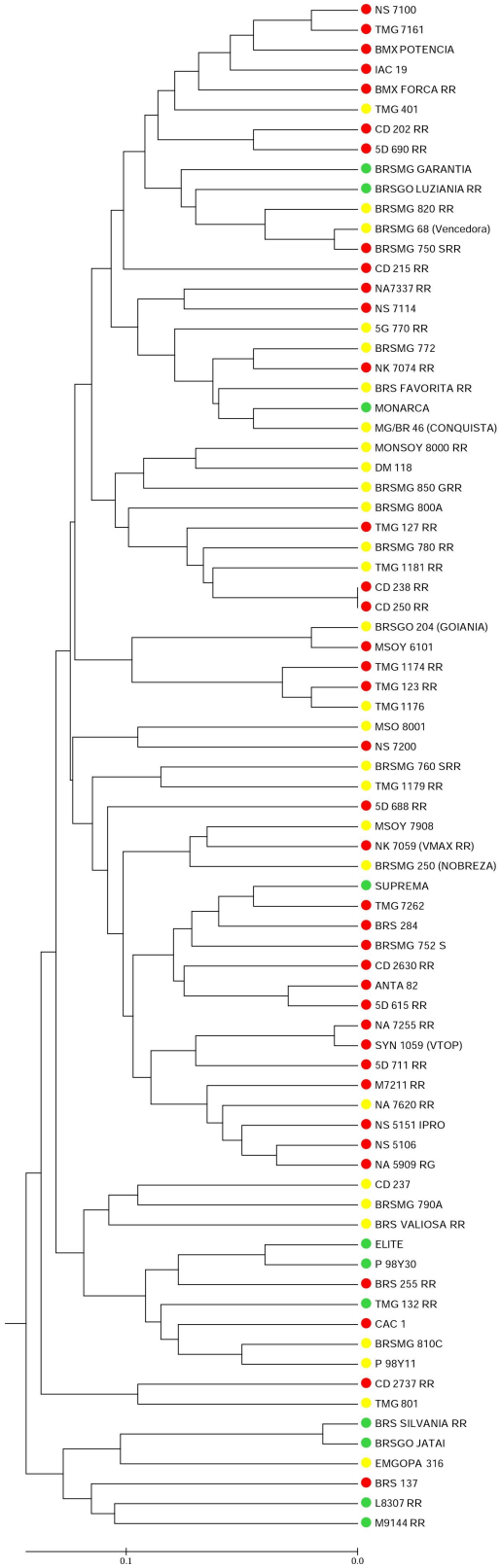
Most of the soybean cultivars analyzed in this study are short season genotypes (RMG I). Nowadays, the short season cultivars become more popular among Brazilian

farmers due to the production system employed that is based on the succession of crops using early maturing cultivars in the summer followed by another crop in a short time gap. This system largely used in Brazil is called “safrinha” and aim to maximize the land use across time. Thus, this production model drives the seed companies to focus on the development of short season cultivars in a faster rate, what might be another factor contributing to the bottleneck.

The lowest Nei diversity index (H=0.3426) was observed in the short season group (RMG I) suggesting the low genetic diversity within this group. The RMG III showed the lowest Shannon-Waever index suggesting low diversity as well. Nevertheless, the RMG III is composed by a smaller number of genotypes, that can underestimate its diversity range (TABLE 3).

In our results, the absence of clear clustering pattern in the dendrogram regarding relative maturity groups, and the small distance between clusters (FIGURE 1), suggested that the cultivars with short and long season might share alleles obtained from the same parents used in the breeding process.

Figure 1 – Dendrogram based on Jaccard similarity coefficient, showing the soybean genotypes from different maturity groups: RMG I (red), RMG II (yellow), RMG III (green)



A pattern was observed in the population structure, where two groups were formed. The first group was composed by 73.9% of individuals from the RMG I and the second group by 61.3% from the RMG II and 25.8% from the RMG III, indicating that a higher allele frequency of alleles associated to a short season in the first group and in the second group a higher frequency of alleles associated to a long season (FIGURE 2).

In general, the result suggested low genetic variability among soybean cultivars, reinforcing previous results. The popularization in development and utilization of short season transgenic cultivars should be followed by mitigating measures to avoid the human-made bottleneck. In this context, the soybean breeding programs need to consider the genetic diversity issue as an important factor instead of focus exclusively in yield and other agronomic traits.

In this regard the wild soybean species as *Glycine soja* Seib. and Zucc., can be considered to increase genetic variation and can be used as an important source of alleles, despite its non-domesticated characteristics (AKPERTEY et al., 2014). Interspecific crosses are commonly used in several crop species with the objective of introducing new alleles to overcome biotic and abiotic stress and increase the genetic diversity (HAUSSMANN et al., 2004). Unfortunately, there is a lack of availability of this genetic material in the Brazilian's germplasm (CARTER et al., 2004).

An international cooperation aiming to share and exchange soybean germplasm should be considered to enrich the Brazilian soybean germplasm pool to avoid the bottleneck many countries of the world. In Brazil, this bottleneck effect should be considered as a worrying fact for the future development of soybean breeding programs.

Expand the soybean genetic base its fundamental to assure the future progression of genetic gains through plant breeding. A more diverse genetic pool and germplasm set can make it easy the development of new varieties capable to overcome a hostile environment.

The analyzed soybean genotypes showedlow genetic diversity, pointing the need to increase the variability in the genetic pool of soybean in the breeding program of Federal University of Lavras.

REFERENCES

AKPERTEY, A.; BELAFFIF, M.; GRAEF, G. L.; MIAN, R.; GROVER SHANNON, J.; CREGAN, P. B.; HUDSON, M. E.; DIERS, B. W.; NELSON, R. L. Effects of selective genetic introgression from wild soybean to soybean. **Crop Science**, v. 54, n. 6, p. 2683-2695, 2014.

CARTER, T.; NELSON, R.; SNELLER, C.; CUI, Z. **Genetic diversity in soybean**. Soybean monograph. American Society of Agronomy, Madison, Wisconsin, USA, p. 303-416, 2004.

CECCARELLI, S.; GRANDO, S.; MAATOUGUI, M.; MICHAEL, M.; SLASH, M.; HAGHPARAST, R.; RAHMANIAN, M.; TAHERI, A.; AL-YASSIN, A.; BENBELKACEM, A. Plant breeding and climate changes. **The Journal of Agricultural Science**, v. 148, n. 06, p. 627-637, 2010.

CLEMENTE, T. E.; CAHOON, E. B. Soybean oil: genetic approaches for modification of functionality and total content. **Plant physiology**, v. 151, n. 3, p. 1030-1040, 2009.

DONG, D.; FU, X.; YUAN, F.; CHEN, P.; ZHU, S.; LI, B.; YANG, Q.; YU, X.; ZHU, D. Genetic diversity and population structure of vegetable soybean (Glycine max (L.) Merr.) in China as revealed by SSR markers. **Genetic resources and crop evolution**, v. 61, n. 1, p. 173-183, 2014.

EARL, D. A. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation genetics resources**, v. 4, n. 2, p. 359-361, 2012.

EVANNO, G.; REGNAUT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Molecular ecology**, v. 14, n. 8, p. 2611-2620, 2005.

HAUSSMANN, B.; PARZIES, H.; PRESTERL, T.; SUSIC, Z.; MIEDANER, T. Plant genetic resources in crop improvement. **Plant genetic resources**, v. 2, n. 1, p. 3-21, 2004.

KAGA, A.; SHIMIZU, T.; WATANABE, S.; TSUBOKURA, Y.; KATAYOSE, Y.; HARADA, K.; VAUGHAN, D. A.; TOMOOKA, N. Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. **Breed Sci**, v. 61, n. 5, p. 566-592, 2012.

LIU, K.; MUSE, S. V. PowerMarker: an integrated analysis environment for genetic marker analysis. **Bioinformatics**, v. 21, n. 9, p. 2128-2129, 2005.

MOHAMMADI, S.; PRASANNA, B. Analysis of genetic diversity in crop plants—salient statistical tools and considerations. **Crop Science**, v. 43, n. 4, p. 1235-1248, 2003.

NEI, M. Analysis of gene diversity in subdivided populations. **Proceedings of the National Academy of Sciences**, v. 70, n. 12, p. 3321-3323, 1973.

PEAKALL, R.; SMOUSE, P. E. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. **Molecular ecology notes**, v. 6, n. 1, p. 288-295, 2006.

PEREIRA, H. S.; SANTOS, J. D.; ABREU, A. D. F.; COUTO, K. R. Informações fenotípicas e marcadores microssatélites de QTL na escolha de populações segregantes de feijoeiro. **Pesquisa Agropecuária Brasileira**, v. 42, n. 5, p. 707-713, 2007.

PHALAN, B.; BERTZKY, M.; BUTCHART, S. H.; DONALD, P. F.; SCHARLEMANN, J. P.; STATTERSFIELD, A. J.; BALMFORD, A. Crop expansion and conservation priorities in tropical countries. **PloS one**, v. 8, n. 1, p. e51759, 2013.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945-959, 2000.

RAY, D. K.; RAMANKUTTY, N.; MUELLER, N. D.; WEST, P. C.; FOLEY, J. A. Recent patterns of crop yield growth and stagnation. **Nature communications**, v. 3, n., p. 1293, 2012.

R CORE TEAM. **R:** A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2014.

SEDIYAMA, C. S.; CARNEIRO, J. E. D. S.; FRITSCHE-NETO, R.; SEDIYAMA, T.; BARBOSA, M. H. P.; GALVÃO, J. C. C.; SOUZA, M. A. D. Contribution of the universities to the development of field crop cultivars. **Crop Breeding and Applied Biotechnology**, v. 12, n. SPE, p. 121-130, 2012.

SLATKIN, M. A measure of population subdivision based on microsatellite allele frequencies. **Genetics**, v. 139, n. 1, p. 457-462, 1995.

SMITH, J. S. C.; DIERS, B. W.; SPECHT, J. E.; CARVER, B. F. **Yield gains in major US field crops**: American Society of Agronomy, 2014

SMITH, S.; BUBECK, D.; NELSON, B.; STANEK, J.; GERKE, J. Genetic diversity and modern plant breeding. In: **Genetic diversity and erosion in plants**: Springer, 2015. p. 55-88.

SNELLER, C. H. Impact of transgenic genotypes and subdivision on diversity within elite North American soybean germplasm. **Crop Science**, v. 43, n. 1, p. 409-414, 2003.

TAMURA, K.; DUDLEY, J.; NEI, M.; KUMAR, S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. **Molecular biology and evolution**, v. 24, n. 8, p. 1596-1599, 2007.

WANG, K.-J.; LI, X.-H.; YAN, M.-F. Microsatellite markers reveal genetic diversity of wild soybean in different habitats and implications for conservation strategies (Glycine soja) in China. **Conservation genetics**, v. 15, n. 3, p. 605-618, 2014.

WYSMIERSKI, P. T.; VELLO, N. A. The genetic base of Brazilian soybean cultivars: evolution over time and breeding implications. **Genetics and molecular biology**, v. 36, n. 4, p. 547-555, 2013.

YEH, F. C.; YANG, R.; BOYLE, T.; YE, Z.; MAO, J. X. **POPGENE**, version 1.32: the user friendly software for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, AB, Canada, 1999.