# Long-read targeted region sequencing and functional studies of the rust resistance gene *R11* in sunflower

Md Shamimuzzaman[1], Guojia Ma[2], and Lili Qi[1]

[1]USDA-Agricultural Research Service, Edward T. Schafer Agricultural Research Center, 1616 Albrecht Blvd. N., Fargo, ND 58102

[2]Department of Plant Sciences, North Dakota State University, 1360 Albrecht Blvd. N., Loftsgard Hall 166, Fargo, ND 58102

## Abstract

Rust, caused by the fungus *Puccinia helianthi* Schwein., is one of the most devastating diseases of sunflower (Helianthus annuus L.) affecting global sunflower production. The rust resistance locus *R11* in the sunflower germplasm line HA-R9 was mapped in the chromosome 13 with an interval of genomic regions having multiple candidate genes. EMS (ethyl methane sulfonate) mutagenesis of HA-R9 was performed to develop rust susceptible lines for identification of the candidate gene that confers resistance to rust. A total of 2,350 HA-R9 seeds were treated with 0.7% EMS for 6 hours. Out of 559 M2 populations tested for their reaction to rust, susceptible plants were identified in 63 M2 populations. The six M3 families showed homozygous susceptible to rust infection. Sequencing of a 60 kb region spanning the *R11* locus on both the *R11*-HA-R9 and the six *R11*-susceptible mutants identified three genes corresponding to the genes HanXRQChr13g0422111, HanXRQChr13g0422121, and HanXRQChr13g0422131 annotated in the XRQ reference genome. The first two genes have exactly same sequences in all three genomes, XRQ, *R11*-HA-R9, and *R11*-susceptible mutants, while three point mutations were found in the third gene corresponding to HanXRQChr13g0422131 in the three *R11*-susceptible mutants. First two point mutations caused premature stop codons, and third mutation caused the changes in amino acid from serine to asparagine. Additional functional studies using comparative RNA sequencing (RNA-Seq) of the *R11*-HA-R9 and the *R11*-susceptible mutants are in progress. Identification of differentially expressed genes will shed further light on the gene regulatory networks in the rust disease resistance pathways in sunflower.

## Introduction

*Puccinia helianthi* Schwein. is a fungal pathogen that causes rust on sunflower. This disease significantly affect the quality and production of sunflower in the northern Great Plains where most of the U.S. sunflower is grown. The rust resistance is conferred by single dominant gene in sunflower. The rust resistance locus *R11* in sunflower line, HA-R9 was mapped in chromosome 13 and it provides resistance to all *P. helianthi* races identified so far in North America (Qi et al., 2012, Ma et al., 2020 unpublished). The genomic region containing *R11* locus was narrowed down and three potential candidate genes (HanXRQChr13g0422111, HanXRQChr13g0422121, and HanXRQChr13g0422131) for $R_{11}$ were identified in the XRQ reference genome (Ma et al., 2020 unpublished). EMS (ethyl methane sulfonate) mutagenesis of HA-R9 was carried out to develop rust susceptible lines to pinpoint the actual *R11* rust resistance gene. To overcome the repetitive nature of the genomic regions harboring *R11* rust resistance gene, we used long-read PacBio target region sequencing to obtain full length gene sequence. In this poster, we summarize our long-read target region sequencing results of both HA-R9 resistance and EMS induced susceptible mutant lines. Also, we report the EMS induced mutations that caused premature stop codons and made the HA-R9 sunflower line susceptible to rust.

## Materials & Methods

**EMS Mutagenesis**

Development of Rust Susceptible HA-R9 Mutants

- HA-R9 (M0) seeds were treated with 0.7% EMS for 6 hours
- 2350 treated (M1) seeds were planted
- 1013 M2 individuals (558 M2 individuals with >20 seeds)
- 63 M2 susceptible plants were identified

**Rust Resistance Evaluation**

- The M2 populations and subsequent M3 families were evaluated for rust resistance in the greenhouse.
- The inbred lines, HA 89 and HA-R9 were used as susceptible and resistant controls, respectively.
- Rust resistance was evaluated 12-14 days after inoculation for both infection types (ITs) based on the 0-4 scale and severity.
- Infection types 0, 1, and 2 combined with pustule coverage of 0 to 0.5% were classified as resistant
- Infection types 3 and 4 with pustule coverage >0.5% were considered as susceptible.

**Long-read Targeted Sequencing**

- DNA was extracted from a rust resistant *R11*-HA-R9 and six *R11*-susceptible mutants.
- About 60 Capture probes were designed to target 60 kb genomic region harboring *R11* gene.
- De novo assembly of long reads was performed using canu assembler to obtain full length *R11* gene sequence.
- EMS induced SNPs were identified when sequences from resistant and susceptible lines were compared.
- Integrative Genomics Viewer (IGV) was used to visualize the SNPs in *R11* gene of susceptible mutants.

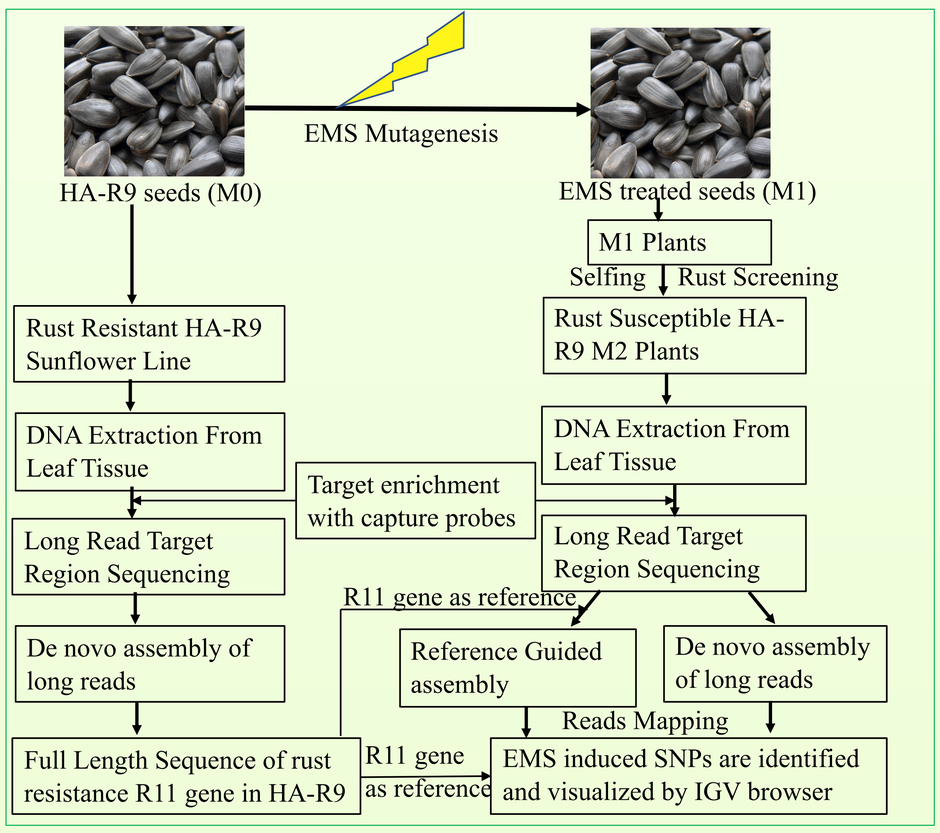


Fig 1. Mutational genomics strategy for sequence-based cloning of the rust resistance R11 gene using long-read (PacBio) sequencing

## Results

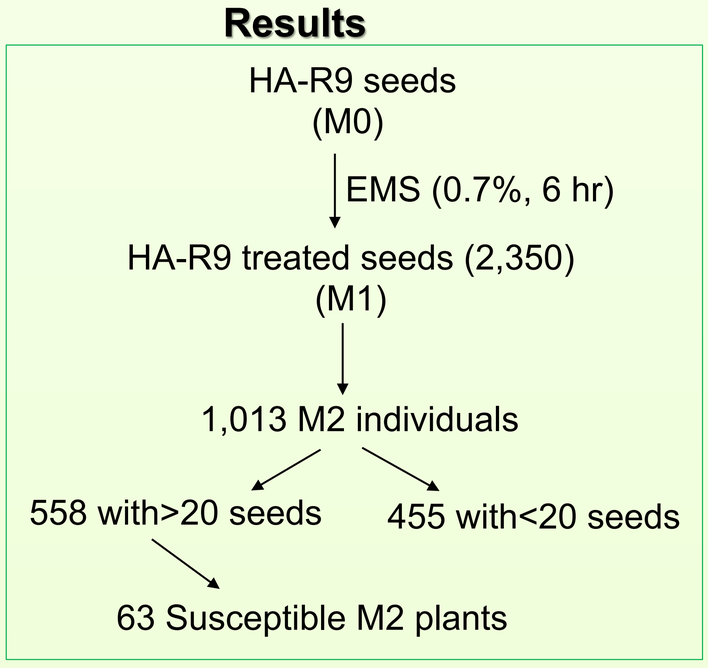


Fig 2. EMS mutagenesis to develop rust susceptible HA-R9 mutants

- 558 M2 populations (24 plants for each M2) were grown
- Total 13,392 plants were screened
- 63 rust susceptible M2 plants were identified

Genomic co-ordinates of target region: HanXRQChr13-181784941..181844941 (~60kb)

3 genes are located in the target region:

- HanXRQChr13g0422111: Probable anthocyanidin 3-O-glucosyltransferase 1 (AOGT)
- HanXRQChr13g0422121: Putative UDP-glucuronosyl/UDP-glucosyltransferase (UGT)
- HanXRQChr13g0422131: Putative NB-ARC family protein
- Long-read targeted region sequencing was conducted
- Probe Design: 120-mer probes were designed every 1kb apart.
- Xgene lockdown probe pools containing ~60 probes were used for target enrichment.

| Sample | DNA Source/Pedigree | Gene Status |
|---|---|---|
| S2 | HA-R9-1 | R11 Gene |
| S3 | 17-055-19a | Mutant R11 |
| S4 | 17-055-20a | Mutant R11 |
| S5 | 17-055-21a | Mutant R11 |
| S6 | 18-004-20a | Mutant R11 |
| S7 | 18-004-34a | Mutant R11 |
| S8 | 18-004-36a | Mutant R11 |

Table 1. Sample Information and status of rust resistance gene R11

- Bioinformatics analysis was performed using PacBio long read sequence data.
- De novo and reference guided assembly was carried out to obtain full length gene sequences located in the target region
- Comparative sequence analysis of both HA-R9 rust resistant line and EMS induced rust susceptible mutants was performed to identify mutations.
- No mutation was found in two genes (AOGT and UGT).
- EMS induced SNPs found in gene encoding NB-ARC family protein


Fig 4. Multiple sequence alignment of deduced protein sequences of R11 gene

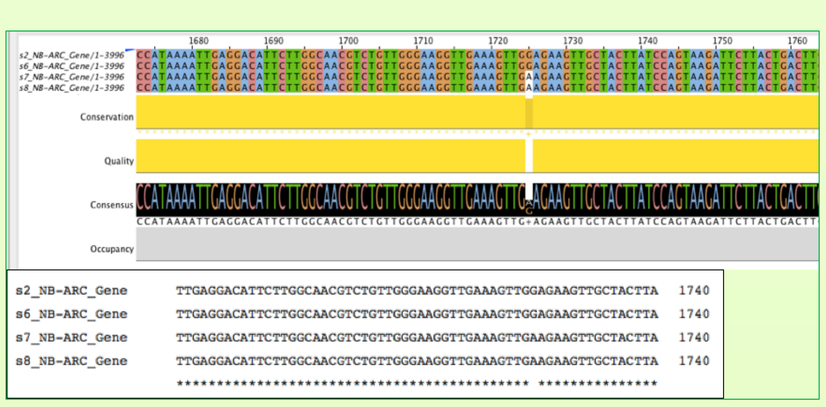


Fig 3. Multiple sequence alignment of *R11* gene assembled from different samples

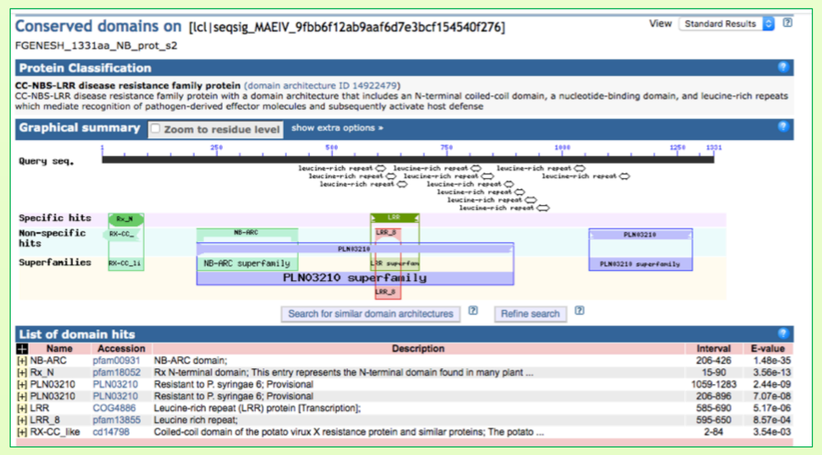


Fig 5. Conserved Domains of *R11* gene encoding CC-NBS-LRR protein

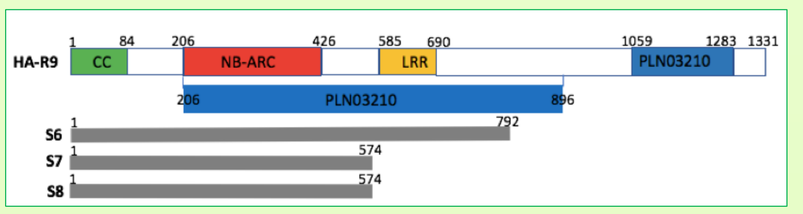


Fig 6. Domains of NB-ARC/CC-NBS-LRR protein affected due to mutations

## Summary:

- Performed PacBio long read target region sequencing using DNA extracted from both HA-R9 rust resistant line and EMS induced rust susceptible mutants
- Bioinformatics analysis Identified mutations in NB-ARC/CC-NBS-LRR gene gene located in the target genomic region and established this gene is as $R_{11}$ gene.
- Three rust susceptible HA-R9 mutant lines (S6, S7, S8) have point mutations at position 1725 G>A, 2378 G>A and 3107 G>A. First two point mutations caused premature stop codons at amino acid position 575 and 793. Third mutation caused the changes in amino acid from serine to asparagine.
- EMS mutation introduces premature stop codon so that the resulting protein was shorter and lack of certain domains (Fig. 8). It causes the plant to become susceptible to rust.

## References

Qi LL, Seiler GJ, Hulke BS, Vick BA, Gulya TJ (2012). Genetics and mapping of the *R11* gene conferring resistance to recently emerged rust races, tightly linked to male fertility restoration, in sunflower (Helianthus annuus L.). Theor Appl Genet 125, 921–932.

Ma G, Long Y, Song Q, Talukder ZI, Shamimuzzaman M and Qi L (2020). Map and sequence-based chromosome walking towards cloning of the male fertility restoration gene Rf5 linked to *R11* in sunflower. (Accepted by Scientific Reports).

## Acknowledgment

We thank Angelia Hogness for technical assistance. This project was supported by the USDA-ARS CRIS Project No.3060-2100-043-00D.