Data in Brief 50 (2023) 109624

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Genomic data of two chickpea populations sharing a potential Ascochyta blight resistance region

Alejandro Carmona*, Patricia Castro, Adrian Perez-Rial, Jose V. Die

Department of Genetics Escuela Técnica Superior de Ingeniería Agronómica y de Montes (ETSIAM), University of Cordoba, Cordoba, Spain

ARTICLE INFO

Article history:
Received 15 June 2023
Revised 19 September 2023
Accepted 20 September 2023
Available online 26 September 2023

Dataset link: Annotation of candidate fungal resistance region (Original data)
Dataset link: RIP10 in chickpea (Original data)
Dataset link: RIP8 in chickpea (Original data)

Keywords:
Ascochyta rabiei
Cicer arietinum
DNA
GO term
RIL
Sequencing
SRA

ABSTRACT

Chickpea (*Cicer arietinum* L.) is one of the most important crops worldwide and a valuable nutritional source. The availability of data from different genotypes and populations is important for the comprehension of the biology and trait control of chickpea. Tissue from young leaves was collected from adult plants and sequenced using an Illumina HiSeq X platform, which provided sequencing data for a total of 169 individuals from two different populations. Furthermore, functional annotation was performed with BLAST2GO software in a candidate region for resistance to Ascochyta blight, a devastating disease that produces huge yield reductions if the growth conditions are optimal for the fungus. A total of 273 different genes in a region spanning ~4.67 Mb in chromosome 4 were fully annotated. The raw DNA sequences and functional annotation data can be reused by the scientific community for the analysis of different agronomic traits of interest in chickpea.

* Corresponding author at: Department of Genetics, Campus Universitario de Rabanales, Ctra. N-IV, Km. 396, 14014 Cordoba, University of Cordoba, Spain.
E-mail address: b62cajia@uco.es (A. Carmona).
Social media: @Aleexcj98 (A. Carmona), @Aprial97 (A. Perez-Rial)

https://doi.org/10.1016/j.dib.2023.109624

Specifications Table

Subject	Agricultural Sciences Biological Sciences
Specific subject area	Agricultural Engineering Agronomy and Crop Science Bioinformatics Genetics: General
Type of data	Figure Tables DNA sequences Gene Ontology data
How the data were acquired	The sequence read archive (SRA) reads were obtained from the sequencing of young leaves tissue via Illumina HiSeq X platform. Quality checks were performed to ensure the reliability of the SRA reads. The candidate fungal resistance region was downloaded from NCBI and genes within the region were functionally annotated using default parameters in BLAST2GO v6.0.3 software.
Data format	Raw
Description of data collection	The data were obtained from two recombinant inbred line (RIL) populations segregating for Ascochyta blight resistance. RIP8 was developed from the cross ILC3279 (Resistant) x WR315 (Susceptible), and RIP10 was derived from the cross JG62 (Susceptible) x ILC72 (Resistant). Genotypic data were obtained for both parents and 84/81 RILs in RIP8 and RIP10 respectively (86/83 individuals in RIP8/RIP10).
Data source location	Department of Genetics, Campus Universitario de Rabanales, Ctra. N-IV, Km. 396, 14014 Córdoba, University of Cordoba, Spain. Lat. 37.91350199528213, Long. -4.7191828471133075
Data accessibility	The SRA data are publicly available under the following repository: Repository name: National Center for Biotechnology Information. Data identification numbers: SRP441161, SRP441763. Direct URL to data: https://trace.ncbi.nlm.nih.gov/Traces/?view=study&acc=SRP441161 https://trace.ncbi.nlm.nih.gov/Traces/?view=study&acc=SRP441763 The data for the functional annotation of the candidate fungal resistance region are available under the following repository: Repository name: Mendeley Data Data identification number: 10.17632/z34fks5sr2.1 Direct URL to data: https://data.mendeley.com/datasets/z34fks5sr2/1
Related research article	A. Carmona, J. Rubio, T. Millan, J. Gil, J.V. Die, P. Castro, Four haplotype blocks linked to Ascochyta blight disease resistance in chickpea under Mediterranean conditions, Front. Plant Sci. 14 (2023): 1183287. DOI: 10.3389/fpls.2023.1183287

1. Value of the Data

- The data presented here contribute valuable insights into the genetic basis of Ascochyta blight resistance, a disease that poses a significant threat to the global chickpea industry. Previous research has revealed that the disease control is polygenic and multiple regions containing potential candidate genes have been described. However a consensus on a shared resistance mechanism across various populations remains elusive. Introducing new candidate regions and genes using different populations may help comparing the resistance mechanism among populations and genotypes, and provide researchers with markers that can be further used to select resistant genotypes in different breeding programs.
- Given that Ascochyta blight affects other legume crops and comparative genome analysis has revealed substantial synteny among legume species, these data offer a foundation for unraveling the mechanism of resistance in legume species beyond chickpea.
- The raw sequence reads can be further processed and analyzed by researchers using their own bioinformatic algorithms for association studies that target various agronomic traits. The assignment of a GO term for a total of 273 genes can give researchers information about the function of other genes that could be associated with a different agronomic trait of interest.

2. Objective

The sequence reads data were used to identify the single nucleotide polymorphisms (SNPs) found in the chickpea populations, allowing the marker-trait association performed in our Original Research Article [1]. Moreover, these data can be further utilized to analyze various traits in both populations, as they also exhibited variation in agronomic characteristics such as flowering time or growth habit. The functional annotation data enhance the value of the candidate gene list reported in [1] by providing additional information on various Gene Ontology (GO) categories, including Biological Process, Cellular Component, and Molecular Function.

3. Data Description

3.1. Raw sequences (SRA data)

Fresh leaf tissue samples (30 mg of freeze-dried tissue) were collected from adult plants (four weeks old) in both parents and 84/81 individuals from two recombinant inbred line (RIL) populations, namely RIP8 and RIP10. The DNA from the samples underwent quality-check procedures (absorbance check: 260/280 ratio of 1.7–2.0 and 260/230 ratio of 2.0–2.2), and raw reads were obtained using tuneable genotyping-by-sequencing (tGBS) with the restriction enzyme Bsp1286I on an Illumina HiSeq X instrument. A quality trimming process was performed to remove low-quality regions at the beginning and end of the reads. Detailed information about the total amount of data generated and the maximum, minimum, and mean number of reads before quality trimming can be found in Tables 1 and 2 for RIP8 and RIP10, respectively. Tables 3 and 4 represent data regarding the quality trimming process for RIP8 and RIP10 after removing those reads with an error rate $\geq$ 3 %. The sequencing data are accessible in the NCBI public repository [Datasets 1,2] under the accessions SRP441161 (RIP8) and SRP441763 (RIP10).

Table 1
Summary of tGBS Reads for RIP8 (ILC3279 × WR315).

Description	No.
Number of individuals	86
Total Reads	2 × 180,737,915
Average Reads per Sample	2 × 1,882,686
Median Reads per Sample	2 × 1,538,578
Minimum Reads per Sample	2 × 446,407
Maximum Reads per Sample	2 × 7,757,117
25 % Pctl. Reads per Sample	2 × 1,084,603

Table 2
Summary of tGBS Reads for RIP10 (JG62 x ILC72).

Description	No.
Number of individuals	83
Total Reads	2 × 187,964,419
Average Reads per Sample	2 × 1,957,962
Median Reads per Sample	2 × 1,558,894
Minimum Reads per Sample	2 × 2,052
Maximum Reads per Sample	2 × 8,450,979
25 % Pctl. Reads per Sample	2 × 1,083,064

Table 3
Quality trimming Summary for RIP8 (ILC3279 x WR315).

	Processed reads			Quality trimmed reads		
	No. Reads	Base Pairs (bp)	Length (bp)	No. Reads	Base Pairs (bp)	Length (bp)
Sum	2 × 180,737,915	2 × 25,817,750,485	143	2 × 180,643,326 (99.9 %)	2 × 25,555,916,837 (99.0 %)	141
Average	2 × 1,882,686	2 × 268,934,900	143	2 × 1,881,701 (99.9 %)	2 × 266,207,467 (99.0 %)	141
Median	2 × 1,538,578	2 × 218,147,147	143	2 × 1,537,967 (100 %)	2 × 216,274,763 (99.1 %)	142

Table 4
Quality trimming Summary for RIP10 (JG62 x ILC72).

	Processed reads			Quality trimmed reads		
	No. Reads	Base Pairs (bp)	Length (bp)	No. Reads	Base Pairs (bp)	Length (bp)
Sum	2 × 187,964,419	2 × 26,826,621,167	143	2 × 187,860,604 (99.9 %)	2 × 26,561,828,135 (99.0 %)	141
Average	2 × 1,957,962	2 × 279,443,970	143	2 × 1,956,881 (99.9 %)	2 × 276,685,709 (99.0 %)	141
Median	2 × 1,558,894	2 × 223,538,800	143	2 × 1,558,133 (100 %)	2 × 221,355,445 (99.0 %)	142

3.2. Functional annotation data

The SNPs obtained from the alignment of raw reads with the reference genome defined a robust candidate region in Ca4 (3.52–8.20 Mb) which was functionally annotated with BLAST2GO software. Fig. 1 shows the distribution of GO terms for the three GO categories (Biological Process, Molecular Function, and Cellular Component). The annotation file provided information regarding the functional annotation of 273 gene sequences in the candidate region and can be accessed under the DOI 10.17632/z34fks5sr2.1 in the Mendeley Data repository [Dataset 3].

4. Experimental Design, Materials and Methods

4.1. Plant material and SRA data generation

Two RIL populations segregating for Ascochyta blight resistance were developed from the crosses ILC3279 (resistant) x WR315 (susceptible), and JG62 (susceptible) × ILC72 (resistant), and were called RIP8 and RIP10, respectively [1]. Tissue from fresh leaves was collected from adult plants (30 mg of freeze-dried tissue from plants with four weeks of development), and samples from parents and the RILs in both populations were quality-checked and sent to Freedom Markers for whole genome sequencing (WGS). The tGBS protocol [2] used the restriction enzyme Bsp1286I for the obtention of raw reads via Illumina HiSeq X instrument. Raw sequence data underwent debarcoding, ensuring that reads were correctly aligned to the corresponding sample. Low-quality regions at the beginning and end of the reads were removed after quality trimming (PHRED quality score $\geq$ 15; error rate $\geq$ 3 %).

4.2. Functional annotation of a candidate fungal resistance region

Raw reads were aligned to the CDC Frontier Reference Genome [3] (ASM33114v1) using GSNAP aligner v2019-09-12 [4] for the obtention of SNPs, which were used for defining candidate fungal resistance regions via direct phenotype-genotype association. A robust candidate

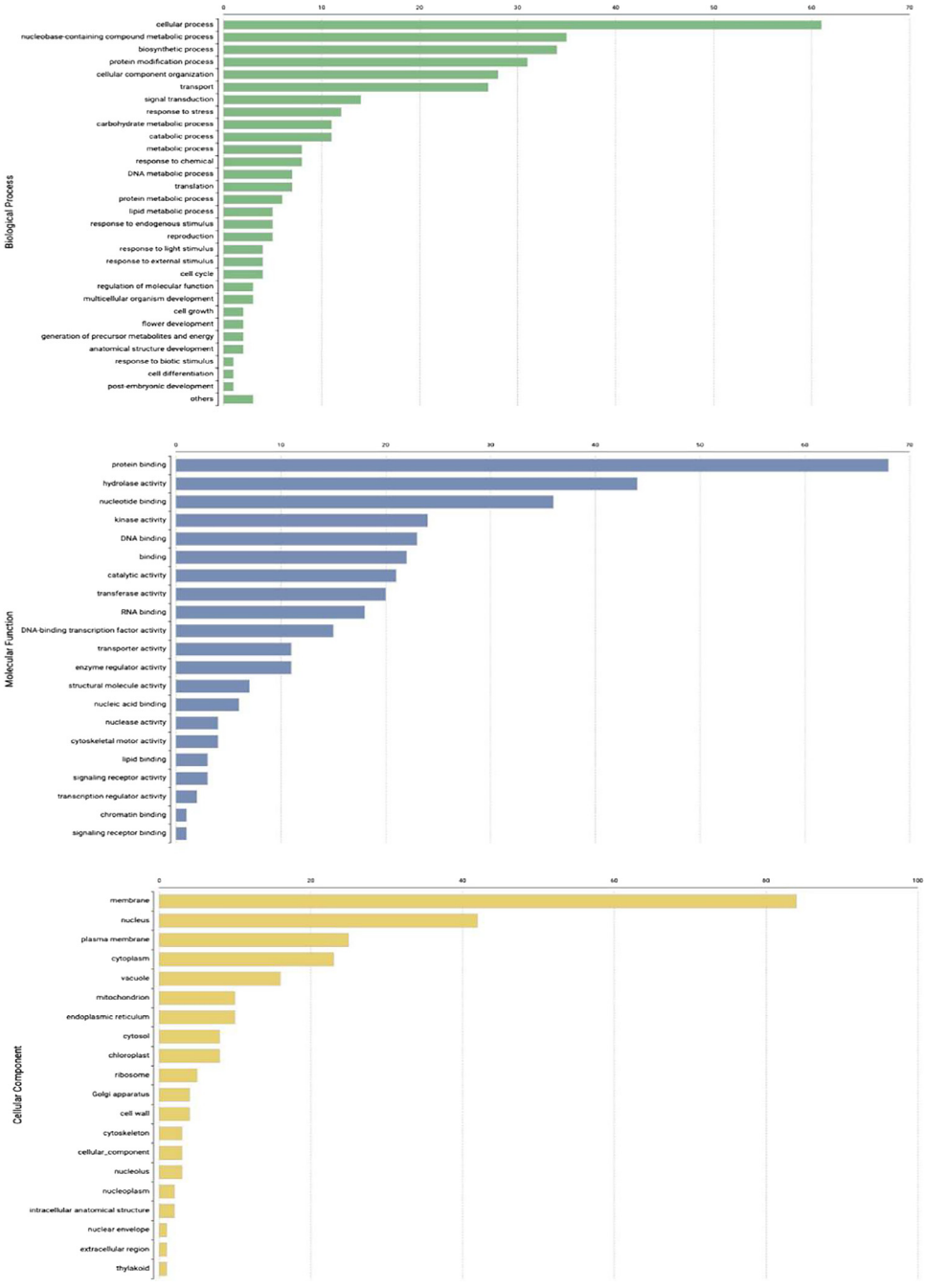

Fig. 1. GO category distribution of BLAST2GO annotated genes in the candidate fungal resistance region (3.52–8.20 Mb; Ca4).

region in Ca4 (3.52–8.20 Mb) was associated in both populations. This region was downloaded from Genome Data Viewer in NCBI [5] and functionally annotated with BLAST2GO software v6.0.3 [6] using default parameters. A complementary InterProScan analysis was performed in BLAST2GO for a more detailed description of the functional annotation.

Ethics Statements

This work and the data generated do not involve human subjects, animal experiments or data collected from social media platforms.

Data Availability

Annotation of candidate fungal resistance region (Original data) (SRA NCBI).
RIP10 in chickpea (Original data) (SRA NCBI).
RIP8 in chickpea (Original data) (SRA NCBI).

CRediT Author Statement

Alejandro Carmona: Formal analysis, Investigation, Data curation, Writing – original draft; **Patricia Castro:** Validation, Investigation, Writing – review & editing, Supervision; **Adrian Perez-Rial:** Investigation, Writing – review & editing, Visualization; **Jose V. Die:** Investigation, Writing – review & editing, Supervision, Funding acquisition.

Acknowledgments

This work was supported by Funding from the programs EU Marie Sklodowska-Curie Action H2020-MSCA-IF-2018 (grant ref. 844431, FORCE) and MCIN/AEI/10.13039/501100011033 (Ref. RYC2019-904 028188-I).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Carmona, J. Rubio, T. Millan, J. Gil, J.V. Die, P. Castro, Four haplotype blocks linked to Ascochyta blight disease resistance in chickpea under Mediterranean conditions, Front. Plant Sci. 14 (2023) 1183287, doi:10.3389/fpls.2023.1183287.
[2] A. Ott, S. Liu, J.C. Schnable, C. Yeh, K.S. Wan, P.S. Schnable, tGBS® genotyping-by-sequencing enables reliable genotyping of heterozygous loci, Nucleic Acids Res. 45 (21) (2017) e178, doi:10.1093/nar/gkx853.
[3] R. Varshney, C. Song, R. Saxena, et al., Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement, Nat. Biotechnol. 31 (2013) 240–246, doi:10.1038/nbt.2491.
[4] T.D. Wu, S. Nacu, Fast and SNP-tolerant detection of complex variants and splicing in short reads, Bioinformatics 26 (2010) 873–881, doi:10.1093/bioinformatics/btq057.
[5] S.H. Rangwala, A. Kuznetsov, V. Ananiev, et al., Accessing NCBI data using the NCBI sequence viewer and genome data viewer (GDV), Genome Res. 31 (1) (2021) 159–169 http://www.genome.org/cgi/doi/10.1101/gr.266932.120.
[6] A. Conesa, S. Götz, Blast2GO: a comprehensive suite for functional analysis in plant genomics, Int. J. Plant Genom. 2008 (2008) (2008) 619832, doi:10.1155/2008/619832.
[7] Universidad de CordobaRIP8 in Chickpea, National Center for Biotechnology Information, 2023 https://trace.ncbi.nlm.nih.gov/Traces/?view=study&acc=SRP441161.
[8] Universidad de CordobaRIP10 in Chickpea, National Center for Biotechnology Information, 2023 https://trace.ncbi.nlm.nih.gov/Traces/?view=study&acc=SRP441763.
[9] [Dataset 3] C.J. Alejandro, Functional Annotation of a Candidate Ascochyta Blight Resistance Region in Chickpea Mendeley Data V1, 2023 https://data.mendeley.com/datasets/z34fks5sr2/1 , doi:10.17632/z34fks5sr2.1.