ISSN No:-2456-2165

Analysis of Transcriptome datasets of Jatropha Curcas Leaves Infected with Colletotrichum Gloeosporioides Using RNA-Seq

KrunalPawar, Dr. SiddhivinayakBarve, Dr. Neeta S. Madan
Kelkar Education Trust's Scientific Research Centre,
V.G.Vaze College Campus, Mithagar Road, Mulund (East),
Mumbai, Maharashtra 400081, India,

Abstract:- Even though jatropha holds enormous potential as a biodiesel feedstock, there are still significant obstacles that can be overcome before it can be economically viable in many parts of the world. Jatropha curcas develops anthracnose due to Colletotrichumgloeosporioides and it affects each genotype in a different way. A gene-expression profiling study was performed with total RNA from two cultivars of Jatropha curcas that had and did not have a Colletotrichumgloeosporioides inoculation line in order to identify potential candidate genes involved in pathogen-plant interactions. There were 213 significant differential gene expressions between the wild type 9-1 and RJ127 cultivar lines inoculated with Colletotrichumgloeosporioides. A comprehensive analysis of these genes provided information regarding the signaling system, hormone biosynthesis and regulation, transcription regulation, and ubiquitin-mediated proteolysis. Gene Ontology (GO:0006952) annotations were found for nine genes associated with defense response, six orthologous genes were associated wi**th pathogen-plant interactions, and 27 transcription factors were associated with fungus responses. The expression levels of genes varied among cultivars when infected with Colletotrichumgloeosporioides, according to gene expression analysis. Transcriptome analyses were used to diagnose candidate genes for transcription factors and structural genes that may be involved in plant-pathogen interactions in Jatropha curcas leaves of the cultivated genotypes. These analyses provided new insights into the defense mechanisms controlling infection associated with cultivar Jatropha curcas 9-1. The structural genes and regulators identified here can also be used as efficient genetic markers for the selection of cultivars with excessive defense mechanisms.**

Keywords:-*Jatropha*; *C. gloeosporioides*; *NGS*; *RNA-Seq*; *DGEs*; *biodiesel*; NGS;

I. INTRODUCTION

Biofuels existed long before vehicles were invented. The relatively modest gas and diesel prices have kept them at the periphery for quite some time. The increase in oil prices, as well as worldwide initiatives to combat the negative effects of environmental change, have lent fresh urgency to the search for spotless, sustainable [1]. Jatropha curcas (J. curcas) is a species endemic to Mexico and Central America and a member of the Euphorbiaceae family [2]. The tropical plant has extraordinary potential to produce biofuel for energy crisis mitigation, environmental management, and sustainable production management in the future [3]. A moderately small level genome ~320.5 MB for the RefSeq version of J.Curcas, available in the NCBI genome J.curcas Annotation Release 101, which syncs with the J. Curcas database (JCDB) and contains gene functional annotation, interaction networks, and expression matrices for use in J. Curcas functional genomics research [4][5][6][7][8][9][10]. A case of anthracnose was reported on the J.Curcas plant in 2010 and 2011, which was the first known case of the disease caused by Colletotrichumgloeosporioides (C. gloeosporioides) on J.Curcas in Korea [11]. C. gloeosporioides causes anthracnose, a serious Curcas disease, causing damage to leaves, stems, and fruits, which adversely affects the production and quality of seeds. Therefore, there is a limitation to the improvement of the biofuel industry due to parasitic diseases [12].

Improvements in innovation and strategies have allowed fast and generally reasonable RNA-Seq dataset generation over the decades [13]. As a result, RNA-Seq in various model frameworks has become an undeniably well-known strategy for transcriptome-wide changes. The development of information examination has propelled improvements in open access programs for quality control, mapping, and differential expression (DE) investigation [14]. NCBI SRA database contains raw data that can be downloaded from the NCBI distributed research archive, which gradually transfers distributed information to openly accessible files. Researchers can now examine RNA-Seq information with fewer obstacles, allowing them to make and test theories with RNA-Seq data that have progressively different logical foundations.

The present study compares the transcriptomic profiles of four Jatropha samples from the NCBI BioProjects PRJNA254929 and PRJNA254930 data using the Illumina HiSeq 2500 platform to identify differentially expressed genes related to resistance and susceptibility of Jatropha to C.gloeosporioides. Based on these results, we can identify some candidate genes for the resistance of two different Jatropha genotypes against C.gloeosporioides, and gain insight into the mechanism of Anthracnose disease on the two genotypes of J. Curcas.

ISSN No:-2456-2165

II. MATERIALS AND METHODS

A. Raw Data Retrieval From Rna Sequencing

This study used data from NCBI BioProject PRJNA254929 and PRJNA254930. To investigate differentially expressed genes, fresh leaf samples of J. curcas strains RJ127 and Cultivar 9-1 had been artificially infected or not with C. gloeosporioides. Following the manufacturer's recommended protocol, RNA was extracted from control, induced tissue of susceptible and tolerant lines after 2hr, 24hr, 96hr, and 144hr of infection. To retrieve data from NCBI, we used SRAToolsKit 2.10.4.

B. Raw Data Processing And Mapping Of Reads To The Reference Genome

To analyze the raw reads, we preprocessed them in fastq format before analyzing them further. Trimming the raw fastq reads was automated using the Trim Galore, which is a wrapper tool around FastQC and Cutadapt to consistently apply quality and adapter trimming. Default settings include trimming all paired-end reads with a Phred score of less than 20, trimming only paired-end reads, and introducing an Illumina adaptor sequence for any sequence with a 5 base pair overlap. This filtered high-quality read data was used for all downstream analyses. STAR aligner with default settings was used to align trimmed reads to J. curcas genome (assembly JatCur_1.0).

C. Identification And Annotation of Differentially Expressed Genes (Degs)

Gene model GTF files were generated from combined transcriptomic sequences by Cufflinks. Each transcript's FPKM value was calculated across all samples. For each sample, we used Cuffmerge to obtain a final GTF file. Using Cuffdiff, we calculated the differential abundance of transcripts among different samples/sites. In order to identify the significance of an expression difference, we used p-values of ≤ 0.05 and log2Ratios greater than 1. Our analysis of significant genes with Gene Ontology (GO) uses Uniprot database after obtaining locusIDs from Cuffdiff result with uniprotIDs from the bioDBnet server.

III. RESULTS

A. Transcription Profiles Of Resistant And Susceptible Cultivars Of J. Curcus Infected With C. Gloeosporioides

In this study, two cultivars of J. curcus reacted contrastingly to the pathogen C. gloeosporioides. Four libraries were developed from C. gloeosporioides infected leaf tissues and non-inoculated leaf tissues to determine which cultivars are receptive to C. gloeosporioides and to design quality articulation designs (Table 1).

TABLE 1. CONTROLS AND SAMPLES INFORMATION OF J. CURCUS CULTIVARS

Abbreviation	Tissue	Condition	Description
Pool A	Leaf	9-1_Control (resistant)	A Control was formed by pooling uninfected genotypes of J. Curcas and tissue from tolerant lines 2, 24, 96, and 144 hours after infection.
Pool B	Leaf	9-1 induced (resistant)	The induced tissue of tolerant lines collected after 2, 24, 96, and 144 hours after infection was pooled as B induced with J. Curcas genotypes showing tolerance to C. gloeosporioides.
Pool C	Leaf	RJ127 control (susceptible)	A control group was made up of uninfected J Curcas genotypes, and tissue of susceptible lines was collected after 2, 24, 96, and 144 hours of infection.
Pool D	Leaf	RJ127 induced (susceptible)	The induced tissue of tolerant lines collected after 2, 24, 96, and 144 hours after infection were pooled as D induced with J. Curcas genotypes showing tolerance to C. gloeosporioides.

From total RNA isolates, cDNA libraries were synthesized to analyze the transcriptome of J. curcas. Illumina sequencing platforms were subsequently used to sequence cDNA libraries. Following the filtering and trimming of low-quality reads and adaptors, clean reads were mapped against the NCBI genome JatCur_1.0. Over 90% of the reads were uniquely mapped. The raw data and the QC processed data are available in table 2. Short reads assembling software Star aligner was used to assemble these clean reads. The FPKM method was used to calculate the expression levels of 25,515 genes based on the total mapped reads. An indication of significant-quality from cuffdiff yields based on p values and FDRs after Benjamini-Hochberg correction was used to identify response against C. gloeosporioide.

Table 2. Statistics of Sequence Output Of J. Curcas Cultivars

Samples	Raw reads	Raw bases	Clean Reads	Clean Bases	Valid ratio (base)	GC content (%)	Uniquely mapped (%)
Pool A	45280396	4573319996	44402527	4377837709	95.7%	43%	92.89%
Pool B	37461637	3783625337	36842449	3632603350	96.0%	43%	92.40%
Pool C	51169379	5168107279	49898983	5003568089	96.8%	43%	94.12%
Pool D	38496992	3888196192	37317006	3678977234	94.6%	43%	91.04%

B. Dges Profiles In Response To C. Gloeosporioides Infection

➢ *Profile Of Gene Expression In Control Cultivars Versus C. Gloeosporioides Infected Cultivars*

A total of 160 significant DGEs were identified between the resistance control the 9-1 cultivar and its inoculated libraries (PoolA vs PoolB). Out of 104 annotated genes, 101 were up-regulated while 3 were down-regulated. In the comparison between RJ127 and the inoculated libraries, 116 DGEs were identified out of 83 annotated DGEs, of which 77 were upregulated and 6 were downregulated (Table 3).

➢ *Dge Profiles Of Both Cultivars In Control Samples*

In the comparison between the resistant control and susceptible control, 94 significant DGEs were found out of 58 annotated DGEs, 46 of which were upregulated and 12 were downregulated in the RJ127 cultivar sample compared to the 9-1 cultivar sample(Table 3).

➢ *Dges PROFILES OF BOTH CULTIVARS WITH C. GLOEOSPORIOIDES INFECTION*

Based on the comparison of the inoculated 9-1 cultivar and the RJ127 inoculated cultivar, 88 significant DGEs were identified out of 63 annotated DGEs, in which there were 53 up-regulated genes and 10 down-regulated genes(Table 3). Overall, the number of DGEs was significantly higher in the 9-1 cultivar than RJ127 in Response to C. gloeosporioides Infection.

C. Gene Ontology (Go) Annotation

An outline of the utilitarian categories associated with infection-associated DGEs was developed using the Gene Ontology (GO) assignments within the UniProt database [*Fig. 1b*]. A total of 78 annotated DGEs associated with 9-1 and its inoculated libraries had been classified in the GO database, whereas 63 annotated DGEs associated with RJ127 and its inoculated libraries had been classified in the GO database. A venn diagram [*Fig. 1a*]that depicted the DGEs results from all 4 pools made it evident that unique as well as shared DGEs were distinguishable among, and between, sets.

TABLE 3. DGES IN DIFFERENT GROUPS

Gene names	**UniProt Accession**	**log2(fold change) PoolA vs PoolB**	**Regulation (PoolA vs PoolB)**	**log2(fold_change) PoolA vs PoolC**	**Regulation (PoolA vs PoolC)**	**log2(fold_change) PoolC vs PoolD**	**Regulation (PoolC vs PoolD)**
JCGZ_14311	A0A067K8D2	-2.91851	Down	---	---	---	---
---	Q8VYU0	-1.71861	Down	---	---	---	---
JCGZ_15401	A0A067LNK5	-1.62426	Down	---	---	---	---
JCGZ_21348	A0A067JAT4	1.42747	UP	---	---	---	---
JCGZ_07055	A0A067KNU6	1.44722	UP	---	---	---	---
JCGZ_17504	A0A067JQW3	1.51161	UP	---	---	---	---
WRKY57 JCGZ_03308	S5CFW3; S5CS97	1.54828	UP	---	---	---	---
---	C9E0E8	1.55778	UP	---	---	---	---
JCGZ_18383	A0A067K4S0	1.56653	UP	---	---	---	---
JCGZ_25752	A0A067JVT3	1.57314	UP	---	---	---	---
JCGZ_10666	A0A067KG70	1.63609	UP	---	---	---	---
JCGZ_24672	A0A067L0C9	1.6536	UP	---	---	---	---
JCGZ_18911	A0A067JV40	1.65432	UP	---	---	---	---
JCGZ_23324	A0A067JHR5	1.68004	UP	---	---	---	---

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_16744	A0A067L4V4	1.68064	UP	---	---	---	---
JCGZ_11537	A0A067K4L2	1.70613	UP	---	---	---	---
JCGZ_26283	A0A067JF33	1.74402	UP	---	---	---	---
JCGZ_16041	A0A067LC07	1.74905	UP	---	---	---	---
JCGZ_11901	A0A067KF78	1.84629	UP	---	---	---	---
WRKY51 JCGZ_14353	S5CKC9	1.8522	UP	---	---	---	---
JCGZ_02957	A0A067L1F8	1.85487	UP	---	---	---	---
JCGZ_25208	A0A067L3L7	1.87466	UP	---	---	---	---
JCGZ_22456	A0A067K1V9	1.88759	UP	---	---	---	---
JCGZ_26850	A0A067L091	1.89331	UP	---	---	---	---
JCGZ_02619	A0A067L678	1.93896	UP	---	---	---	---
JCGZ_13272	A0A067K8C1	1.97321	UP	---	---	---	---
JCGZ_09700	A0A067LAG1	1.97504	UP	---	---	---	---
JCGZ_02050	A0A067KV70	1.99654	UP	---	---	---	---
LOX2 JCGZ_16708	A0A067L4M8	2.00344	UP	---	---	---	---
JCGZ_08745	A0A067KIP6	2.08499	UP	---	---	---	---
JCGZ_09271	A0A067KIY5	2.10078	UP	---	---	---	---
JCGZ_25773	A0A067JJH0	2.11164	UP	---	---	---	---
WRKY47 JCGZ_14332	S5CKC4	2.11916	UP	---	---	---	---
WRKY08 JCGZ_00400	M9TNQ5	2.20164	UP	---	---	---	---
JCGZ_07439	A0A067KFV0	2.21122	UP	---	---	---	---
JCGZ_03925	A0A067L6T7	2.23813	UP	---	---	---	---
JCGZ_24882	A0A067L0X2	2.24577	UP	---	---	---	---
JCGZ_02823	A0A067LCX8	2.27533	UP	---	---	---	---
JCGZ_19400	A0A067JZK3	2.27653	UP	---	---	---	---

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_02588	A0A067KX45	2.30864	UP	---	---	---	---
JCGZ_01012	A0A067L449	2.31447	UP	---	---	---	---
JCGZ_19656	A0A067JYJ8	2.3289	UP	---	---	---	---
WRKY28 JCGZ_07091	S5CFS7	2.33565	UP	---	---	---	---
JCGZ_01716	A0A067JTW8	2.34832	UP	---	---	---	---
JCGZ_08703	A0A067KX47	2.3508	UP	---	---	---	---
JCGZ_09328	A0A067KRX2	2.37089	UP	---	---	---	---
JCGZ_11530	A0A067KH50	2.41079	UP	---	---	---	---
JCGZ_01715	A0A067JJI5	2.42259	UP	---	---	---	---
JCGZ_07968	A0A067LEP7	2.5398	UP	---	---	---	---
JCGZ_00441	A0A067JTI6	2.62832	UP	---	---	---	---
PAL4 JCGZ_07452	A0A067KCD4	2.71673	UP	---	---	---	---
JCGZ_02813	A0A067LCW8	2.92729	UP	---	---	---	---
JCGZ_12474	A0A067KAG7	3.0316	UP	---	---	---	---
JCGZ_02113	A0A067L6J1	3.16631	UP	---	---	---	---
JCGZ_24951	A0A067KXS0	3.36643	UP	---	---	---	---
JCGZ_14447	A0A067JXI1	3.38977	UP	---	---	---	---
JCGZ_15016	A0A067L9M0	3.39243	UP	---	---	---	---
JCGZ_16811	A0A067LH44	3.45701	UP	---	---	---	---
JCGZ_20671	A0A067JZT2	3.57009	UP	---	---	---	---
JCGZ_15163	A0A067LDL0	3.61511	UP	---	---	---	---
JCGZ_18946	A0A067JV70	3.99563	UP	---	---	---	---
JCGZ_09023	A0A067KTI1	4.2537	UP	---	---	---	---
JCGZ_07576	A0A067KNY6	4.52988	UP	---	---	---	---
JCGZ_20670	A0A067K136	4.77342	UP	---	---	---	---
JCGZ_24689	A0A067L877	5.12327	UP	---	---	---	---

ISSN No:-2456-2165

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_16815	A0A067L8E6	5.39239	UP	---	---	---	---
JCGZ_20633	A0A067JNL7	3.04133	UP	---	---	9.04184	Up
JCGZ_18942 JCGZ_18944	A0A067JVF0	3.73551	UP	---	---	4.00984	Up
JCGZ_21144	A0A067JD40	3.538	UP	---	---	3.70795	Up
JCGZ_05111	A0A067L2R2	2.27356	UP	---	---	3.53599	Up
JCGZ_22904	A0A067K243	2.95129	UP	---	---	3.12113	Up
JCGZ_25080	A0A067JKK3	2.70873	UP	---	---	2.85497	Up
CLA1 JCGZ_23012	A0A067LHX1	2.29141	UP	---	---	2.84838	Up
JCGZ_12476	A0A067KIB2	1.66776	UP	---	---	2.46842	Up
JCGZ_00615	A0A067JGC3	2.73234	UP	---	---	-2.77065	Down
JCGZ_07518	A0A067KCS8	1.83993	UP	---	---	-3.28346	Down
JCGZ_04087	A0A067L3V1	2.21339	UP	---	---	-3.86326	Down
JCGZ_17396	A0A067LMZ1	2.91394	UP	2.17882	Up	-2.69737	Down
JCGZ_15810	A0A067L2E4	2.11219	UP	2.45413	Up	-2.28623	Down
JCGZ_21418	A0A067JB04	1.49165	UP	1.55317	Up	---	---
JCGZ_23841	A0A067LFL5	1.60033	UP	1.4277	Up	---	---
WRKY07 JCGZ_16684	M9TGB8	1.74768	UP	1.61453	Up	---	---
JCGZ_16894	A0A067L5A5	1.77871	UP	2.45938	Up	---	---
JCGZ_18287	A0A067JZG6	1.78904	UP	1.76848	Up	---	---
JCGZ_26050	A0A067JRT8	1.80388	UP	1.71348	Up	---	---
JCGZ_20215	A0A067JTU3	1.81124	UP	1.64682	Up	---	---
JCGZ_16315	A0A067L7T6	1.82742	UP	1.71468	Up	---	---
JCGZ_11380	A0A067KGN0	1.84507	UP	1.61488	Up	---	---
JCGZ_01313	A0A067LC73	1.84923	UP	1.82433	Up	---	---
JCGZ_08772	A0A067KM96	1.87745	UP	1.68813	Up	---	---

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_21718	A0A067JBV0	1.91043	UP	1.71733	Up	---	---
JCGZ_20572	A0A067JN47	2.01654	UP	1.51071	Up	---	---
JCGZ_15872	A0A067LAD0	2.0847	UP	1.99865	Up	---	---
JCGZ_21761	A0A067JMR1	2.15307	UP	2.17196	Up	---	---
JCGZ_15809	A0A067KZ11	2.16436	UP	2.72198	Up	---	---
JCGZ_08268	R4NHR7	2.30741	UP	1.92321	Up	---	---
WRKY27 JCGZ_26990	S5CKF6	2.31891	UP	1.71705	Up	---	---
JCGZ_09675	A0A067LAD6	2.32451	UP	2.38589	Up	---	---
JCGZ_03451	A0A067L5Z5	2.35703	UP	1.87669	Up	---	---
JCGZ_15808	A0A067KYX4	2.41251	UP	2.56576	Up	---	---
GRAS17 JCGZ_00787	A0A067L3E6	2.53187	UP	1.85568	Up	---	---
JCGZ_08750	A0A067KIQ1	3.10243	UP	2.36395	Up	---	---
JCGZ_04802	A0A067L249	3.42307	UP	3.20016	Up	---	---
WRKY54 JCGZ_20472	S5CKD4	4.05806	UP	3.42691	Up	---	---
JCGZ_21250	A0A067JMQ3	---	---	-4.62991	Down	---	---
JCGZ_00555	A0A067JG74	---	---	-4.54832	Down	---	---
JCGZ_06967	A0A067L0K1	---	---	-3.55157	Down	---	---
JCGZ_23980	A0A067JM79	---	---	-3.05669	Down	---	---
JCGZ_05646	A0A067LIV1	---	---	-2.63614	Down	---	---
DHAR JCGZ_06589	R9QAK3	---	---	-2.40991	Down	---	---
JCGZ_26055	A0A067JRU1	---	---	-2.18077	Down	---	---
JCGZ_26788	A0A067LBB3	---	---	-2.16355	Down	---	---
JCGZ_24918	A0A067L9X5	---	---	-1.97304	Down	---	---
JCGZ_00267	A0A067LD96	---	---	-1.86521	Down	---	---
JCGZ_16709	A0A067L4S5	---	---	-1.6963	Down	---	---

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_16557	A0A067K293	---	---	-1.5609	Down	---	---
JCGZ_03320	A0A067JQ58	---	---	1.54494	Up	---	---
JCGZ_08815	A0A067KIW7	---	---	1.69431	Up	---	---
JCGZ_20298	A0A067K4Y8	---	---	1.73078	Up	---	---
JCGZ_25601	A0A067JWD4	---	---	1.73781	Up	---	---
JCGZ_18739	A0A067K1G2	---	---	1.75928	Up	---	---
JCGZ_06225	A0A067KZA6	---	---	1.92552	Up	---	---
JCGZ_13243	A0A067KBE6	---	---	1.95642	Up	---	---
JCGZ_26795	A0A067L047	---	---	2.00783	Up	---	---
JCGZ_16640	A0A067JZ78	---	---	2.2488	Up	---	---
JCGZ_22457	A0A067JQK6	---	---	2.28603	Up	---	---
JCGZ_18325	A0A067KC42	---	---	2.30001	Up	---	---
JCGZ_03363	A0A067JD10	---	---	2.57497	Up	---	---
JCGZ_23064	A0A067JH55	---	---	2.72935	Up	---	---
JCGZ_17563	A0A067K2B0	---	---	2.76153	Up	---	---
JCGZ_02540	A0A067L4T7	---	---	2.7958	Up	---	---
JCGZ_24823	A0A067L9Q0	---	---	3.19224	Up	---	---
JCGZ_08894	A0A067KK72	---	---	3.78448	Up	---	---
JCGZ_02198	A0A067L6V1	---	---	4.03899	Up	---	---
RHF	G9DD71	---	---	---	---	-2.88633	Down
JCGZ_25989	A0A067JH80	---	---	---	---	2.23663	Up
JCGZ_21344	A0A067JDN5	---	---	---	---	2.23891	Up
JCGZ_12091	A0A067KLX7	---	---	---	---	2.25554	Up
JCGZ_26104	A0A067JHM0	---	---	---	---	2.30378	Up
JCGZ_18967	A0A067JVH2	---	---	---	---	2.43708	Up
JCGZ_26020	A0A067JRP4	---	---	---	---	2.47485	Up
JCGZ_11478	A0A067K4N4	---	---	---	---	2.49398	Up

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_10173	A0A067LG90	---	---	---	---	2.50453	Up
JCGZ_00663	R4N5K2	---	---	---	---	2.55145	Up
JCGZ_21093	A0A067K1C8	---	---	---	---	2.58763	Up
JCGZ_21103	A0A067K1D9	---	---	---	---	2.63856	Up
JCGZ_06867	R4NFR5	---	---	---	---	2.67368	Up
JCGZ_07779	A0A067KGW9	---	---	---	---	2.69994	Up
JHL06P13.16 JCGZ_06164	E6NUC5	---	---	---	---	2.70591	Up
JCGZ_15477	A0A067LBY7	---	---	---	---	2.73333	Up
JCGZ_04776	A0A067KT36	---	---	---	---	2.81096	Up
JCGZ_16092	A0A067LB16	---	---	---	---	2.93859	Up
JCGZ_06603	A0A067LCF8	---	---	---	---	2.97213	Up
JCGZ_07815	A0A067KR22	---	---	---	---	2.98905	Up
JCGZ_26324	A0A067JI47	---	---	---	---	3.09123	Up
JCGZ_21766	A0A067JMR4	---	---	---	---	3.1424	Up
JCGZ_09009	A0A067KH55	---	---	---	---	3.19027	Up
JCGZ_07216	A0A067KMZ2	---	---	---	---	3.20441	Up
JCGZ_26879	A0A067L086	---	---	---	---	3.23976	Up
JCGZ_05632	A0A067LHU5	---	---	---	---	3.34604	Up
JCGZ_01195	A0A067LJ82	---	---	---	---	3.39571	Up
JCGZ_21899	A0A067JF68	---	---	---	---	3.42101	Up
SAM JCGZ_22479	K9UUN6	---	---	---	---	3.45222	Up
JCGZ_00759	A0A067KS90	---	---	---	---	3.58836	Up
JCGZ_12329	A0A067KA58	---	---	---	---	3.61842	Up
JCGZ_00868	A0A067KSB6	---	---	---	---	3.62234	Up
JCGZ_04632	A0A067L1K4	---	---	---	---	3.65871	Up

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_26412	A0A067JF88	---	---	---	---	3.68785	Up
JCGZ_18676	E2CXJ0	---	---	---	---	3.77831	Up
JCGZ_07557	A0A067KCN2	---	---	---	---	3.86772	Up
JCGZ_18517	A0A067K4H1	---	---	---	---	3.88523	Up
JCGZ_20694	A0A067JRS2	---	---	---	---	4.00472	Up
JCGZ_01427	A0A067L934	---	---	---	---	4.03707	Up
JCGZ_04841	A0A067KTE0	---	---	---	---	4.06179	Up
JCGZ_25536	A0A067JX93	---	---	---	---	4.08249	Up
JCGZ_13644	A0A067KDF3	---	---	---	---	4.11712	Up
JCGZ_15634	R4N7L5	---	---	---	---	4.11958	Up
JCGZ_15836	A0A067LBD6	---	---	---	---	4.15701	Up
JCGZ_10110	A0A067LN77	---	---	---	---	4.2811	Up
WRKY35 JCGZ_21688	S5CH50	---	---	---	---	4.30072	Up
JCGZ_24448	A0A067JZ76	---	---	---	---	4.40172	Up
JCGZ_11389	A0A067K7J2	---	---	---	---	4.45745	Up
JCGZ_10284	A0A067LPI9	---	---	---	---	4.5021	Up
JCGZ_02159	A0A067KVC5	---	---	---	---	4.62141	Up
JCGZ_20507	A0A067JMZ0	---	---	---	---	4.62345	Up
JCGZ_21644	A0A067JEH9	---	---	---	---	4.84044	Up
JCGZ_22279	A0A067K3C6	---	---	---	---	4.86939	Up
JCGZ_12482	A0A067K6Z9	---	---	---	---	4.90456	Up
JCGZ_03482	A0A067KUR0	---	---	---	---	4.96825	Up
JCGZ_07142	A0A067KBH4	---	---	---	---	5.24689	Up
JCGZ_22364	A0A067L5S7	---	---	---	---	5.31491	Up
JCGZ_00771	A0A067L4J4	---	---	---	---	5.35103	Up
JCGZ_21253	A0A067JAJ5	---	---	---	---	5.40441	Up

Gene names	UniProt Accession	log2(fold change) PoolA vs PoolB	Regulation (PoolA vs PoolB)	log2(fold_change) PoolA vs PoolC	Regulation (PoolA vs PoolC)	log2(fold_change) PoolC vs PoolD	Regulation (PoolC vs PoolD)
JCGZ_11762	A0A067K581	---	---	---	---	5.41423	Up
JCGZ_25081	A0A067JKU4	---	---	---	---	5.45967	Up
JCGZ_01044	A0A067KT39	---	---	---	---	5.91992	Up
JCGZ_05485	A0A067L9U2	---	---	---	---	6.05053	Up
JCGZ_25953	A0A067JE72	---	---	---	---	6.4195	Up
JCGZ_07402	A0A067KC92	---	---	---	---	6.64945	Up
JCGZ_21402	A0A067JAU2	---	---	---	---	6.7041	Up
JCGZ_07961 JCGZ_21019	A0A067LQR9	---	---	---	---	6.82357	Up
JCGZ_18289	D2D954	---	---	---	---	6.95559	Up
JCGZ_24827	A0A067L0T4	---	---	---	---	8.96372	Up
JCGZ_14551	A0A067K938	---	---	---	---	9.93983	Up

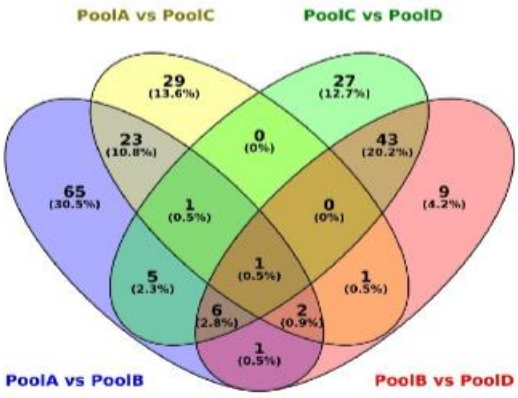

Figure a

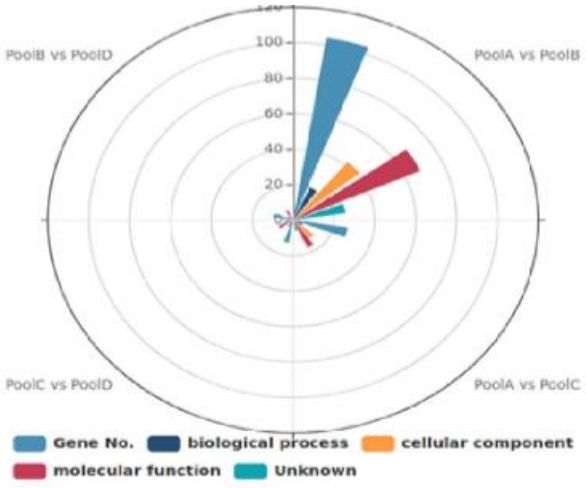

Figure b

Fig 1: Gene ontology distribution

Fig 1: Gene ontology distribution: Gene ontology distributions for DGEs of J curcas related to plant physiology of the molecular function, cellular component, and biological process. Figure 1a: Venn Diagram of DGEs genes in all 4 pools. Figure 1b: Gene ontology distributions for DGEs of J curcas. A gene may be associated with several GO terms.

D. Go Terms For Defense Response Gene Ontology (Go:0006952)

Gene Ontology term GO:0006952 was assigned to nine genes for defense response. Five DGEs were annotated when 9-1 cultivar and its inoculated libraries were compared. When comparing RJ127 cultivar and its inoculated libraries, 1 DGE was annotated with GO:0006952. A total of two DGEs were annotated for both controls and 4 genes for both infected cultivars (*Fig. 2*).

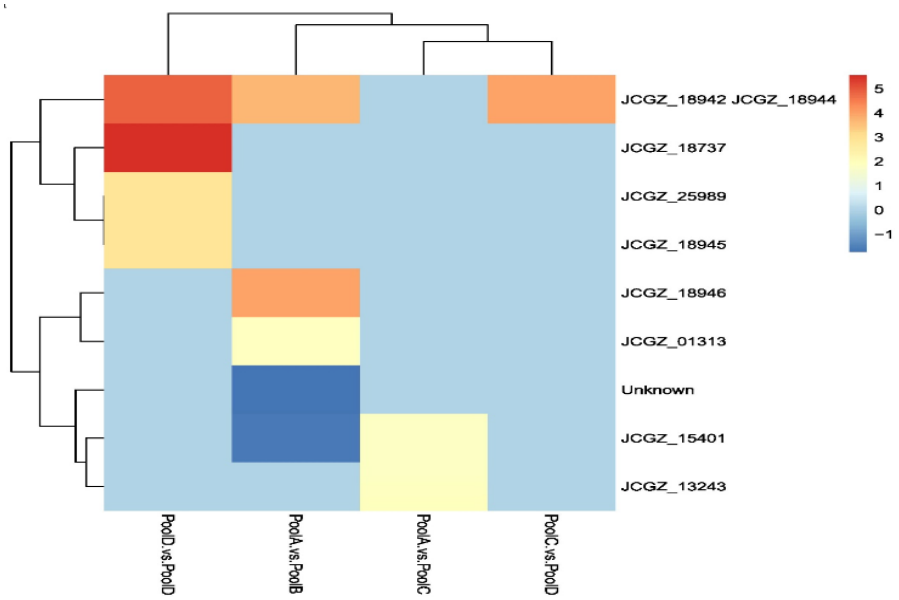

Fig. 2: DGEs of Gene Ontology [GO:0006952]:DGE distribution for GO:0006952 to understand the defense response against fungi.

E. Pathway Enrichment Analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified according to DGEs with a corrected $p < 0.05$. A total of 72 KEGG pathways were significantly enriched, of which 17 were enriched when 9-1 cultivar and its inoculated libraries were compared, 46 were enriched when RJ127 cultivar and its inoculated libraries were compared, while 33 were enriched in control samples for both cultivars were compared e.g. 9-1 cultivar compared with RJ127 cultivar without inoculation, while 31 were enriched both cultivars with C. gloeosporioides Infection i.e. 9-1 cultivar compared with RJ127 cultivar with inoculation were compared, while 5 are enriched all the time (*fig. 3*). DGEs related to 'Plant metabolism' (KO01100) had the highest representation, followed by DGEs for 'Secondary metabolite biosynthesis' (KO01110), then those for 'Plant-pathogen interactions' (KO04626), 'Plant hormone signal transduction' (KO04075), 'Phenylpropanoid biosynthesis' (KO00940), and 'MAPK signaling pathway – plant' (KO04016).

These findings suggest that J curcus has adapted a range of molecular defenses that prevail in disease, and C gloeosporioides contamination causes predominant changes in numerous metabolic pathways. The process is clearly complex and is likely to be regulated by a number of pathways, such as the plant-pathogen interaction.

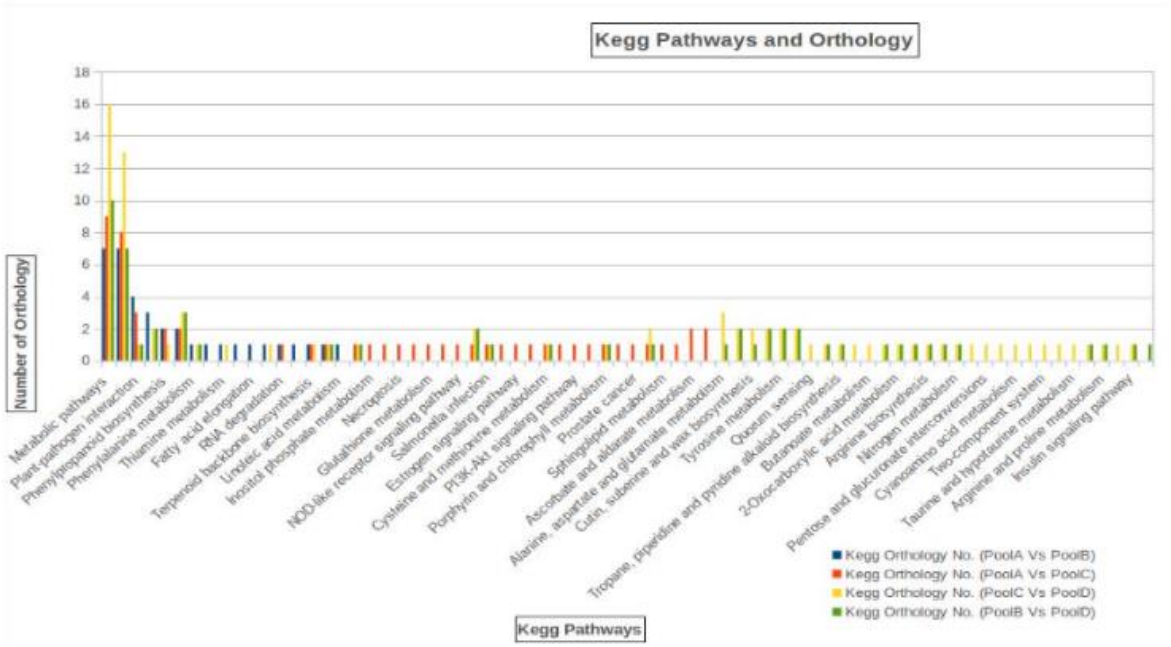

Fig. 3:KEGG pathways and Orthology

F. Dges Involved In Plant-Pathogen Interaction

Based on *fig. 4*, there were 4 gene orthology WRKY33, WRKY22, CML and KCS were significantly upregulated when 9-1 cultivar and its inoculated libraries were compared however CML is downregulated when RJ127 cultivar and its inoculated libraries were compared. The related KEGG map id is KO 04626.When we were looking resistant control with susceptible control comparison, there WRKY33 and RBOH were significantly upregulated and HSP90A was downregulated. While we compared cultivars with C. gloeosporioides Infection CML is upregulated in 9-1 cultivar.

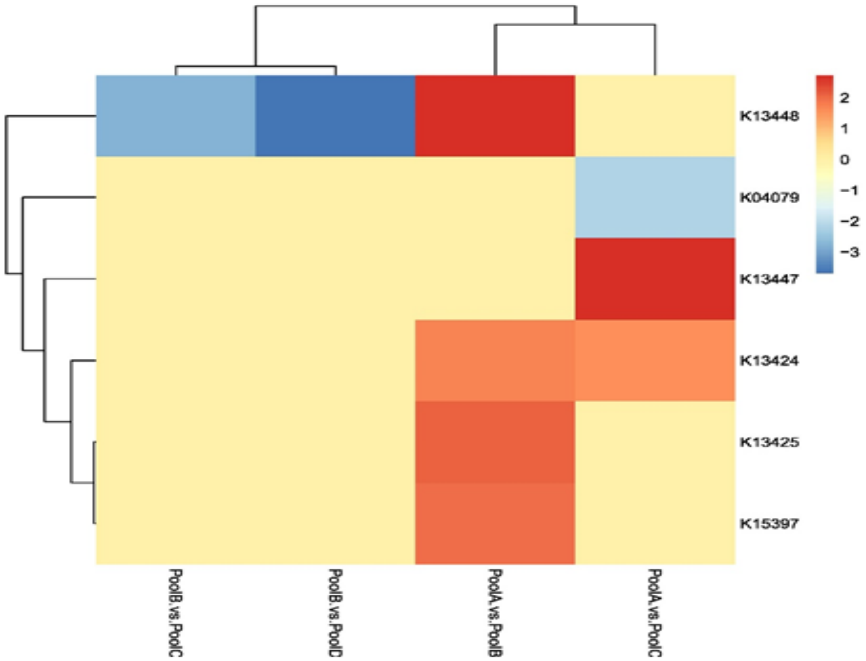

Fig. 4: KO:04626 related orthology regulation

G. Tfsrelated To C. Gloeosporioides Responses

TF families, such as WRKY (a conserved N-terminal sequence of WRKYGQK in conjunction with TF a Zn finger-like motif.), NAC (a (N-acetyl cysteine), AP2/ERF (APETALA2/ethylene with a Zn finger-like motif.), bHLH (basic (homeodomain-leucine helix-loop-helix), HSF (heat shock transcription factor) play vital roles in regulating plant resistance mechanisms under abiotic stress.

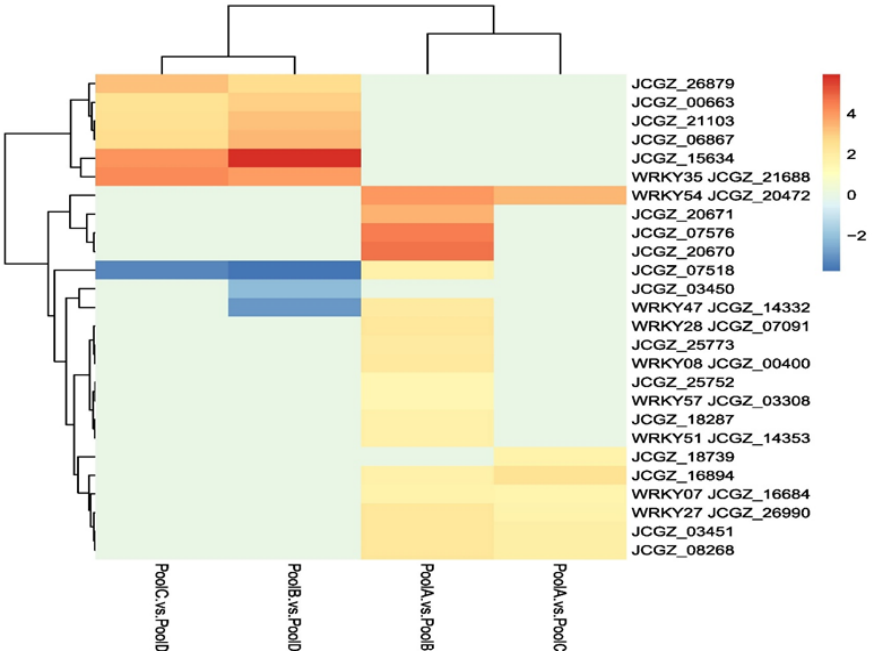

Fig. 5: TFsResponse Against C. gloeosporioides

In the present study, when 9-1 cultivar and its inoculated libraries were compared, 17 TFsfrom 4 families and one TF from an unknown family were detected. The RJ27 cultivar and its inoculated libraries showed seven TFsfrom four families. A comparison of resistant control with susceptible control revealed five TFsfrom three families and three from an unknown family. When inoculated libraries of the 9-1 cultivar were compared with inoculated libraries of the RJ127 cultivar, nine TFsfrom three families were identified. Figure 5 illustrates the expression of the related TFsin J.curcus induced by C.gloeosporioides.

IV. DISCUSSION

The pathogen Colletotrichum is one of the most important causes of anthracnose on trees, grasses, and a range of other types of plants[15]. There have been few studies of J.Curcus for disease resistance mechanisms against C gloeosporioides that could provide insight into anthracnose [16]. In this study, gene expression levels were compared between two cultivars of J.Curcas leaves subjected to C.gloeosporioides infection and those grown without the pathogen. A total of 308 DGEs were identified.

In plants, inborn resistance is activated by means of the reaction of PRRs (pattern-recognition receptors) to PAMPs (pathogen-associated molecular patterns), along these lines giving the principal line of inducible pathogen-related molecular PTI (PAMP-triggered immunity), further it triggers MAPK (Mitogen-activated protein kinase) signaling pathway that activates defense genes for antimicrobial compounds. In our study MAPK signaling pathway genes, those responsible for early/ late response of pathogen infection and cell death due to pathogen attack were upregulated in 9-1 J. curcas cultivar. In RJ-127 cultivar genes related to stress adaption signaling and defense response activated by ethylene signaling were upregulated when both cultivars were infected by C.gloeosporioides. Expansion in the cytosolic Ca2+ fixation also acts as a controller for the creation of receptive oxygen species and limited customized cell death/hypersensitive reaction. Ca2+ fixation genes from plant-pathogen interaction induced stomatal closure were upregulated in inoculated 9-1 cultivar.

Plants can incite transcription factors to DNA groupings as a technique for changing in accordance with natural and abiotic stresses [17]. Pathogen-reaction qualities, for example, WRKY, assume urgent jobs in plant reactions. The WRKY family has been recognized in the plant realm, and diverse WRKY families are found in higher plants [18]. The transcript investigation uncovered that the WRKY transcript gets upregulated because of saltiness, parchedness, salicylic acid (SA), methyl jasmonate (MeJa), and the collar rot fungus Macrophomina[19]. Similarly, AP2/ERF transcription factors assume critical jobs in plant development, improvement and reactions to biotic and abiotic stresses [20]. In this study, certain DGEs encoding WRKY and AP2/ERF transcription factors exhibited different expression patterns, and we found that WRKY transcripts were accumulated in response to C. gloeosporioides*(Fig. 5)*.

V. CONCLUSIONS

Using Illumina sequencing technology, a total of 25,515 genes were obtained and annotated. Analyzing the DGE library provided comprehensive, valuable information regarding the defense mechanism against anthracnose in J. curcas. The expression pattern analysis using RNA-Seq suggested the following: WRKY transcription factors, calcium-binding protein CML and KCS; 3-ketoacyl-CoA synthase upregulation activated defense-related genes induction in 9-1 cultivar which contributes to defense mechanism. Overall, our data provide valuable new clues and information about the C. gloeosporioides induced anthracnose in 2 different cultivars of J. curcas, and further study on the anthracnose-related genes will be helpful in elucidating the defense mechanism of J. curcas.

VI. ABBREVIATIONS

TF: Transcription factor
GO: Gene ontology
SRA: Sequence read archive
DGEs: Differential Gene Expression
FPKM: Fragments Per Kilobase Million
GTF: Gene transfer format
NCBI: National Center for Biotechnology Information
KEGG: Kyoto Encyclopedia of Genes and Genomes

VII. DECLARATIONS

a. AVAILABILITY OF DATA AND MATERIALS
The datasets supporting the conclusions of this article are included within the article and RNA-Seq data is available as accession number SRP044808 in the NCBI SRA database (https://www.ncbi.nlm.nih.gov//sra/?term=SRP044808).

b. AUTHOR'S CONTRIBUTIONS
Dr. Neeta and Dr. Siddhivinayakconceived the study. Mr. Krunal carried out the analysis and wrote the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

Computation support for this study was provided by Sedeer Medical, Fereej Kulaib Street, PO Box: 6782, Doha – Qatar.

CONFLICT OF INTEREST

The authors declare that they have no Conflict of Interest.

REFERENCES

[1]. Ong HC, Mahlia TM, Masjuki HH, Norhasyima RS. Comparison of palm oil, Jatropha curcas and Calophylluminophyllum for biodiesel: a review. Renewable and Sustainable Energy Reviews. 2011 Oct 1;15(8):3501-15.

[2]. Pandey VC, Singh K, Singh JS, Kumar A, Singh B, Singh RP. Jatropha curcas: A potential biofuel plant for sustainable environmental development. Renewable

and Sustainable Energy Reviews. 2012 Jun 1;16(5):2870-83.
[3]. Hui W, Yang Y, Wu G, Wang Y, ZakyZayed M, Chen X. Differential gene expression analyses related to fruit yield of Jatropha curcas L. using RNA-seq. Biotechnology & Biotechnological Equipment. 2018 Sep 3;32(5):1126-33.
[4]. Carvalho CR, Clarindo WR, Praça MM, Araújo FS, Carels N. Genome size, base composition and karyotype of Jatropha curcas L., an important biofuel plant. Plant Science. 2008 Jun 1;174(6):613-7.
[5]. Sato S, Hirakawa H, Isobe S, Fukai E, Watanabe A, Kato M, Kawashima K, Minami C, Muraki A, Nakazaki N, Takahashi C. Sequence analysis of the genome of an oil-bearing tree, Jatropha curcas L. DNA research. 2011 Feb 1;18(1):65-76.
[6]. Hirakawa H, Tsuchimoto S, Sakai H, Nakayama S, Fujishiro T, Kishida Y, Kohara M, Watanabe A, Yamada M, Aizu T, Toyoda A. Upgraded genomic information of Jatropha curcas L. Plant Biotechnology. 2012:12-0515.
[7]. Wu P, Zhou C, Cheng S, Wu Z, Lu W, Han J, Chen Y, Chen Y, Ni P, Wang Y, Xu X. Integrated genome sequence and linkage map of physic nut (Jatropha curcas L.), a biodiesel plant. The Plant Journal. 2015 Mar;81(5):810-21.
[8]. Ha J, Shim S, Lee T, Kang YJ, Hwang WJ, Jeong H, Laosatit K, Lee J, Kim SK, Satyawan D, Lestari P. Genome sequence of Jatropha curcas L., a non-edible biodiesel plant, provides a resource to improve seed-related traits. Plant biotechnology journal. 2019 Feb;17(2):517-30.
[9]. Kancharla N, Jalali S, Narasimham JV, Nair V, Yepuri V, Thakkar B, Reddy VB, Kuriakose B, Madan N. De Novo sequencing and hybrid assembly of the biofuel crop Jatropha curcas L.: identification of quantitative trait loci for geminivirus resistance. Genes. 2019 Jan;10(1):69.
[10]. Zhang X, Pan BZ, Chen M, Chen W, Li J, Xu ZF, Liu C. JCDB: a comprehensive knowledge base for Jatropha curcas, an emerging model for woody energy plants. BMC genomics. 2019 Dec;20(9):1-9.
[11]. Kwon JH, Choi O, Kim J, Kwak YS. First report of anthracnose disease on Jatropha curcas caused by Colletotrichumgloeosporioides in Korea. Journal of Phytopathology. 2012 May;160(5):255-7.
[12]. Ferreira-Pinto MM, Silva MJ, Santos MR. Screening of Jatropha curcas genotypes to anthracnose caused by Colletotrichumgloeosporioides. Communications in agricultural and applied biological sciences. 2011;76(4):629-34.
[13]. Doughty T, Kerkhoven E. Extracting novel hypotheses and findings from RNA-seq data. FEMS Yeast Research. 2020 Mar;20(2):foaa007.
[14]. Chowdhury HA, Bhattacharyya DK, Kalita JK. Differential Expression Analysis of RNA-seq Reads: Overview, Taxonomy and Tools. IEEE/ACM transactions on computational biology and bioinformatics. 2018 Oct 1.
[15]. Ajay Kumar G. Colletotrichumgloeosporioides: biology, pathogenicity and management in India. Journal of Plant Physiology and Pathology. 2014;2:2-11.
[16]. Rahman M, Ahmad SH, Mohamed MT, Rahman MZ. Extraction of Jatropha curcas fruits for antifungal activity against anthracnose (Colletotrichumgloeosporioides) of papaya. African Journal of Biotechnology. 2011;10(48):9796-9.
[17]. Atkinson NJ, Urwin PE. The interaction of plant biotic and abiotic stresses: from genes to the field. Journal of experimental botany. 2012 Jun 13;63(10):3523-43.
[18]. Ülker B, Somssich IE. WRKY transcription factors: from DNA binding towards biological function. Current opinion in plant biology. 2004 Oct 1;7(5):491-8
[19]. Agarwal P, Dabi M, Agarwal PK. Molecular cloning and characterization of a group II WRKY transcription factor from Jatropha curcas, an important biofuel crop. DNA and cell biology. 2014 Aug 1;33(8):503-13.
[20]. Tang Y, Qin S, Guo Y, Chen Y, Wu P, Chen Y, Li M, Jiang H, Wu G. Genome-wide analysis of the AP2/ERF gene family in physic nut and overexpression of the JcERF011 gene in rice increased its sensitivity to salinity stress. PLoS One. 2016 Mar 4;11(3):e0150879.