International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 8, August - 2014

Statistical Analysis of Microarray Data to Elucidate the Differential Gene Expression of Puccinia Striiformis F.sp. Tritici.

Shwetha P[1]
Bioinformatics(BBI)
GM Institute of Technology
Davangere, India.

Dr. H. Gurumurthy[2]
HOD, Dept. of Biotechnology
GM Institute of Technology
Davangere, India.

Prof. Manjunath Dammalli[3]
Dept. of Biotechnology
Siddaganga Institute of Technology
Tumkur, India.

***Abstract*—The Yellow rust of wheat (*Puccinia striiformis* f.sp. *tritici*), is one of the most destructive diseases causing extensive yield loss throughout the world. The present study deals with transcript profiling using Affymetrix Wheat Genome Array GeneChip. Molecular function enrichment analysis suggested that the differentially regulated genes were mainly related to protein degradation and modification, cell signaling and stress related mechanisms. The knowledge and comprehension of currently applied methods is one of the central criteria for a successful work. This study would be helpful in identification of early induced genes in wheat pathogens by the information of the resistance genes.**

Keywords*—Triticum aestivum; Wheat blast; Yellow rust; Leaf rust resistance; Bioconductor R; Resistance genes.

I. INTRODUCTION

Triticum aestivum, common or bread wheat, is an annual grass in the Poaceae (grass family) native to the Mediterranean region and southwest Asia, which is one of several species of cultivated wheat, grown in temperate climates worldwide. Wheat is one of the top two cereal crops grown in the world for human consumption, along with rice (*Oryza sativa*). It is one of the most ancient of domesticated crops, with archaeological evidence of the cultivation of various species in the Fertile Crescent dating back to 9,600 B.C. The various species have been developed into thousands of cultivars that differ in chromosome number from the primitive diploid types, with 7 pairs of chromosomes, to hybrid allopolyploids, with 14, 21, and 28 chromosome pairs[1].

Cultivars are variously categorized according to their horticultural requirements (spring vs. winter wheat), texture and food uses (hard wheat, which often contains more gluten and is used for bread; vs. pastry or flour wheat, used for cakes, biscuits, and cookies), or by growth form and seed characteristics. Wheat (*Triticum aestivum*) is high in carbohydrates, protein (although it lacks several essential amino acids), and vitamins B and E (if the grain is left whole) is used in countless breads and baked goods, and is an important source of calories. Wheat can be refined into starch and wheat-germ oil, and wheat gluten (the proteins that make it sticky) is used in many products. The straw is traditionally used for thatching and wickerwork; it can also be utilised to make pulp for paper etc. or as fuel. Wheat is also used to make beer and as animal fodder[2]. The FAO estimates that global commercial production of all types of wheat was 650.9 million metric tons in 2010, harvested from 217.0 million hectares; it is grown on around 4% of the planet's agricultural land. Leading producers were China, India, the U.S., the Russian Federation and France. The cereal grain wheat is subject to numerous wheat diseases including bacterial, viral and fungal diseases. The rusts of wheat (*Triticum aestivum*) cause common and widespread wheat diseases that can be found in most areas of the world where wheat is grown. Wheat stem rust is caused by *Puccinia graminis* f. sp. *tritici*, wheat leaf rust by *Puccinia triticina*, and wheat stripe rust is caused by *Puccinia striiformis*[3]. The blast fungus *Magnaporthe grisea* causes a serious disease on a wide variety of grasses including rice, wheat, and barley[4].

Yellow rust of wheat (*Puccinia striiformis f.sp. tritici*), a basidiomycete belonging to the uredinales, is the cause of stripe rust on cereal crops and grasses like wheat, corn or maize as shown in the left of fig. 1. Several formae speciales of *P. striiformis* West. var. striiformis have been successively named on the basis of physiological specialization: *P.striiformis f.sp. tritici* collected from wheat[5]. Like other cereal rusts, *P. striiformis* forms races which are usually identified with a differential set of wheat cultivars[6]. Wheat blast, caused by *Magnaporthe oryzae Triticum* pathotype (wheat isolates), was first reported in

the State of Parana in Brazil in 1985. This fungus has since become a major pathogen. The disease also occurs on triticale, barley and black oats. The pathogen can infect all above-ground parts of wheat plant, but damage in the field comes mainly from head (spike) blast, which produces shriveled seeds or totally prevents grain filling as in the right of fig1. Symptoms closely resemble Fusarium head blight[7]. Yield losses to this disease range from low, when the weather doesn't favor disease, to as high as 100% when conditions favor disease. Effective resistance is generally lacking for the wheat blast disease and fungicide treatments are unreliable when weather favors disease[8].

Fig.1. Yellow rust and blast fungus

Microarrays are used to measure gene expression levels in different ways. An experiment was designed by which the microarray experiment was carried out and data were generated. The analysis of microarray data to produce lists of differentially expressed genes has several steps which can differ based on the type of data being assayed. However, all data follows the same general pipeline which involves reading raw data, quality assessing the data, removing bad spots/arrays from further analysis, preprocessing the data and calculating differential expression by statistical analysis[9]. In addition, higher level analysis may involve various methods relevant to the biological samples and the information required. The data provide information on RNA expression levels, not on mechanism or casuality. Data analysis usually leads to new hypotheses that are tested in follow-up experiments which identify relevant metabolic and signaling pathways. Thus, the list of differentially expressed genes can subsequently be annotated with useful information that explains the various genes function, for example, gene ontology[10]. In this paper, the common resistance genes are identified so that this work can help in making a foundation in further studies.

II. MATERIALS AND METHODS

The current study has planned to take gene expression data from Gene Expression Omnibus (GEO) and aimed to get genes which were not found by the studies conducted previously(http://www.ncbi.nlm.nih.gov/geo). The series matrix of the sample is downloaded and they are saved in ZIP/winRAR format. For microarray analysis, we used a series of R/Bioconductor packages (http://www.bioconductor.org). Briefly, the CEL files were imported into R environment and the robust multi-array average (RMA) methodology, as implemented in the affy package, was used for microarray normalization. Following normalization, a non-specific filtering step was carried out.

For the given gene list, The **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (**DAVID)** tool was used to identify enriched biological themes, particularly GO terms and also discover enriched functional-related gene groups and convert gene identifiers from one type to another(http://david.abcc.ncifcrf.gov). The common gene list are found by using Jvenn(http://www.bioinfo.cau.edu.cn/jvenn/). The agriGO (http://bioinfo.cau.edu.cn/agriGO), a web-based tool and database is used for the gene ontology analysis. The complex networks can be visualized using Cytoscape (http://cytoscape.org/) an open source software platform. The UniProt Knowledgebase (UniProtKB) was referred for the collection of functional information on proteins, with accurate, consistent and rich annotation (http://www.uniprot.org/uniprot/).

III. RESULTS AND DISCUSSION

The goal was to identify a set of genes which are common and differentially expressed in the fungal diseases of wheat. The samples for different fungal diseases like yellow rust(GSE31761) and blast fungus(GSE31760) are downloaded from GEO database in order to perform expression analysis. The National Center for Biotechnology Information's(NCBI) Gene Expression Omnibus(GEO) database was queried for datasets of wheat involving two infectious fungal pathogens: *Puccinia striiformis* and *Magnaporthe grisea*. GEO datasets were selected based on the following inclusion criteria: Both the datasets are of the same organism *triticum aestivu*m. The samples must be originated from Affymetrix Wheat Genome Array GeneChip. Each dataset must have atleast 3 groups and the supplementary files must be of the type .CEL file. All criteria for dataset inclusion in the final analysis were chosen prior to the analysis.

A. Pre-processing

The installation and loading of packages from the libraries are done using Bioconductor R. The current working directory is set in the beginning. The datasets will be unzipped in .CEL files format and screened in a folder. Further preprocessing and analysis was performed using the .CEL files. The data preprocessing was done in bioconductor R after the .CEL files were imported into RMA for further processing. The .CEL files from the folder specified are read by using the ReadAffy command in R programming.

B. Normalization

The LIMMA package were used to normalize the microarray data. Subsequent background adjustment, quantile normalization of the raw data and estimation of probe sets signal intensities were to be done. Thus, probeset was summarized and the expression values were determined. This was done by using GeneChip RMA (GC-RMA), an improved form of Robust Multiarray Averaging

(RMA) method of microarray normalization and summarization that is able to use the sequence specific probe affinities of the GeneChip probes to attain more accurate gene expression values. The boxplot then appears after normalization as in fig.2.

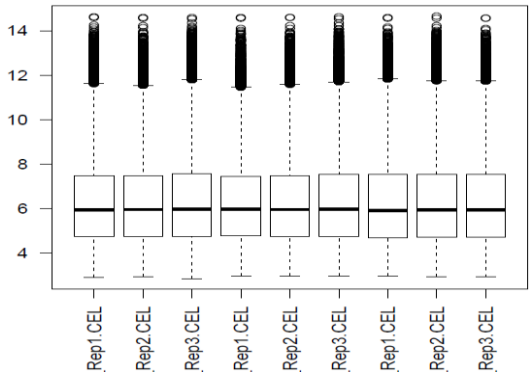

Fig.2. Boxplot after normalization for GSE31761

C. *Quality Control*

To run the statistical algorithm in bioconductor R, a matrix design was built which was created by grouping of the samples. List of possible number of comparisons was made before running the statistical algorithm. The significant probe sets were extracted that met the criteria of 0.001 (p-value). The heat map was then generated. The Hierarchical clustering was done by generating a Heat Map by using heatmap function of R package. The clustering of samples are shown horizontally above the heatmap and the probeset ID's are shown vertically in the left side of the heatmap as shown in the fig.3. In the cluster analysis of the probe sets of dataset GSE31761, red colour indicates the highly expressed probe sets and green colour indicates the less expressed probesets.

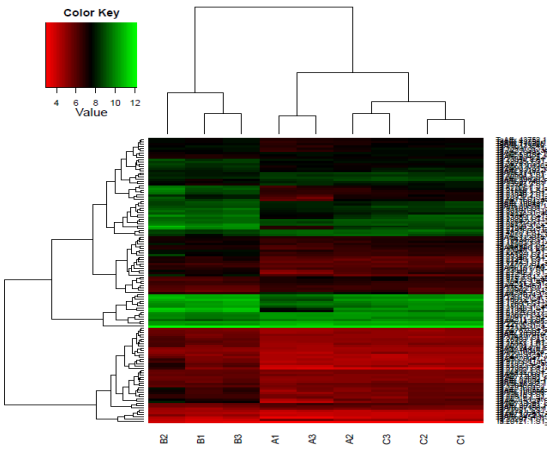

Fig.3. Heatmap generated for GSE31761

D. *Differentially Expressed Genes*

Statistical tests are carried out that will be used to identify the genes that are differentially expressed among the two datasets. The significant p-values are selected as parameters. For this analysis, p-value is used which is a measure that allows us to control how big a proportion of false positives (genes that we think are differentially expressed but really are not) we are willing to accept. Often the results of microarray experiments are verified using other methods, and then we may want to filter out genes that exhibit differences in expression that are so small that we will not be able to verify them with another method. This is done by adding one last criterion to the filter: Difference should have a significant value higher than 0 or lower than 0, as we are working with log transformed data, the group mean difference is really the fold change, so this filter means that we require a fold change above 0 and below 0. Note that the significant value > is important because the difference could be negative as well as positive. The result is that we end up with a list of genes that are likely candidates to exhibit differential expression in the two groups.

A number of summary statistics are computed for each gene. The log-fold change is the log expression level for that gene. The AveExpr is the average expression level for that gene across all the arrays and channels. Differential expression analysis of genes was performed by means of the moderated *t* statistics(t) using Benjamini-Hoschberg false discovery rate (FDR) correction. The moderated t-statistic(t) is the ratio of the M value to its standard error. In addition, p value threshold of <0.001 was used for the comparison, in order to extract the significantly differentially expressed genes. Each p-value has an adjacent p-value for each of the gene. The log odd statistics(B) is shown for each gene. For the dataset GSE31761, the list of genes >0 are the up-regulated value and the list of genes <0 are the down-regulated values. The process of the data analysis, pre-processing, normalization, quality control and the differential genes expressed for the dataset GSE31761 must also be carried out for the other dataset in the same way. The combined results of both the datasets are used in the further work.

E. *Common Gene List*

The common genes between the two datasets are acquired using Jvenn that is represented in venn diagram and a bar chart. The up-regulated and down-regulated probe ids of both the datasets are pasted in the given box. Venn diagram shows the overlap of up-regulated and down-regulated genes in response to the two datasets: GSE31760(green) and GSE31761(blue). Area of overlaps is not proportional to the overlap. The numbers of genes in each region of the diagram are indicated. There are 40 genes that are common between the two datasets as shown in fig.4.

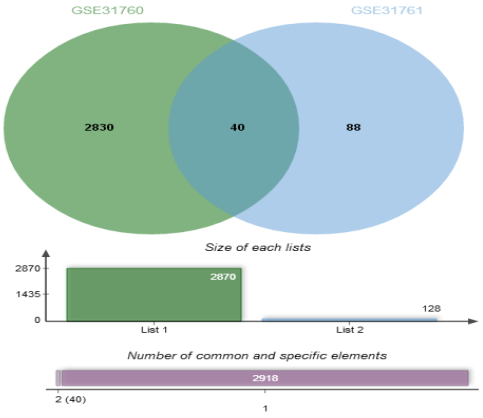

Fig.4. Common genes between both the datasets.

Out of those common genes, 37 common genes are up-regulated and 2 common genes are down-regulated. Some genes may not be shown as they belong to the species of different organisms. The list of common genes between the two datasets is shown in Table1. The gene names that are not shown are uncharacterized proteins.

Table 1. Common gene list

AFFYMETRIX_3PRIME_IVT_ID	Name
Ta.18203.1.S1_at	blue copper-binding protein homolog
Ta.27762.1.S1_x_at	thaumatin-like protein
Ta.21342.1.S1_x_at	chitinase 3
Ta.24501.1.S1_at	thaumatin-like protein
Ta.28.1.S1_at	glucan endo-1,3-beta-D-glucosidase
TaAffx.15327.1.S1_at	glucan endo-1,3-beta-D-glucosidase
Ta.82.1.S1_at	peroxidase
Ta.8447.1.S1_a_at	No Homology
Ta.21281.1.S1_at	No Homology
Ta.97.1.S1_at	No Homology
TaAffx.107507.1.S1_at	No Homology
Ta.8674.1.A1_at	No Homology
TaAffx.24475.1.S1_x_at	No Homology
Ta.13.1.S1_at	No Homology
Ta.15072.1.A1_at	No Homology
Ta.22615.1.S1_at	No Homology
Ta.97.2.S1_x_at	No Homology
Ta.3133.1.S1_x_at	No Homology
Ta.30501.1.S1_at	No Homology
TaAffx.28302.2.S1_at	No Homology
TaAffx.110196.1.S1_s_at	No Homology
TaAffx.108437.1.S1_at	No Homology
Ta.5518.1.S1_at	No Homology
Ta.14946.1.S1_at	No Homology
TaAffx.6454.1.S1_at	No Homology
Ta.27314.1.S1_at	No Homology
Ta.8990.1.S1_at	No Homology
TaAffx.110081.1.S1_x_at	No Homology
Ta.22619.1.S1_x_at	No Homology
TaAffx.107979.1.S1_at	No Homology
TaAffx.108437.1.S1_x_at	No Homology
Ta.21556.1.S1_at	No Homology
TaAffx.28047.1.S1_at	No Homology
Ta.21556.1.S1_x_at	No Homology
Ta.13991.1.S1_x_at	No Homology
Ta.11087.1.S1_at	No Homology
Ta.30731.1.S1_at	No Homology

F. *Functional Annotation*

The total differentially expressed genes; Up regulated and Down regulated, were mapped to DAVID open source database, this indexing will give curated evidence and confirmation of these genes as differentially expressed. The annotation results show that the probe id list has two functional categories and three protein domains. In functional categories, the functions of 13 genes are given from the SwissProt-Protein Information Resource and the sequence features of 11 genes are from the UniProt database. There are 13 genes having Protein domains shown from Interpro. The Protein Information Resource superfamily has 10 genes and there are 6 genes having protein domains in Smart. The probe ids are converted to gene names using the gene accession conversion tool. This makes it possible to know the interaction of the genes by the columns From and To.

G. *Gene Ontology Annotation*

The gene list will be annotated to see the patterns in the biological annotations of the genes in the list of candidate differentially expressed genes. Each of the two groups of genes, that is, showing parental dominance expression and non-additive expression, both in their entity and as further categorized subgroups, were analyzed with Gene Ontology (GO) annotation using AgriGO, a web-based database tool for gene ontology annotations of agricultural crops. The Singular Enrichment Analysis (SEA) tool was used to perform the GO annotations and statistical analysis for GO term-enrichment. The SEA analysis computed GO term enrichment in one set of genes by comparing it to another set, then named the target and reference list, respectively. The Fisher exact test was used for statistical analysis with Yekutieli FDR based multi-test adjustment method with the significance of P-value < 0.01. The GO processes can be represented in a graphical way as shown in fig. 5.

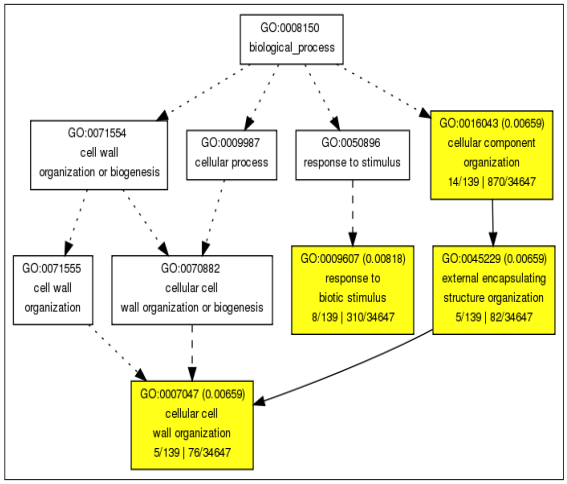

Fig. 5 Graphical results

AgriGO also displays the results by representing them in a bar chart as shown in fig.6. In the biological process about 42% of the genes were mostly enriched in the cellular process and the metabolic process. The catalytic activity and the binding activity had the highest percent of genes enriched (40%) in the molecular function. The cell

part and the organelles were more enriched in the cellular component which had more than 56% of the genes.

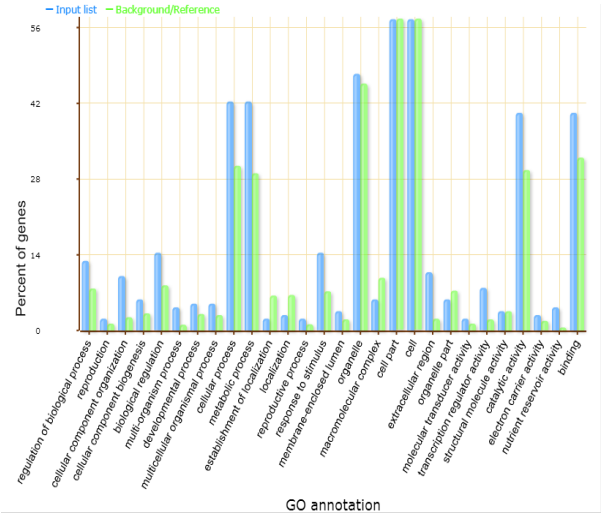

Fig. 6 Bar chart of the biological annotation

There were 11 significant GO terms shown. An important annotation is the Gene title which describes the gene and is much more informative. The fig.7 summarizes the GO annotation terms along with its ontology and description.

GO term	Ontology	Description	Number in input list	Number in BG/Ref	p-value	FDR
GO:0016043	P	cellular component organization	14	870	1.3e-05	0.0066
GO:0007047	P	cellular cell wall organization	5	76	1.9e-05	0.0066
GO:0045229	P	external encapsulating structure organization	5	82	2.7e-05	0.0066
GO:0009607	P	response to biotic stimulus	8	310	4.5e-05	0.0082
GO:0050162	F	oxalate oxidase activity	4	5	3e-08	8.3e-06
GO:0016623	F	oxidoreductase activity, acting on the aldehyde or oxo group of donors, oxygen as acceptor	4	10	2.4e-07	3.3e-05
GO:0030145	F	manganese ion binding	5	77	2e-05	0.0019
GO:0005576	C	extracellular region	15	777	7.2e-07	0.00012
GO:0030312	C	external encapsulating structure	6	170	7.9e-05	0.0033
GO:0048046	C	apoplast	5	98	6.1e-05	0.0033
GO:0005618	C	cell wall	6	166	7e-05	0.0033

Fig.7. Gene Otology terms list

H. Mapping of Differentially Expressed Genes to Pathway Databases.

The genes are interconnected to different genes which can be predicted by the network topology using the cytoscape software. The network statistics can be done in cytoscape for network analysis of pathways. The file is imported and the interactions are being defined in a new window which appears after importing the file. The network can also be assigned names at the top of the window. The file has to be imported and the columns are set as the source and target. After importing the file a network appears. The edges and nodes are shown in the network. The analysis of the network statistics is done by using the network analyzer in tool box.

The network has parameters that are set to visualize the network. This is done by clicking visualize parameters below. A new window appears showing the parameters that can be set. The arrows between the nodes can be directed or undirected. The edges, nodes and arrows can be set in different colours. The red colour represents highly expressed genes and the green colour represents low expressed genes as shown in fig.8. The nodes or genes can be zoomed to view the names of the genes expressed in the topology. The list of highly expressed genes along with its gene ID, gene symbol and the function of those genes are shown in Table2.

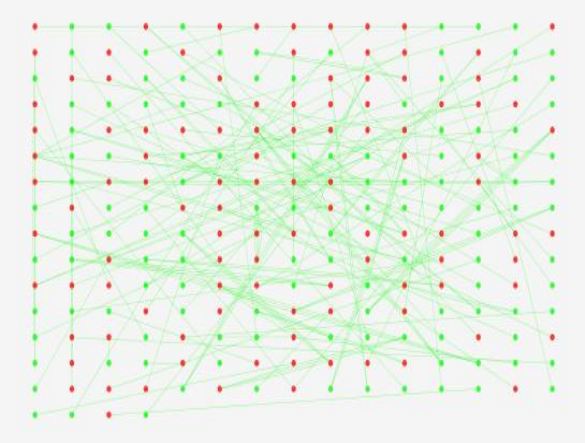

Fig.8 Network topology

Table 2. List of highly expressed genes

Gene ID	Gene Name	Gene Symbol
606343	tousled-like protein kinase	TLK1
100146083	calreticulin	CRT
542771	wpk4 protein kinase	wpk4
100127073	PTF1	LOC100127073
543292	thaumatin-like protein	LOC543292
543090	allene oxide synthase	TaAOS
100037657	CBFIIIc-D3	LOC100037657
543498	germin protein precursor	LOC543498
542994	glutathione transferase	gstu3
780696	alternative splicing regulator	RSZ38
543153	AML15	TaAML15
606342	ribosomal protein S29	LOC606342
100125682	sulfur-rich/thionin-like protein	LOC100125682
606326	NAC domain transcription factor	NAC2
543365	peroxidase	LOC543365
543380	phenylalanine ammonia-lyase	wali4
606315	glycosyltransferase	a3a
542826	blue copper-binding protein homolog	S85
542788	glucose-6-phosphate dehydrogenase	g6pdh
606333	ribosomal protein L6	LOC606342
100146081	homeobox-like resistance protein	HLRG
100136972	cryptochrome 2	Cry2
543422	pathogenesis-related protein 1	LOC543422
100037560	flavanone 3-hydroxylase	LOC100037560
542892	metallothionein	LOC542898
543216	ubiquitin carrier protein	LOC543216
543491	S-adenosyl-L-homocysteine hydrolase	SH6.2
100049026	WRKY transcription factor	WRKY10
543321	pSBGer1 protein	pSBGer1
780664	U2AF small subunit	LOC780664
54330	glucan endo-1,3-beta-D-glucosidase	LOC543330
606311	histone H1 WH1A.3	TAc41

The functions of highly expressed genes in the pathways are shown in Table 3. The differentially expressed genes were mapped to their pathway. This gave the information about the genes and the pathway on which the gene acts.

Table 3 Functions of highly expressed genes

Gene Symbol	Function
TLK1	perform cell autonomous functions
CRT	plays a key role in many cellular processes
wpk4	shows increased transcript levels in response to multiple stimuli
LOC100127073	DNA binding
LOC543292	defense response, response to biotic stimulus
TaAOS	heme binding, iron ion binding, oxidoreductase activity, acting on paired donors
LOC100037657	recognition of interaction partners and transactivation potential of a specific set of CBF proteins.
LOC543498	play an important role in several aspects of plant growth and defense mechanisms.
gstu3	glutathione transferase activity, metabolic process
RSZ38	nucleic acid binding, nucleotide binding, zinc ion binding
TaAML15	positive regulation of growth, positive regulation of meiosis, nucleic acid binding, nucleotide binding
LOC606342	translation, metal ion binding, structural constituent of ribosome
LOC100125682	plant defense response
NAC2	tolerances to drought, salt, and freezing stresses
LOC543365	response to environmental stresses such as wounding, pathogen attack and oxidative stress.
wali4	produces environmental stresses such as wounding, HgC12, UV light, and funga1 elicitors
a3a	transferase activity, transferring glycosyl groups
S85	copper ion binding, electron carrier activity.
g6pdh	salt stress response
LOC606342	translation, structural constituent of ribosome.
HLRG	involved in race-specific responses to stripe rust
Cry2	Subcellular Localization and Involvement in Photomorphogenesis and Osmotic Stress Responses
LOC543422	extracellular region
LOC100037560	oxidoreductase activity, with incorporation or reduction of molecular oxygen.
LOC542898	metal ion binding,
LOC543216	protein modification; protein ubiquitination
SH6.2	control of methylations via regulation of the intracellular concentration of adenosylhomocysteine
WRKY10	sequence-specific DNA binding transcription factor activity
pSBGer1	manganese ion binding, nutrient reservoir activity
LOC780664	RNA binding, metal ion binding, nucleotide binding.
LOC543330	carbohydrate metabolic process, Catalysis of the hydrolysis.
TAc41	nucleosome assembly, DNA binding

I. Identification of Common Resistance Genes

Plants have evolved R genes (resistance genes) whose products allow recognition of specific pathogen effectors, either through direct binding of the effector or by recognition of the alteration that the effector has caused to a host protein. Resistance genes help in identifying the need of benefits in agriculture and for further studies.

The list of common resistance genes for both datasets is shown in the Table 4. The gene id and gene names are shown for the respective genes. The resistance genes are given for each gene which are resistant to that particular gene in the process. The function of each gene represents protein coding.

Table 4 Common resistance genes

Gene ID	Gene Name	Resistance gene
543330	glucan endo-1,3-beta-D-glucosidase	Yr5
542826	blue copper-binding protein homolog	Lr34/Yr18
542780	chitinase 3	Sr5/Sr24
543342	thaumatin-like protein	Yr26
543285	peroxidase	Sr5/Sr6
543330	glucan endo-1,3-beta-D-glucosidase	Yr5
543342	thaumatin-like protein	Yr26

IV. CONCLUSION

High throughput technologies, such as gene expression arrays and protein mass spectrometry allow one to simultaneously evaluate thousands of potential biomarkers that distinguish different tissue types. The common genes responsible to cause the rust in wheat have been identified using a technique called DNA microarray analysis. As an overview of the entire process, relevant data from GEO is acquired, tabulated and subjected to various analysis tools that could generate relevant annotations. Additionally, connections to related metabolic pathways and common differentially expressed genes are shown in the results. The study looks forward of investigating the common genes responsible to cause the rust disease in wheat. The Functional annotation and expression profiling can implicate subsets of genes in compatibility and incompatibility of leaf rust in wheat. Extensive studies on other related genes will help to understand their role in leaf rust infection in wheat. Many new genes have to be identified that can be useful for future studies.

ACKNOWLEDGMENT

The authors would like to thank Siddaganga Institute of Technology, tumkur for their technical support and assistance. We are grateful to GM Institute of Technology for their guidance in this work.

REFERENCES

[1]. http://eol.org/pages/1115240/overview
[2]. http://www.cropscience.bayer.com/en/Crop-Compendium/Crops/Wheat.aspx
[3]. Kolmer, J., "Leaf Rust of Wheat: Pathogen Biology, Variation and Host Resistance", Forests, Vol 4, 2013, pp 70-84.
[4]. Talbot, N.J., "On the trail of a Cereal Killer: Exploring the Biology of Magnaporthe grisea", Annu. Rev. Microbiol, Vol 57, 2003, pp 177–202.
[5]. Coram, T.E., Settles, M.L., and Chen, X., "Transcriptome analysis of high-temperature adult-plant resistance conditioned by Yr39 during the wheat–Puccinia striiformis f. sp. tritici interaction", Molecular Plant Pathology, Vol 9, 2008, pp 479–493
[6]. Bozkurt, O., Akkaya, M., "Transcription profiling of wheat interacting with incompatible and compatible yellow rust in a Yr1-containing genotype", 2011.
[7]. Bockus, W., Cruz, C., Stack, J., Valent, B., Farman, M., Hershman, D., Paul, P., Peterson, G., Pedley, K., and Magarey, R., "Recovery Plan For Wheat Blast Caused by Magnaporthe oryzae Triticum pathotype", June 2013
[8]. Tufan, H.A., McGrann, G.R.D., Magusin, A., Morel, J.B., Miche, L., and Boyd, L.A., "Wheat blast: histopathology and transcriptome reprogramming in response to adapted and nonadapted Magnaporthe isolates", New Phytol, Vol 184(2), 2009, pp 473-84
[9]. Slonim, D.K., Yanai, I., "Getting Started in Gene Expression Microarray Analysis", PLoS Computational Biology, Vol 5, October 2009, issue 10, e1000543.
[10]. Clarke, J.D., and Zhu, T., "Microarray analysis of the transcriptome as a stepping stone towards understanding biological systems: practical considerations and perspectives", The Plant Journal, Vol 45, 2006, pp 630–650.