Journal of Multidisciplinary Scientific Research, 2014,2(6):05-09
ISSN: 2307-6976

Available Online: http://jmsr.rstpublishers.com/

# EVALUATION OF MICROARRAY SENSITIVITY AND SPECIFICITY IN GENE EXPRESSION DIFFERENTIAL ANALYSIS BY RNA-Seq AND QUANTITATIVE RT-PCR.

**Dago Dougba Noel*[1,2],Giovanni Malerba[3] , Alberto Ferarrini[2] and Massimo Delledonne[2]**

1)UPGC University of Korhogo Cote d'Ivoire Unité Formation Recherche Sciences Biologiques,Ivory Coast.
2)Department of Biotechnology,University of Verona, Italy Strada le Grazie 15, Cà vignal 1,Italy.
3)Department of Life and Reproduction sciences, Section of Biology and Genetics, University of Verona, Italy.
*Email: dgnoel7@gmail.com*

**Received:14 ,Aug,2014** **Accepted: 21, Nov,2014**

## Abstract

*There are several techniques for quantifying the amount of transcribed mRNA, all of them relying on the fundamental property of complementary base pairing. The widely used tools are gene expression microarrays, allowing for the snapshot of the entire genome. The reliability of microarray technology in gene expression analysis and diagnostic process have been fully discussed and critiqued, allowing the development of several microarray technology design strategies with the aim of improving its accuracy in transcriptome and genomic studies. Hence, we are evaluating the sensitivity and the specificity of four previously developed grape microarray design strategies based either on multiple and/or single long and/or short oligonucleotide probe per gene model transcript by RNA-Seq and quantitative RT-PCR (qRT-PCR) technologies; this is due to their advantages in detection sensitivity, sequence specificity, the large dynamic as well as their high precision and reproducible quantitation compare to microarray. Our results showed that (i) regardless of the array design strategies used, microarray gene expression technologies are less specific and less sensitive for the purpose of detecting lower expressed gene (differential expressed genes (DEGs)) associated with small variation change; and (ii) for the highly expressed genes, microarray gene expression platforms, exhibited a high reliability and a good level of agreement with both RNA-Seq and qRT-PCR tools in gene expression differential analysis.*

**Keywords**: *Sensitivity, Specificity, Microarray, Rna-Seq, Qrt-Pcr.*

## INTRODUCTION

While the multiplicity of microarray platforms offers an opportunity to expand the use of the methodology and make it more easily available to different laboratories, the comparison and integration of data sets obtained with different microarray remains a challenge. Diversity arises from the technology features intrinsic to chip manufacturing, from the protocols used for sample processing and hybridization, from detection systems, as well as from approaches applied to data analysis. On the one hand, the combined use of multiple platforms can overcome the inherent biases of each approach, and may represent an alternative that is complementary to quantitative reverse tanscription (qRT-PCR) for identification of the more robust changes in the gene expression profiles on the other hand, the comparison of data generated using different platforms may represent a significant challenge, particularly when considering systems that vary greatly. The publication of studies with dissimilar or altogether contradictory results, obtained using different microarray platforms to analyze identical RNA sample, has raised concerns about of the reliability of this technology. The Microarray Quality Control (MAQC) project was initiated to address these concerns, as well as other performance and data analysis issues. The MAQC project has generated a rich data set that, when appropriately analyzed, reveals promising results regarding the consistency of microarray data between laboratories and cross platforms [1]. What is expected from microarrays? Searching for determinants of a phenotype using gene expression levels requires suitable exposure of the genome coupled with reasonable reproducibility, accuracy and sensitivity in the technology employed. These limitations matter less if microarrays are used for screening because changes in gene expression can be verified independently. However, the stakes were raised once microarrays were put forward as a diagnostic tool in molecular disease classification [2] because regulatory agencies, such as the Food and Drug Administration (FDA), require solid, empirically supported data related to the accuracy, sensitivity, specificity, reproducibility and reliability of diagnostic techniques. The first decade of microarray technology produced rather limited data pertinent to these issues. Nevertheless, transcriptional profiling using microarray technology is a powerful genomic tool that is widely used to characterize biological systems. Despite the increasing reliance on this technology by the scientific community, the issue concerning the reproducibility of microarray data between laboratories and across platforms has yet to be fully resolved. The issues of data reproducibility and reliability is crucial with regard the generation of, and ultimately to the utility of, a large database of microarray results. Several consortiums such as MAQC [1] have coordinated an impressive effort to develop guideline in order to assess the performance of different microarray technologies. However, even under normal conditions, microarray technology in its current state would face significant limitations for several reasons; (i) first, the relationship between probe sequences [3], target concentration and probe intensity is rather poorly understood. (ii) Second, splice variants constitute another dimension that can pose a problem for microarray analysis. It is estimated that at least half of the human genes are alternatively spliced, and might have many potential splice variants [4]. A given short oligonucleotide probe is targeted at either a constitutive exon (present in all splice variants) or at an exon specific for certain splice variants. (iii) Third, folding of the target transcripts [5] and cross-hybridization [6] can also contribute to the variation between different probes targeting the same region of a given transcript. It has been shown previously that a large proportion of the microarray probes produce significant cross-hybridization signals [7]. Even a limited stretch of sequence complementarity might be sufficient to enable binding between two unrelated sequences. However, evaluating the overall impact of cross-hybridization on the accuracy of microarray measurements is not

easy.While the reliability of the arrays is considered high and satisfactory, it is less clear if their heterogeneity of microarray design may affect the final results when searching for differentially expressed genes (DEGs). This issue has been marginally explored by earlier studies. The Microarray Quality Control (MAQC) consortium [1] evaluated the correlation of results obtained from different microarray platforms focusing on gene expression levels and concluded that the stability of gene lists correlates with endpoint predictability, subtly suggesting that microarray platforms are all similar in defining gene expression profiles [1]. However, despite their high degree of inter-platform data reproducibility, microarrays are extremely heterogeneous tools due to (i) their makeup and (ii) the bioinformatics and the statistical approaches used, for their data processing and analysis. Profiling mRNAs of a few hundred genes by qRT-PCR (qRT-PCR; reverse transcription followed by real-time PCR) has prove to be a viable option. In principle, qRT-PCR has higher specificity and accuracy than microarrays. More recently, next generation massively parallel sequencing technologies, have made it feasible to quantitate mRNA by direct sequencing of cDNAs and count each of the mRNA species. It is interesting to note that MAQC-III also known as Sequencing Quality Control (SEQC) is the third phase of the MAQC project (MAQC-III), also referred to interchangeably as SEQC. This project aims to assess the technical performance of next-generation sequencing platforms by generating benchmark datasets with reference samples and evaluating advantages and limitations of various bioinformatics strategies in RNA and DNA analysis. Moreover, as we know, fewer studies investigated the accuracy of microarray in detecting DEGs with respect to both qRT-PCR and RNA-Seq gene expression profiling technologies together. Therefore we evaluated the sensibility and the specificity of four previously developed grape microarray designs based on two different probe design strategies [8], by the integration of microarray gene expression data with those of both qRT-PCR and RNA-Seq gene expression profiling tools, recognized to be more specific and more sensitive with regard to microarray technologies in gene expression differential analysis.

## MATERIALS AND METHODS

Samples from two development stages of grape (veraison and ripening) have been profiled for global gene expression using four different microarray design strategies each based on either duplicate probes or different probes for the genes essayed [8]. Samples were previously profiled using next generation technology (RNAs-Seq) as reported in Zenoni et *al*.,2010 [9]. However, for this analysis we performed the RNA-Seq gene expression differential analysis using the R software DESeq package (version1.6.1,http://bioconductor. org/packages/release/ bioc/ html/DESeq.html).Normexp_saddle and quantile normalization parameters (R *limma* package) have been used to processed and analyzed microarrays data (background subtraction and normalization procedures). Results of gene expression differential analysis from each microarray were then compared with the results obtained from RNA-Seq and qRT-PCR.

### 1.Grape array hybridization and data processing/ normalization

Microarray design strategies (single and multiple long (60 mer) and/or short (35-40mer) oligonucletide probe per gene transcript model) based on two different custom microarray platforms have been described in Dago N. 2012 [8]. Microarray hybridization process and data processing and differential analysis performed between two *vitis vinifera* development stage repining and véraison used in this work have been reported in the table below.

**Table 1:** Summary of microarray hybridization process and data processing and differential analysis

| Grape microarray design | Background correction | Probe average | Package and statistical method analysis |
|---|---|---|---|
| Grape array design 1 (*GAD1*): single specific replicate long probe per gene | Normexp_sad -dle (*limma*) | Mean | *Limma* (moderated t test) |
| Grape array design 2 (*GAD2*): 4 different long probes per gene | Normexp_sad -dle (*limma*) | Mean | *Limma* (moderated t test) |
| Grape array design 3 (*GAD3*): single specific short probe per gene | Normexp_sad -dle (*limma*) | Mean | *Limma* (moderated t test) |
| Grape array design 4 (*GAD4*): 3 different short probes per gene | Normexp_sad -dle (*limma*) | Mean | *Limma* (moderated t test) |

### 2. Differential gene expression analysis

Differential gene expression analysis between ripening and véraison was performed by comparing arrays processed with the same background subtraction and data normalization combination. Differential gene expression analysis was conducted by applying linear models on the log-expression values and then an empirical Bayes moderated t-statistics on each gene. The "*lmFit"* and "*eBayes"* functions of the *limma* R package (version 3.10.3) were used [10]. RNA-Seq gene expression differential analysis for same analyzed samples (repining and véraison) was performed by using the R bioconductor packag DESeq (version1.6.1,http:/ /bioconductor. org/packages / release/ bioc/ html/ DESeq.html). The raw data of RNA-Seq is available at SRA009962 (or data can also be accessed on Genome Browser at URL http://ddlab.sci.univr.it/cgi-bin/gbrowse/grape). The False Discovery Rate (FDR) suggested by Benjamini and Hochberg [11] was adopted to control the FDR since multiple comparisons were computed for both microarray and RNA-Seq gene expression differential analysis. A gene was considered as differentially expressed (DEG) when showing a mean difference of the expression value greater than or equal to two folds between repining and véraison grape *Vitis vinifera* berry development stages at a False Discovery Ratio (FDR) ≤ 0.05.

### 3.Quantitative Real Time PCR for gene expression data validation

We designed RT-PCR primers (forward and reverse) on 10 genes region within 1 kb upstream of the 3'end for 10 randomly selected genes to validate RNA-Seq and microarray expression data. As a primer design template we used the 12x grape genome assembly [12]. RNA samples have been treated with DNase using the Turbo DNA-free kit (Applied Biosystem). Superscript II reverse Transcriptase Invitrogen kit for cDNA synthesis has been used for cDNA synthesis (3 different reactions have been performed for each considered grape *Vitis vinifera* development stage). qRT-PCR was performed in 25 µl reaction containing SYBR green master mix (invitrogen), 1 µl each primer and 2 µl of above prepared cDNA

template. PCR was performed in a MX 3000 Fast Real Time PCR system (ABI Instrument) in three technical replicates for each sample. The PCR involved a 50°C hold for 2 min and a 95°C hold for 10 min followed by 40 cycles at 95 °C for 30 s, 55 °C for 30 s and 72 °C for 20 s. The detection threshold cycle for each reaction was determined using a standard curve, after normalization of the results using qRT-PCR result of actine primes TC81781 (TIGR, Release 6.0) for the actine gene, which was proved manifest a constant expression level across ripening and veraison berry development stage. Amplification efficiency was calculated from raw data using LingRegPCR software [13]. The relative expression ratio value was calculated according [14] equation. Standard Error (SE) values were calculated according Pfaffl et *al.,* 2002 methods [15].

## RESULTS

**1 Comparison between array and RNA-Seq in gene expression differential analysis**

Significantly differential expressed genes (DEGs) have been detected at a false discovery rate (FDR) ≤0.05 for both microarrays and RNA-Seq technologies. The four analyzed microarray designs exhibit heterogeneous agreement when compared to RNA-Seq gene expression data (see Fig. 1), suggesting their heterogeneity discriminating DEGs in gene expression differential analysis. However, the Pearson correlation analysis based on log2-fold change (log2-FC) measurement of DEGs with a |log2-FC| ≥ 1 detected by both microarray designs 2, 3 and 4 and RNA-Seq (red dots in Fig.1) range between 0.85-0.88 ($R^2$:0.85-0.88). These results show a homogeneity of these microarray designs in discriminating DEGs in gene expression differentially analysis for gene set associated with a high variation change between repining and véraison grape *Vitis vinifera* development stage (Fig. 1). Further, we showed that the four analyzed microarray designs contrast with RNA-Seq gene expression platforms calling DEGs with a low fold change value (|log2_FC| < 1) (see dots blue and green in Fig.1). As observed in Figure 1 (blue dots), a considerable number of DEGs with a weak variation change have been detected exclusively by RNA-Seq. This result suggest a low sensitivity of microarray technologies especially for microarray design 1 and 2 based on short oligonucleotide probes (35-40mer) per gene model transcript calling DEGs with a weak variation change with respect to RNA-Seq approach in gene expression differential analysis.

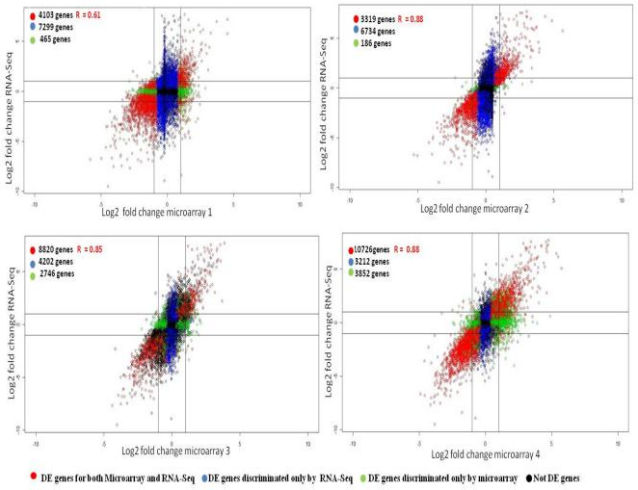


**Figure 1**. log2 fold change (log2-FC) comparison, between the four analyzed microarray design strategies and RNA-Seq in gene expression differential analysis

**2 Microarray specificity and sensitivity and gene expression level**

Previous results (Fig.1) showed that sets of genes differentially expressed detected by microarray designs and RNA-Seq technologies did not overlap for a consistent portion of DEGs (genes with a low fold change value) [8]. To further explore these discrepancies, a Receiver Operating Characteristics (ROC) curve was constructed for each microarray design assuming RNA-Seq expression data set as the reference. Each point on the ROC curve of a given microarray design and platforms represents the sensitivity on *Y-axis* (True Positive Rate: TPR) and the specificity on *X-axis* (False Positive Rate: FPR). We investigated the relationship between genes expression level expressed in Fragments Per Kilobase of exon per Million reads mapped (FPKM) and the area under curve (AUC) formed by each point of the ROC curve of each analyzed microarray. Sensitivity and specificity of microarray designs in discriminating DEGs have been performed at an FDR ≤ 0.05 and |log2-FC| ≥ 1 . Our findings showed that the AUC of all analyzed microarray designs increased with the increasing of gene expression level (Fig. 2) suggesting that microarrays exhibit a good sensitivity and specificity in gene expression differential analysis for highly expressed genes associated with a consistent fold change value. For gene expression value ≥ 15 FPKM, microarray designs 2, 3 and 4 exhibit an AUC value ≥ 90% (Fig. 2). In other words, these microarray designs carry out the same performance with regard to RNA-Seq technology in gene expression differential analysis for highly expressed genes. Also, these results demonstrated that arrays are less sensitive and less specific discriminating small variation change in gene expression differential analysis especially for lower expressed gene.

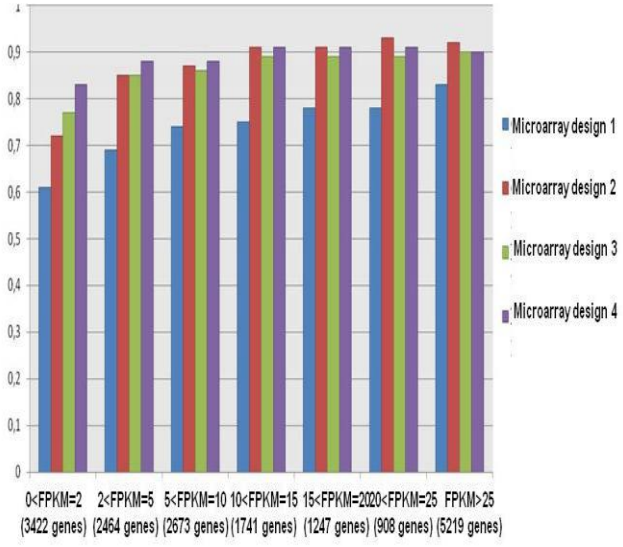


**Figure 2**. Relationship between RNA-Seq gene expression level (FPKM) and microarray AUC of ROC Curve analysis in gene expression level signal detection in microarray differential analysis.

**3 Microarray and RNA-Seq gene expression data validation by quantitative RT-PCR**

We validated the expression data of microarrays and RNA-Seq expression experiments selecting 10 genes whose expression was in agreement and in disagreement concerning the two approaches. These genes have been tested by qRT-PCR. In fact, 8 genes out the

10 selected for the qRT-PCR analysis were in disagreement between the four analyzed microarray designs and RNA-Seq (Fig. 3). It is noteworthy to observe that the expression level of these genes was less than 15 FPKM (lower expressed genes). Interestingly, this analysis showed a good agreement between qRT-PCR and RNA-Seq in fold change measurement analysis disregarding the gene expression level (Fig. 3). In fact, for the 10 analyzed genes, qRT-PCR and RNA-Seq agree for 8 and contrast for two, while qRT-PCR and microarrays (considering the four analyzed microarray designs) agree for only 2 genes JGVV1.1082 and JGVV301.10 that exhibit a high expression level (FPKM > 15). In view of the foregoing, we can suppose that the agreement between microarrays and RT-PCR in gene expression differential analysis could be depend on the gene expression level. Moreover, the Person correlation analysis based on log2 fold change measurement of the 10 randomly selected genes for the qRT-PCR analysis is higher between RNA-Seq and qRT-PCR ($R^2$:0.75) with respect to between qRT-PCR and microarray ($R^2$:0.015). This result suggests that microarray technology failed in differentially analysis for a portion of significantly differentially expressed genes probably because of the low sensitivity and specificity and the limit of microarray technology to detect accurately small variation change in differential analysis.

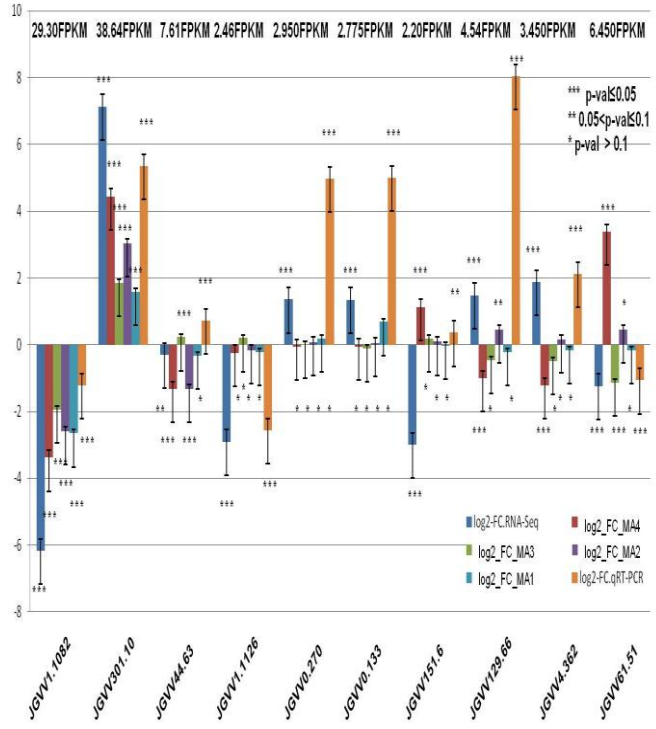


**Figure 3.** the 10 randomly chosen genes whose expression are in agreement or in disagreement concerning the two analyzed microarray and RNA-Seq approaches have been tested by quantitative RT-PCR.

## DISCUSSION AND CONCLUSION

The use of microarray and other global profiling technologies has led to a significant number of exciting new discoveries in the field of biology. Nonetheless, it is important that investigators continue to optimize microarray methodologies and that they develop new approaches aimed to producing accurate and experimentally valid data. When evaluating microarray expression data investigators must ask themselves whether the results are valid in other words, accurate for the particular biological system under study. Measurement of gene expression profiles using microarray technologies is popular among the biomedical research community. Although there has been great progress in this field, investigators are still confronted with a difficult question after completing their experiments. Further, there is no consensus concerning an approach to statistical analysis, and thus array results are analyzed in a variety of different ways [16]. However, at the very least, certain basic methods should be applied. Numerical management of the data permits removal of artifacts caused by low gene expression and low ratios. Following data pre-processing and numerical management, a statistical approach must be chosen to determine the significance of the changes in expression levels of individual genes [17]. The goal of all of these efforts is accurate identification of differences in gene expression between the sample sets, and maximal use of the information toward a better understanding of the biological process(es) under study. The reliability of the microarray results is being challenged due to the existence of different technologies and non-standard methods of data analysis and interpretation. In the absence of a reference method for the gene expression measurements, studies evaluating and comparing the performance of various microarray platforms have often yielded subjective and conflicting conclusions. To address this issue we evaluated the capacity of our four previously developed microarray design strategies based on either long and/or short multiple and/or single oligonucleotide probe per gene model transcript to accurately identified DEGs by using both RNA-Seq next generation sequencing (NGS) and qRT-PCR gene expression profiling technologies as reference because of their advantage in detection sensitivity, sequence specificity, large dynamic as well as their high precision and reproducible quantitation compare to microarray. We first evaluated the detection sensitivity and specificity of the four analyzed microarray design strategy platforms using RNA-Seq gene expression assays data set reported in Dago N 2012 work as the reference [8]. A gene was considered as differentially expressed when showing a mean difference of the expression value greater than or equal to two folds between the two considered grape *Vitis vinifera* development stages ripening and véraison at a FDR ≤ 0.05. We showed that the four analyzed microarray platforms displayed different results among them, when compared with RNA-Seq technology in gene expression differential analysis (Fig. 1). However they exhibited a reasonably good sensitivity and specificity for highly expressed genes (FPKM ≥ 15) calling DEGs associated with a consistent fold change value (fold change ≥2) in gene expression differential analysis when RNA-Seq gene expression profiling tool was assumed as reference (Fig.2), suggesting a good performance of microarray technologies in detecting DEGs for genes with a high expression level (Fig. 2) [18]. Considered as a whole, these results support the good reliability of microarray gene expression result for highly expressed gene set associated with a high gene variation change in gene expression differential analysis. We next evaluated microarray platforms performance in gene expression differential analysis by the integration of their gene expression differential analysis data with those of RNA-Seq and qRT-PCR. In this survey, only genes commonly expressed between microarrays and RNA-Seq gene expression platforms have been considered. The person correlation analysis in log2 fold change measurement between (i) RNA-Seq and qRT-PCR and (ii) microarray and qRT-PCR, of 10 randomly chosen genes out the gene set commonly expressed between microarray and RNA-Seq, proved a good correlation between RNA-Seq and qRT-PCR for lower expressed gene

associated with a small variation change between repining and véraison development stage (Fig.3). In view of the foregoing, these results demonstrated some of the limitations of microarrays to accurately detect the small gene change variation particularly for lower expressed gene. Then we validated that microarrays have acceptable sensitivity, specificity and then a good reliability in detecting differential expression, for genes with a high expression levels. This study also confirms that a certain level of fold change compression is expected for microarray platforms due to various technical limitations, including limited dynamic range, signal saturations, and cross-hybridizations. This survey as we know is the first one that evaluated different custom microarray platform designs sensitivity and specificity in discriminating DEGs in gene expression differential analysis by the integration of microarray gene expression data with those of both RNA-Seq and qRT-PCR gene expression profiling tools.

## CONCLUSION

In conclusion, this survey confirmed the superiority of RNA-Seq and qRT-PCR with respect to microarray technologies in detecting accurately DEGs in gene expression differential analysis especially for lower expressed genes. Microarrays have an acceptable sensitivity and specificity (good reliability) in detecting differential expression in particular for genes with a high expression levels and for detecting high fold change (fold change >2) in signal detection, while their specificity and sensitivity tends to be relatively low, mainly for lower expressed gene associated with a small fold change value in gene expression differential analysis.

### References

[1]MAQC Consortium et al. 2006 Nature Biotechnology (September 2006).

[2] Wang, Y., Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365, 671–679; (2005).

[3] Dago Dougba Noel, Alberto Ferrarini, Diarassouba Nafan, Fofana Inza Jésus, Silué Souleymane, Giovanni Malerba and Massimo Delledonne Probe specificity in array design influence the agreement between microarray and RNA-seq in gene expression analysis African Journal of Science and Research, (3) 5:08-12 ISSN: 2306-5877 Available Online: http://ajsr.rstpublishers.com/ (2014)

[4] Modrek, B. et al. Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res*. 29, 2850–2859 (2001).

[5] Mir, K.U. and Southern, E.M. Determining the influence of structure on hybridisation using oligonucleotide arrays Nat. Biotechnol 17 788-792 (1999)

[6] Zhang, J. Evolution for bioinformaticians and bioinformatics for evolutionists. *Evolution* 59: 2281-2283 (2005)

[7] Torres, T.T., Metta, M., Ottenwälder, B., Schlötterer, C. (2008). Gene expression profiling by massively parallel sequencing. Genome Res. 18(1): 172-177. (2008).

[8] Dago Noel. Performance assessment of different microarray designs using RNA-Seq as reference Id prodotto: 67051; Id Ugov: 404537. (2012).

[9] Sara Zenoni, Alberto Ferrarini, Enrico Giacomelli, Luciano Xumerle, Marianna Fasoli, Giovanni Malerba, Diana Bellin, Mario Pezzotti, and Massimo Delledonne: Characterization of Transcriptional Complexity duringBerry Development in Vitis vinifera Using RNA-SeqPlant Physiology vol. 152 n° 4 1787-1795. (April 2010).

[10] Smyth, G. K., Michaud, J., and Scott, H. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 21(9), 2067-2075. (2005).

[11] Benjamini, Yoav; Hochberg, Yosef Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B 57 (1): 289-300 (1995).

[12] Jaillon O. et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla; 449 (7161): 463-467. Epub 2007 Aug 26. PMID: 17721507 PubMed-indexed for MEDLINE. Nature. (2007 Sep 27).

[13]Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett. 339(1):62-6 (2003 Mar 13).

[14] Pfaffl MW A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 1;29(9):e45 (2001 May).

[15] Michael W. Pfaffl, Graham W. Horgan & Leo Dempfle Relative Expression Software Tool (REST) for group wise comparison and statistical analysis of relative expression results in real-time PCR Nucleic Acids Research 2002 30(9): E36 (2002).

[16] Slonim, D.K. From patterns to pathways: gene-expression data analysis comes of age. Nature Genet. 32, 502–508 (2002).

[17] Mutch, D.M., Berger, A., Mansourian, R., Rytz, A. & Roberts, M.A. Microarray data analysis: a practical approach for selecting differentially expressed genes. Genome Biol. 2, preprint0009 (2001).

[18] Catalin C Barbacioru, Yulei Wang, Roger D Canales, Yongming A Sun, David N Keys, Frances Chan, Karen A Poulter and Raymond R Samaha Effect of various normalization methods on Applied Biosystems expression array system data BMC Bioinformatics,7:533 doi:10.1186/1471-2105-7-533 (2006).