

## Molecular Classification of Breast Cancer Patients using Formalin-Fixed Paraffin-Embedded derived RNA Samples

Godfrey Grech<sup>1\*</sup>, Shawn Baldacchino<sup>1</sup>, Christian Saliba<sup>2</sup>, Keith Sacco<sup>1</sup>, Brendan Yee<sup>3</sup> and Christian Scerri<sup>4,5</sup>

<sup>1</sup>Department of Pathology, Medical School, University of Malta, Malta

<sup>2</sup>Centre for Molecular Medicine and Biobanking, University of Malta, Malta

<sup>3</sup>Affymetrix Inc, Santa Clara, CA, USA

<sup>4</sup>Department of Physiology and Biochemistry, University of Malta, Malta

<sup>5</sup>Molecular Genetics Clinic, Mater Dei Hospital, Malta

Rec Date: Jan 12, 2016, Acc Date: Jan 18, 2016, Pub Date: Jan 30, 2016

\*Corresponding author: Godfrey Grech, Department of Pathology, Medical School, University of Malta, Msida, MSD2090 Malta, Tel: 00356 23403047; E-mail: [godfrey.grech@um.edu.mt](mailto:godfrey.grech@um.edu.mt)

Copyright: © 2016 Grech G, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

### Abstract

The use of archival formalin-fixed paraffin-embedded (FFPE) material to analyse gene expression is limited by the low quality of extracted RNA. In this paper, we utilised an RNA based assay to quantify expression of luminal and basal markers, together with ERBB2 probes, in FFPE archival tissue from 2009 to 2010, all of which had clinical and therapeutic information of more than 5 years. Receptor status of the patients was characterised using the QuantiGene® Plex assay with 100% concordance to immunohistochemical (IHC) and fluorescence *in situ* hybridisation (FISH) results. A panel of molecular markers known to classify luminal and basal tumours were used and correlated with receptor status of the tumours. As expected, the triple negative breast cancer (TNBC) samples were classified as basal and oestrogen receptor (ER) positive cases as luminal. In summary, the QuantiGene® Plex technology provides a platform to quantitate novel panels of biomarkers on archival material. Moreover, multiplex analysis allows the use of minimal amounts of material providing an opportunity to utilise laser micro-dissected material. FFPE tissue samples are an invaluable resource for retrospective studies to interrogate current novel biomarkers, particularly to generate disease free survival and overall survival graphs to measure predictive value using well annotated retrospective samples with full clinical and pharmacological outcomes.

**Keywords:** Archival FFPE material; Multiplex assay; Biomarkers HER2; Breast cancer; Molecular classifiers

### Introduction

Investigations using RNA from archival formalin-fixed paraffin-embedded (FFPE) material is challenging due to the extent of degradation in these samples [1] and also due to the processing variables during and following cut-up of the surgical samples. Formalin cross-links nucleic acids and proteins to preserve tissue integrity, resulting in poor quality of RNA [2]. Despite improvements in quantitative PCR (qPCR) technology [3,4], the poor quality of RNA extracted from FFPE material hinders the potential of gene expression studies in this invaluable resource available in pathology archives.

The branched-chain DNA (bDNA) assay QuantiGene® technology provides a platform to perform expression studies from minimal amount of archival FFPE material without amplification of target sequences [5], overcoming the limiting factor of expression studies using degraded RNA. This assay replaces enzymatic amplification of target RNA with hybridization of specific probes followed by amplification of the reporter signal, through DNA molecule scaffold formation on bound probe-target sequences. The capture and detection probes provide increased specificity during hybridization and the short recognition sequences are designed to capture short fragments of target RNA [6]. In addition, the use of tissue homogenates directly into the assay, overcomes the low yield of RNA following RNA extraction and purification. The bDNA assay provides

the possibility to multiplex the quantification of a panel of targets [6], generating datasets within one run. Analyses using mathematical algorithms, derive the best combination of genes to normalize the target expression values and enhance the power of prediction.

Breast cancer classification is today supported by molecular markers categorising patients into four molecular classes, namely luminal A, luminal B, the human epidermal growth factor receptor 2 (HER2)-enriched and the basal types. Luminal A subtype is positive for oestrogen receptor (ER) and/or progesterone receptor (PgR) expression with low expression of Ki-67, while luminal B, apart from having an ER/PgR positive expression, includes also HER2 positive and negative subgroups associated with high Ki-67 expression. The HER2-enriched are well defined, with an exclusive high expression of HER2 receptor, due to the ERBB2 gene amplification, combined with low or absent ER and PgR. The basal type are in general negative for the latter 3 receptors, the triple negative breast cancer (TNBC) subtype [7,8]. The transcriptome analysis resulting in the definition of the luminal subtypes, along with the basal epithelial-like subgroup provided evidence of the prognostic value based on 8-year overall survival data [9]. Validation of the prognostic value of the subgroups was provided using 599 microarrays giving evidence that the molecular subtypes significantly predict the overall survival and disease free survival [10]. In this study we assessed the use of the bDNA QuantiGene® technology to classify breast cancer patients into well established molecular groups using previously defined biomarkers. The luminal and basal subgroups were correlated with receptor status and clinical outcomes to generate disease free survival

curves. The use of well annotated archival FFPE material will provide the means to reduce the time-lag between discovery of novel biomarkers and validation of prognostic or predictive value, due to the readily available clinical and pharmacological information.

## Materials and Methods

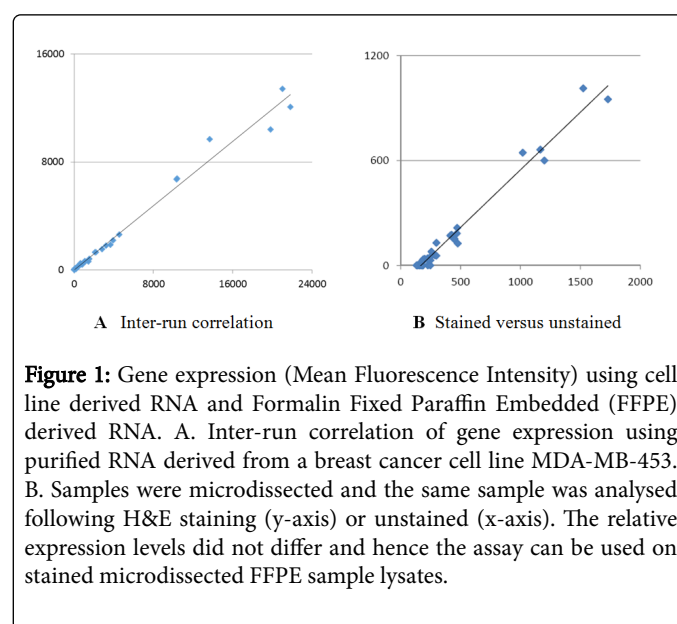
Tissue homogenates from archival breast cancer FFPE sections (6  $\mu\text{m}$ ) were prepared according to the procedure as described in the QuantiGene Sample Processing Kit for FFPE Tissues (Affymetrix Inc., Santa Clara, CA). Excess paraffin from the slides were removed by soaking the slides in xylene for 5 minutes followed by soaking in 70% ethanol for 5 minutes. The slides were air dried and the tissue from the slides were removed using a clean razor blade and placed into 1.5 mL tube. 300  $\mu\text{L}$  of homogenizing solution supplemented with 2  $\mu\text{L}$  of proteinase K (50 ug/  $\mu\text{L}$ ) were added to the tubes. The tubes were vortexed for 30 seconds and incubated at 65°C for 4 hours. The tissue homogenate was separated from the debris by brief centrifugation, and then transferred to a new tube.

Gene expression analysis was performed using a QuantiGene 10 Plex assay (Affymetrix Inc, Santa Clara, CA). The 12 genes measured in this study included EN1, FOXC1, GABRP, CA12, AGR2, TTF3, GATA3 ESR1 SCNN1A, ERBB2 and the housekeeping genes PPIB and HPRT. Tissue homogenates were transferred to a 96 well hybridization plate containing QuantiGene Plex probe and Luminex beads sets. Each bead type was coated with a different single-strand DNA capture probe (CP). Several other components of the QuantiGene Probe set are also comprised of single-strand DNA oligonucleotides including the capture extenders (CEs), label extenders (LEs), and blocking probes (BPs). Parts of CE oligonucleotides are complementary to the target mRNA (covering ~500 bases). The QuantiGene Plex Assay was performed according to the manufacturer's manual with all of the reagents and consumables supplied by the manufacturer. (QuantiGene Plex 2.0 Assay, Affymetrix Inc, Santa Clara, CA). For each well of the 96 well hybridization plate, 40  $\mu\text{L}$  of tissue homogenate were transferred to the plate where each well already contained 60  $\mu\text{L}$  of working bead mix (33.3  $\mu\text{L}$  of lysis mixture, 18.5  $\mu\text{L}$  of nuclease free water, 0.2  $\mu\text{L}$  of proteinase K solution, 2  $\mu\text{L}$  of blocking reagent, 5  $\mu\text{L}$  of probe set, 1  $\mu\text{L}$  of magnetic Luminex beads; customised QuantiGene Plex Set Panel). Hybridization was performed overnight for 18 hours at 54°C, with shaking at 600 rpm. The hybridization mixtures were then transferred to a 96-well flat bottom plate. An Affymetrix handheld magnetic bead washer (Affymetrix P/N QP0702) was used to wash the beads, thus removing all unbound materials. 100  $\mu\text{L}$  of preamplifier reagent was added to each assay well. The magnetic separation plate was sealed with adhesive backed foil and incubated for 1 hour at 50°C and 600 rpm. The unbound preamplifier reagent was removed, and beads were washed three times with 100 ml of wash buffer using the handheld magnetic washer. This was followed by similar incubation and washing steps with 100  $\mu\text{L}$  of Amplifier reagent, followed by 100  $\mu\text{L}$  of Label Probe reagent, and finally followed by 100  $\mu\text{L}$  SAPE working reagent. Signals from the beads were measured with a Luminex 200 (Luminex Corp., Austin, TX), after re-suspending the beads in 130  $\mu\text{L}$  of SAPE wash buffer, using dd gate settings of 5,000–25,000. Per target, 50–100 beads were measured in a sample volume of 100  $\mu\text{L}$ .

## Results

### Branched-chain DNA technology provides consistent quantitation of gene expression

To evaluate the inter-run variability of the bDNA assay using the customised probe set, the analysis was done using RNA derived from cell lines. In addition to the 12 capture probes used in this study, we measured RNA expression of an additional 10 genes that were previously quantified using qPCR. Inter-run regression analysis using the 22-plex assay provide evidence that the gene expression results using the customised probe sets are performing well between runs with an  $r^2 > 0.99$  (Figure 1A). In addition, correlation ( $r^2$ ) of normalised expression data using the bDNA assay and qPCR, on the previously quantified genes, was  $> 0.86$  (data not shown). We then evaluated the possibility to use haematoxylin and eosin (H&E) staining to microdissect tumour sites with higher precision. Using serial sections of the same sample and the same coordinates from the H&E stained slide we compared stained versus unstained tissue. Regression analysis show a high correlation of gene expression using the 22-plex assay with an  $r^2 > 0.98$  (Figure 1B).



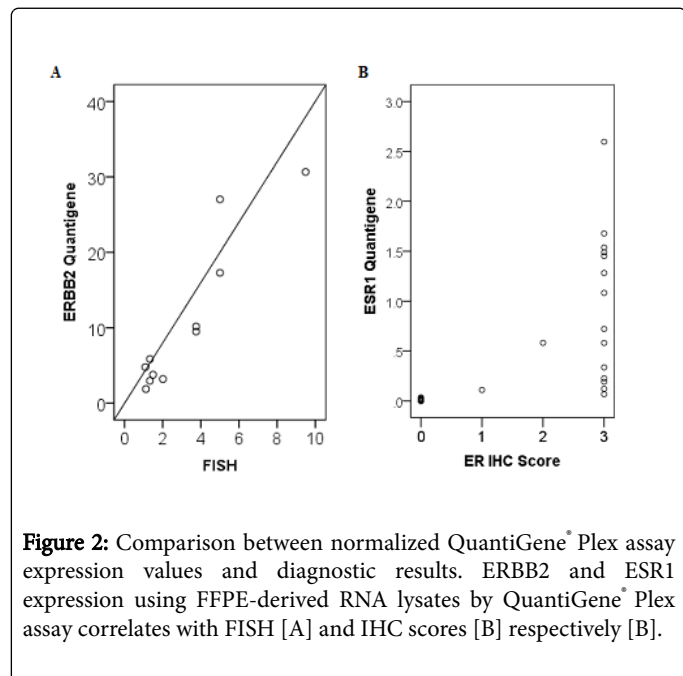
**Figure 1:** Gene expression (Mean Fluorescence Intensity) using cell line derived RNA and Formalin Fixed Paraffin Embedded (FFPE) derived RNA. A. Inter-run correlation of gene expression using purified RNA derived from a breast cancer cell line MDA-MB-453. B. Samples were microdissected and the same sample was analysed following H&E staining (y-axis) or unstained (x-axis). The relative expression levels did not differ and hence the assay can be used on stained microdissected FFPE sample lysates.

### ERBB2 expression correlates with that of other well established/benchmark techniques

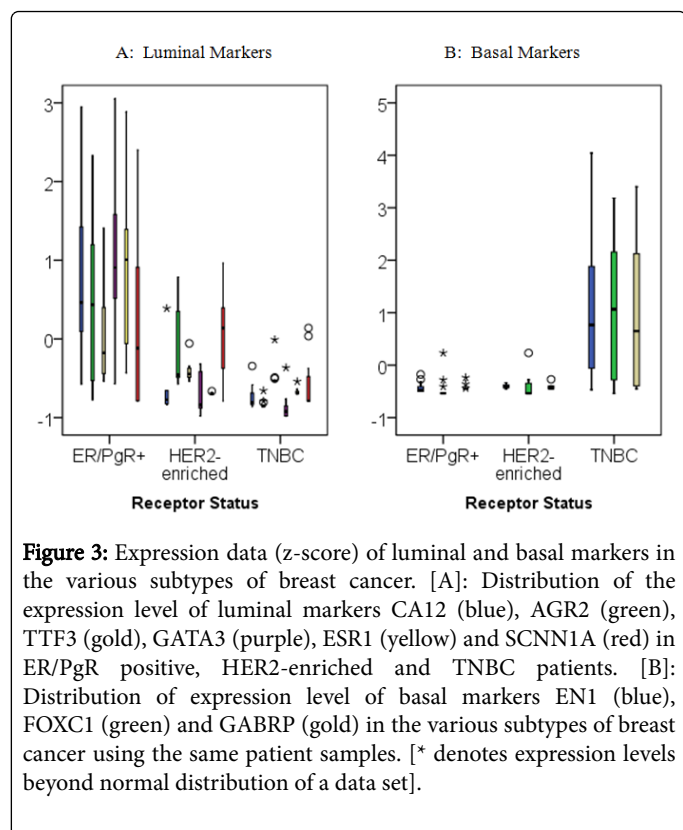
The receptor status of tumours was used as an internal control of the expression runs. The results obtained using the bDNA QuantiGene® Plex assay were correlated with known ERBB2 status using immunohistochemical (IHC) and fluorescence in situ hybridisation (FISH) results (defined as a signal ratio of ERBB2 probe to CEP17 control probe). The expression data correlates with both IHC (data not shown; n=37) and FISH (Figure 2A; n=11) results using Spearman's rho non-parametric correlation and Pearson's parametric correlation, respectively. In addition, oestrogen receptor (ESR1) expression correlates strongly with IHC results (Spearman's rho: 0.878 p-value  $< 0.000$ ); (Figure 2B).

### Basal markers are exclusively upregulated in TNBC patients

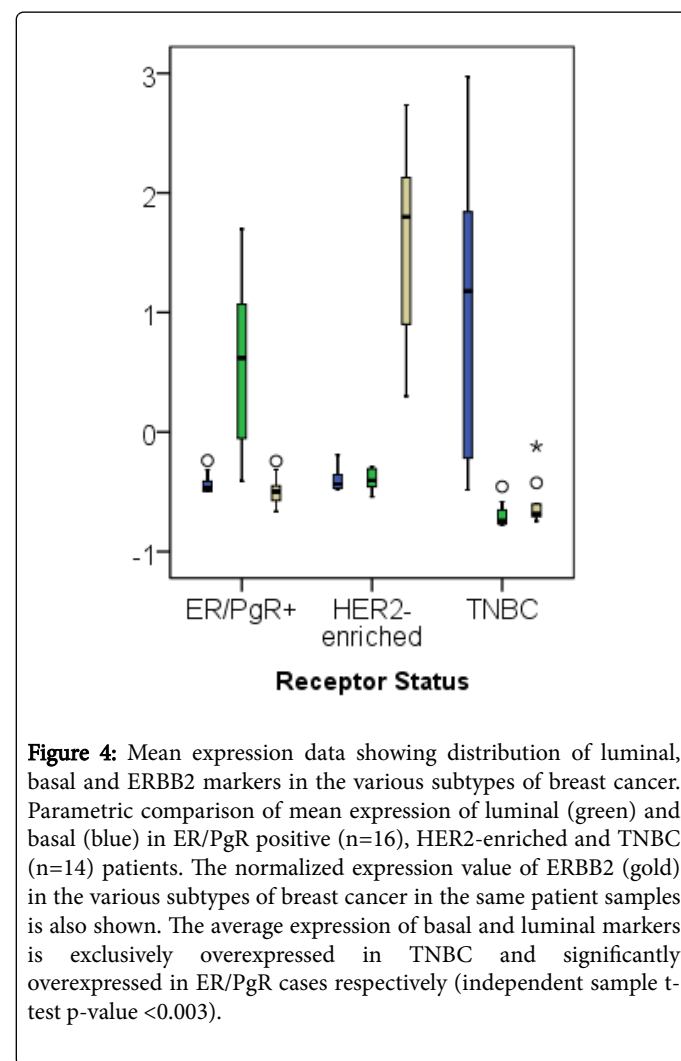
The expression values of 6 luminal markers, 3 basal markers and ERBB2 were normalised against a set of reference genes.



The luminal marker ESR1 (oestrogen receptor; yellow), as expected show a significantly higher expression in the ER/PgR positive group only (Figure 3A).

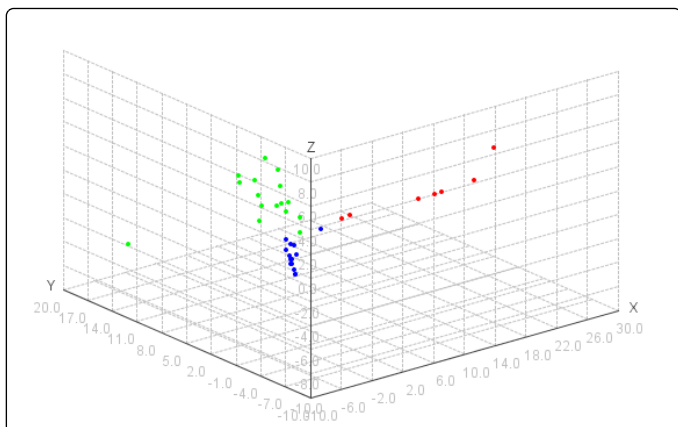


Similarly all luminal markers investigated in this panel were expressed relatively higher in the ER/PgR patients as compared to the TNBC patients, while 4/6 were moderately higher in the HER-2 enriched subtype. Of interest, there are some patients that have expression of some markers that are beyond the normal distribution of a specific expression dataset. One of these patients was shown to have heterogeneous tumour sites (data not shown). The basal markers EN1 (blue), FOXC1 (green) and GABRP (gold) are exclusively upregulated in the TNBC patients, known to be generally of a basal-type (Figure 3B). Moreover, the mean expression value of ERBB2 is significantly higher only in the HER2-enriched subgroup (Figure 4).

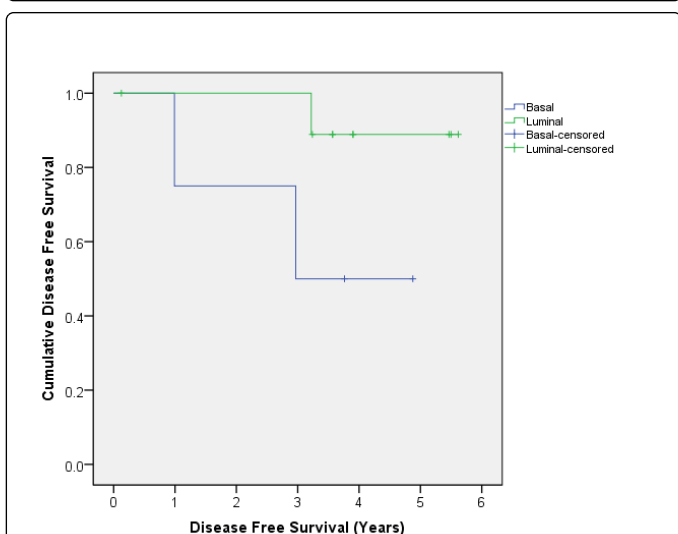


### Expression values of the 10-genes, molecularly classify breast cancer patients

As shown above, the luminal and basal markers matched the expected receptor status. Principal component analysis (PCA), generated well defined luminal (green), basal (blue) and ERBB2 (red) groups from QuantiGene® Plex assay-derived normalised expression values of the 10 classifier genes (Figure 5). Based on this classification, correlation of clinical outcome data show that the basal subgroup have a worse prognosis (Figure 6). Hence, using the QuantiGene® Plex assay to measure expression data of 10 classifier genes, provides well defined patient subtypes using archival material.



**Figure 5:** Principal component analysis (PCA) showing luminal, basal and ERBB2 subgroups. Normalised expression values of the 10 classifier genes generated by the QuantiGene® Plex assay predict the various breast cancer subgroups. The blue subgroup represent patients with high expression of basal markers; the green subgroup represent patients with high expression of luminal markers and the red subgroup represent a high ERBB2 expression.



**Figure 6:** Disease free survival across luminal and basal subgroups. Disease free survival data of a subset of patients diagnosed in 2009 (n=42) were used to derive a 5-year disease survival Kaplan Meier curve following subgrouping using the biomarker panel.

## Discussion

Various attempts to modify protocols aiming to extract RNA and amplify target sequences in FFPE-derived samples provided proof of principle modifications [11], but require laborious validation protocols to ensure minimal bias due to PCR efficiency in between samples. Criteria to select FFPE-derived RNA with sufficient quality for gene expression studies show that less than 25% of the extracted RNA could be used for microarray analysis [12].

Previous studies show that matched frozen and FFPE-derived RNA provide material to successfully perform qPCR assays [13,14]. RNA yield in FFPE material is too low to perform multiple gene analysis and require RNA amplification protocols [15]. Such protocols yield gene expression profiles with high sensitivity and reproducibility, but still require validation assays and an extensive workload, which is unsuitable for clinical applications. The outcome of these studies provide evidence for reproducibility and sensitivity of detection based on statistical driven algorithms to compare RNA from different origins. Ideally the outcome should be measured against known and well validated gene panels that are known to be differentially expressed between patient groups and derive prognostic information. Comparison of the bDNA assay and qPCR showed a significantly higher coefficient of reliability for the bDNA assay (93-100%) and a 10-fold increase in sensitivity using the bDNA assay [16].

In this study we utilise the QuantiGene® Plex assay to measure the expression of 10 genes simultaneously using micro-dissected material from a set of serial sections used also for IHC and histological analysis. Hence, the RNA studies were done of the same material used for diagnosis. In addition, micro-dissection was performed using H&E stained sections ensuring the isolation of a specific tumour site and matched normal tissue from the same section. Tumours with multi-loci were also micro-dissected in different tubes, if present. The 12-plex assay measures well validated luminal and basal classifier genes, the ERBB2 gene and 2 reference genes, providing statistical evidence of the expected receptor status. In addition there was full concordance between ERBB2 QuantiGene® Plex expression and the respective FISH results (Figure 2). The readily available clinical outcome data provided the means to generate disease free survival plots immediately (Figure 6). Novel biomarkers can be used in such an assay to generate prognostic value using retrospective studies. Of interest, using the tissue homogenate allows re-running of the same sample for 6 additional times in the case of whole FFPE sections or 2 additional times when using micro-dissected material. Hence, novel biomarkers can be assessed together with receptor status and other gene classifiers using the same tissue homogenate.

Our study highlights that the bDNA QuantiGene® Plex assay provides the platform to combine histopathology and biomarker analysis at the clinical diagnostic laboratory, providing a robust tool to significantly improve patient management in the clinic. In addition, molecular classification into therapeutic groups shall provide a better understanding of actionable pharmacogenetic biomarkers [17], and potentially targeting new proteins for better treatment [18]. The proper use of biomarkers and the utilization of archival material, with readily available clinical outcome data, provides proper implementation of personalised medicine [19].

## Acknowledgements

This study was funded by the Faculty of Medicine and Surgery, University of Malta and the Italia Malta Genome Breast Cancer Cross Border Risk Surveillance (ImaGenX) project financed under Operational Program Italia Malta 2007-2013 and co-financed by the University of Malta.

## References

1. Abrahamsen HN, Steiniche T, Nexø E, Hamilton-Dutoit SJ, Sørensen BS, et al. (2003) Towards quantitative mRNA analysis in paraffin-embedded tissues using real-time reverse transcriptase-polymerase chain reaction: a

- methodological study on lymph nodes from melanoma patients. *J Mol Diagn* 5: 34-41.
2. Macabeo-Ong M, Ginzinger DG, Dekker N, McMillan A, Regezi JA, et al. (2002) Effect of duration of fixation on quantitative reverse transcription polymerase chain reaction analyses. *Mod Pathol* 15: 979-987.
  3. Cronin M, Pho M, Dutta D, Stephans JC, Shak S, et al. (2004) Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am J Pathol* 164: 35-42.
  4. Masuda N, Ohnishi T, Kawamoto S, Monden M, Okubo K (1999) Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res* 27: 4436-4443.
  5. Yang W, Maqsoodi B, Ma Y, Bui S, Crawford KL, et al. (2006) Direct quantification of gene expression in homogenates of formalin-fixed, paraffin-embedded tissues. *Biotechniques* 40: 481-486.
  6. Flagella M, Bui S, Zheng Z, Nguyen CT, Zhang A, et al. (2006) A multiplex branched DNA assay for parallel quantitative gene expression profiling. *Anal Biochem* 352: 50-60.
  7. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, et al. (2011) Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer. *Ann Oncol* 22: 1736-1747.
  8. Schnitt SJ (2010) Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy. *Mod Pathol* 23 Suppl 2: S60-64.
  9. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98: 10869-10874.
  10. Kapp AV, Jeffrey SS, Langerød A, Børresen-Dale AL, Han W, et al. (2006) Discovery and validation of breast cancer subtypes. *BMC Genomics* 7: 231.
  11. Li J, Smyth P, Cahill S, Denning K, Flavin R, et al. (2008) Improved RNA quality and TaqMan Pre-amplification method (PreAmp) to enhance expression analysis from formalin fixed paraffin embedded (FFPE) materials. *BMC Biotechnol* 8: 10.
  12. Penland SK, Keku TO, Torrice C, He X, Krishnamurthy J, et al. (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. *Lab Invest* 87: 383-391.
  13. Glenn ST, Jones CA, Liang P, Kaushik D, Gross KW, et al. (2007) Expression profiling of archival renal tumors by quantitative PCR to validate prognostic markers. *Biotechniques* 43: 639-640, 642-643, 647.
  14. Rogerson L, Darby S, Jabbar T, Mathers ME, Leung HY, et al. (2008) Application of transcript profiling in formalin-fixed paraffin-embedded diagnostic prostate cancer needle biopsies *BJU Int* 102: 364-370.
  15. Tyekucheva S, Martin NE, Stack EC, Wei W, Vathipadikeal V, et al. (2015) Comparing Platforms for Messenger RNA Expression Profiling of Archival Formalin-Fixed, Paraffin-Embedded Tissues. *J Mol Diagn* 17: 374-381.
  16. Knudsen BS, Allen AN, McLerran DF, Vessella RL, Karademos J, et al. (2008) Evaluation of the branched-chain DNA assay for measurement of RNA in formalin-fixed tissues. *J Mol Diagn* 10: 169-176.
  17. Sacco K, Grech G (2015) Actionable pharmacogenetic markers for prediction and prognosis in breast cancer. *EPMA J* 6: 15.
  18. Baldacchino S, Saliba C, Petroni V, Fenech AG, Borg N, et al. (2014) Deregulation of the phosphatase, PP2A is a common event in breast cancer, predicting sensitivity to FTY720. *EPMA J* 5: 3.
  19. Grech G, Zhan X, Yoo BC, Bubnov R, Hagan S, et al. (2015) EPMA position paper in cancer: current overview and future perspectives. *EPMA J* 6: 9.