Application Note

Rescue of low RIN RNA from FFPE samples and improved RNA sequencing using QIAseq® RNA-seq solutions

Desmots Fabienne[1,2], Santamaria Kathleen[1], Chettouh Hamza[3] and Fest Thierry[1,2]

[1] INSERM, UMR1236, Université de Rennes 1, 2 avenue du Pr. Léon Bernard, Rennes, F-35065, France, [2] Laboratoire d'hématologie, Centre Hospitalier Universitaire, 2, rue Henri Le Guilloux, 35033 Rennes Cedex 9, France; [3] QIAGEN GmbH, Hilden, Germany

Introduction

RNA sequencing (RNA-seq) has become a central tool in transcriptome analysis. In studies comparing healthy individuals and follicular lymphoma (FL) patients, RNA-seq can differentiate between modifications of transcriptomic signatures and compare patient responses to new epigenetic therapeutic approaches. Most of these studies are based on FFPE samples from large retrospective patient cohorts. However, the quality of FFPE samples is generally poor and depends on several parameters, such as fixation or sample storage conditions [1]. Despite typically poor RNA integrity and low yields, FFPE samples are a critical source of molecular gene expression data for retrospective studies.

In this study, we successfully performed RNA-seq using novel QIAGEN® RNA-seq technologies on FFPE tumor samples from FL patients, either at diagnosis or after receiving an HDACi treatment. These FFPE samples were rejected by two NGS service providers for library preparation and sequencing. We also performed identical RNA-seq on other challenging samples of B-cell origin, such as a very small number of sorted cultured cells from tonsils and lymph nodes. Collectively, QIAseq RNA-seq technology for rRNA removal and stranded RNA library preparation enables the discovery of expected B-cell signatures and other interesting signatures in the context of epigenetic treatment.

Materials and methods

Samples

FL tumor samples were obtained from patients recruited for a phase I/II study involving an oral pan-HDACi drug [2]. The samples (FFPE-treated) were obtained via surgical biopsy either at diagnosis (non-treated) or after HDACi treatment at the time of their refractory/relapsed FL (treated). Primary B cells were isolated from human blood samples from healthy volunteers using B Cell Isolation Kit II, human (Miltenyi Biotec) and purified using BD FACSAria™ II system (BD Biosciences). For the Burkitt-derived L3055 cell line, 200,000 cells/ml were seeded in 24-well plates in RPMI1640 plus 3% fetal calf serum (FCS). All patients were recruited under institutional review board approval and informed consent process according to the Declaration of Helsinki.

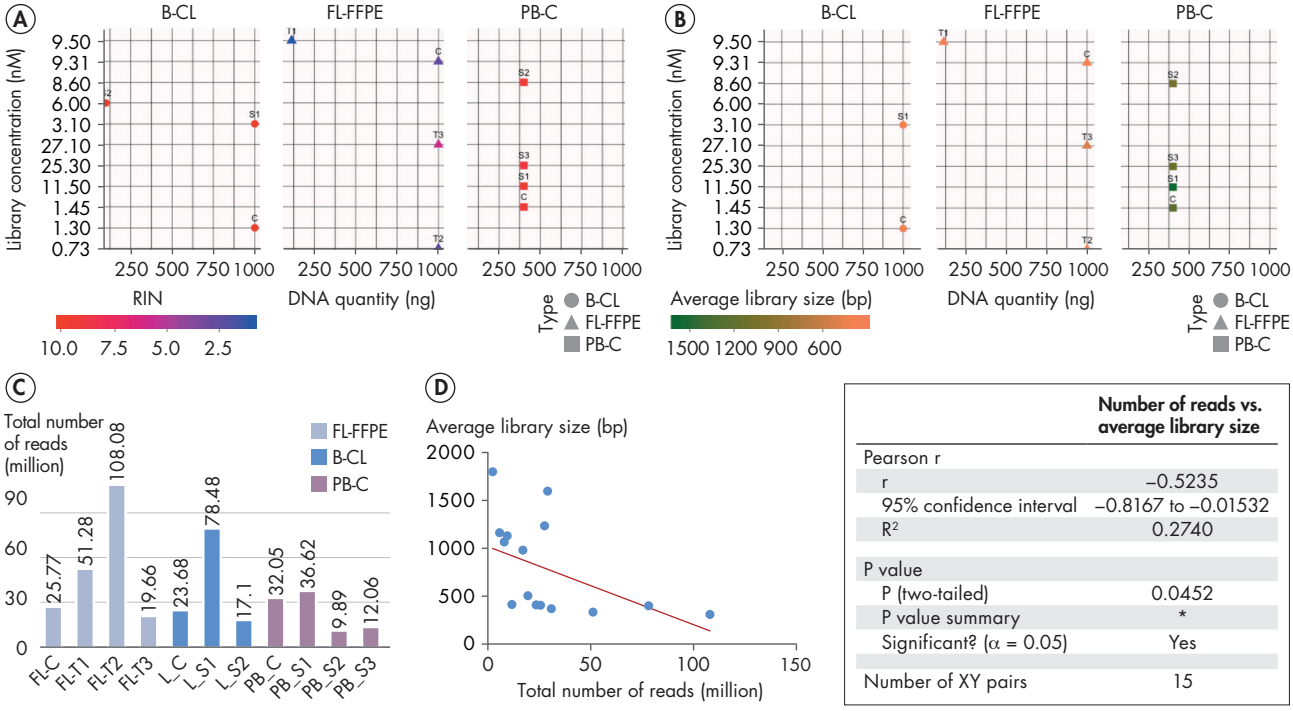
RNA extraction and sequencing

For RNA extraction from primary B-cells and the cell line, the NucleoSpin® RNA XS (Macherey-Nagel) and the AllPrep® DNA/RNA Mini Kit from QIAGEN were used, respectively. For FFPE samples, we used the RecoverAll™ Total Nucleic Acid Isolation Kit (Thermo Fischer Scientific; All multi-sample RNA/DNA workflow; 2 to 5 tissue sections per sample).

To prepare stranded RNA-seq libraries, we used the QIAseq Stranded Total RNA Lib Kit (QIAGEN) in tandem with the QIAseq FastSelect™ –rRNA HMR Kit (QIAGEN), which removes rRNA. Sample fragmentation was not performed for the FFPE samples, and the number of cycles (for the amplification step) was set up according to the DNA concentration (11 cycles for 100 ng and 16 cycles for 1000 ng). Samples were sequenced on the Illumina® NextSeq™ 500 using a paired-end dual indexing mid-output flowcell (2 x 150 bp).

RNA-seq analysis

For RNA-seq data analysis, an in-house pipeline from the investigator's lab was used. It notably integrates the Spliced Transcripts Alignment to a Reference (STAR) tool [3] for alignment to the Hg38 genome reference and the feature Counts tool [4] to estimate the number of reads that are assigned to the genome. The RSeQC package was used to estimate the quality, the genome distribution of reads and the alignment to ribosomal gene coordinates. DESeq2 was used to estimate the differential expression between samples. Conditions were set to extract only those genes with a 1.5-fold change in expression and with a p-value of 0.05. Upon applying this limit, we extracted 9558 differentially expressed genes. For data visualization, 800 most differentially expressed genes were selected and analyzed using R packages such as d3heatmap, heatmaply, EMA and ggplot2. For sashimi plots, IGV interface was used.



	Number of reads vs. average library size
Pearson r	
r	–0.5235
95% confidence interval	–0.8167 to –0.01532
R^2	0.2740
P value	
P (two-tailed)	0.0452
P value summary	*
Significant? (α = 0.05)	Yes
Number of XY pairs	15

Figure 1. Quality parameters of QIAseq RNA libraries prepared from cultured B cell-line (B-CL), primary B- cell (PB-C) or FFPE follicular lymphoma FFPE samples (FL-FFPE). [A] RIN score and [B] average library size after fragmentation in base pairs (bp) relative to final library concentration; [C] Total number of reads sequenced per samples; FL-FFPE samples at diagnosis and after treatment are noted FL_C and FL_T1, T2 or T3; 3 different patients respectively, B cell-line samples control or in stimulated conditions are noted L_C and L_S1 or S2 (2 samples) respectively, primary B- cell control or in stimulated conditions are noted PB_C and PB_S1, S2 or S3 (3 samples), respectively [D] correlation with average library size in bp.

Results and discussion

Improved RNA-seq from poor-quality and low-yield samples

Initially, we approached two commercial NGS service providers to isolate RNA and prepare RNA libraries from the FFPE patient samples. However, both commercial providers, rejected these samples for library preparation, due to low amounts of material and low RIN numbers. We then attempted RNA library construction using the QIAseq FastSelect –rRNA HMR Kit for removal of ribosomal RNA followed by RNA-seq library construction with the QIAseq RNA Stranded Library Kit. RNA-seq performed using QIAseq solutions were suitable for samples of poor quality and low concentrations: (1) B-cell-origin samples such as cultured B cell-line (B-CL) or primary B cell (PB-C) in basal (C) or stimulated (S) conditions with various cytokines and (2) FFPE samples from FL patients (FL-FFPE) collected at diagnosis (C) or after receiving HDACi treatment (T). In each case, we evaluated how the quality of samples influenced the final library concentration (Figure 1A). In the case of FFPE samples, sample T1 – with poor RNA integrity and low yield – gave the highest library concentration compared to all other samples. Starting with higher RNA concentration for FFPE samples (1000 ng of FFPE-RNA) did not always lead to higher library concentrations. Likewise, for PB-C and other samples, similar starting quantities of RNA and higher RIN scores gave variable library concentrations. We, therefore, asked whether fragmentation size could influence the library preparation (Figure 1B). The FFPE samples showed an average size of 600 bp with no fragmentation. We observed that fragment size is inversely proportional to library yield, particularly for PB-C samples. Hence, fragmentation optimization is an important step in the preparation of the libraries using the QIAseq procedure. After sequencing, the number of reads per sample was heterogeneous among samples (Figure 1C) and correlated to the fragmentation size of samples (P=0.0452) (Figure 1D).

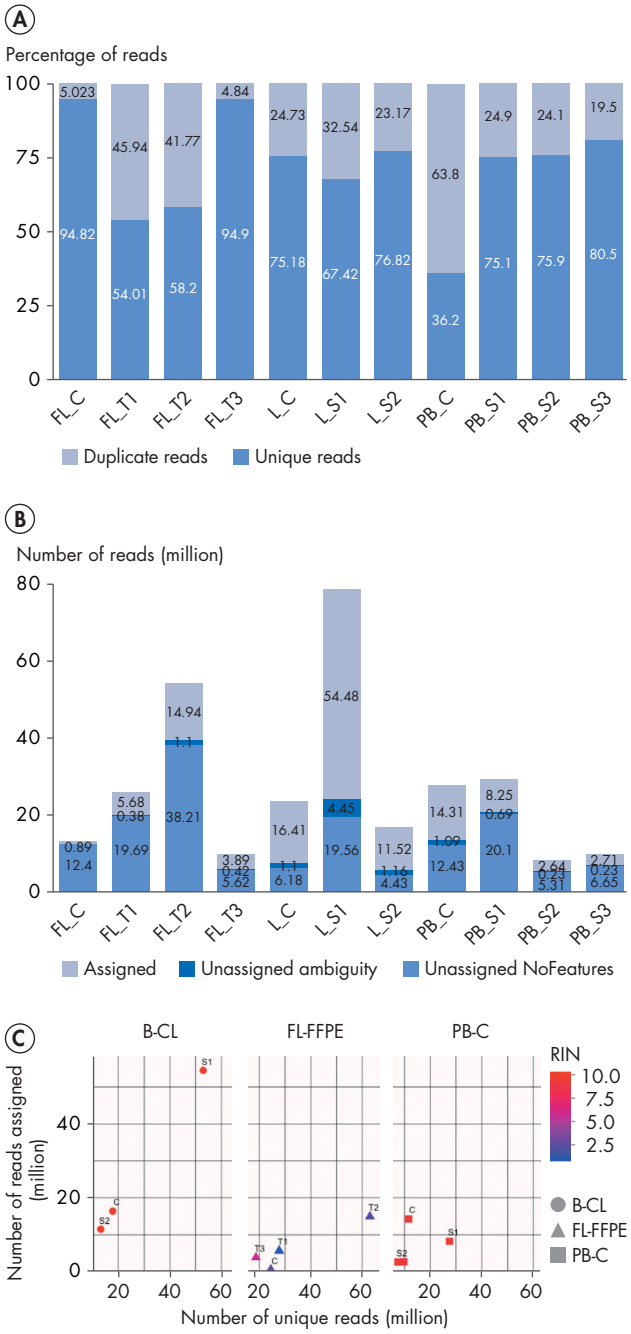


Figure 2. Sequenced read quality of libraries prepared from cultured B cell-line (B-CL), primary B cell (PB-C) or FFPE follicular lymphoma FFPE samples (FL-FFPE). [A] Proportion of unique and duplicate reads per samples; [B] STAR alignment results; [C] Number of reads assigned in unique reads relative to RIN scores. FL-FFPE samples at diagnosis (FL_C) and after treatment (FL_T1, FL_T2 or FL_T3 from 3 different patients); B cell-line samples control (L_C) or in stimulated conditions (L_S1 or L_S2; 2 samples); primary B-cell control (PB_C) or in stimulated conditions (PB_S1, PB_S2 or PB_S3; 3 samples).

High library complexity and complete rRNA removal

STAR alignment scores indicate that a high library complexity – 54–95% of reads uniquely mapped to the genome – can be obtained with FFPE samples (Figure 2A). However, the reads assigned to a gene from the annotation GRCh38-Ensembl 94 were highly heterogeneous between samples, when duplicate reads were removed using a custom pipeline. For FFPE samples, only 1–15 million reads could be assigned to known genes (Figure 2B). Better results were obtained with the cell line samples; they had lesser duplicate reads and 4–55 million reads could be assigned to known genes. Therefore, for FFPE samples, only very few reads were assigned, likely due to the poor quality of the RNA (Figure 2C).

To further assess the quality of data generated from FFPE samples, we compared the read distribution between samples. Sashimi plot of read density for three different samples along exons of the *XBP1* gene showed few unexpected reads between exons in FFPE samples. For B-cell samples, poor-quality reads couldn't be aligned to coding regions (Figure 3A). For FFPE samples, the majority of the reads were assigned to exons (in CDS, 3'UTR or 5'UTR) and a small fraction to introns (Figure 3B). None of the reads originated from ribosomal RNA (Figure 3C).

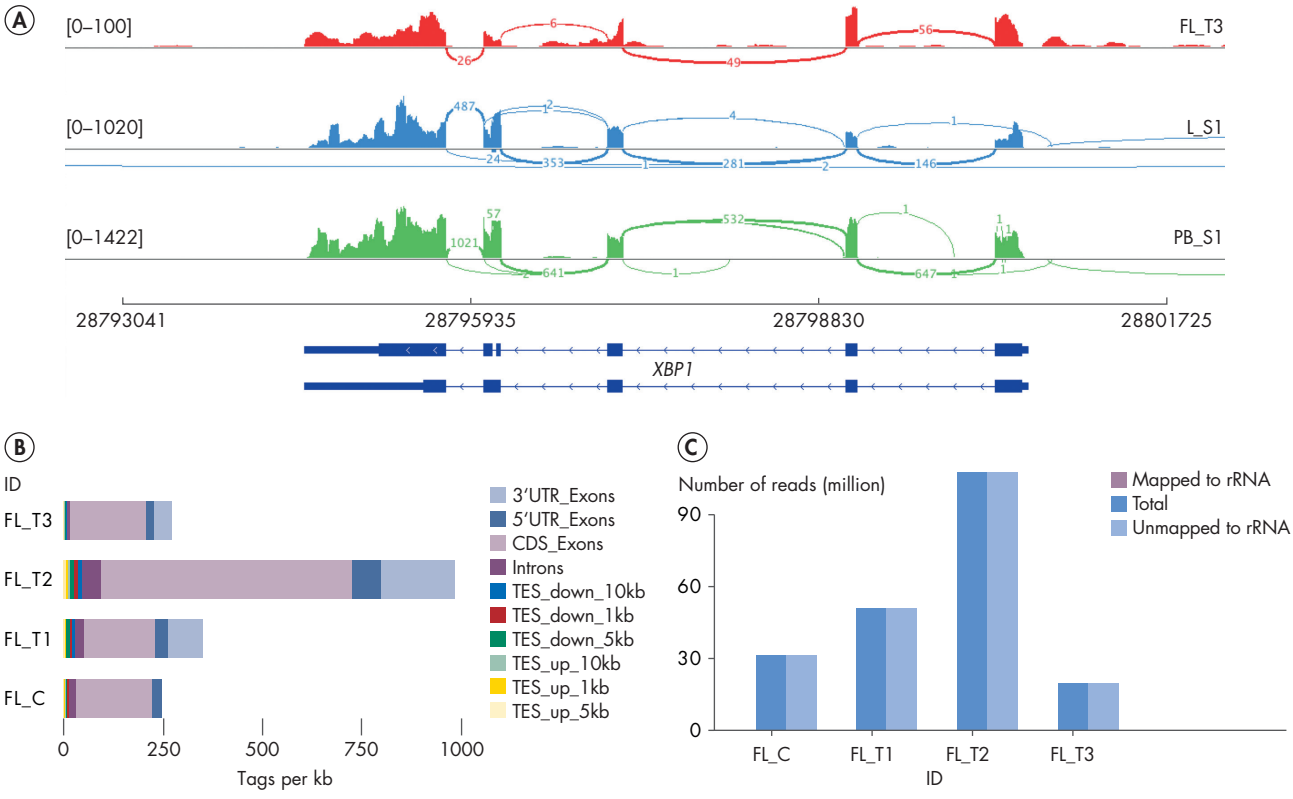


Figure 3. Read distribution and rRNA contamination. [A] Sashimi plot of read density for three different samples along exons of *XBP1* gene; [B] Read distribution over genome features; [C] Estimation of reads mapped on ribosomal RNA. FL-FFPE samples at diagnosis (FL_C) and after treatment (FL_T1, FL_T2 or FL_T3 from 3 different patients); B cell-line samples control (L_C) or in stimulated conditions (L_S1 or L_S2; 2 samples); a primary B-cell in stimulated conditions (PB_S1).

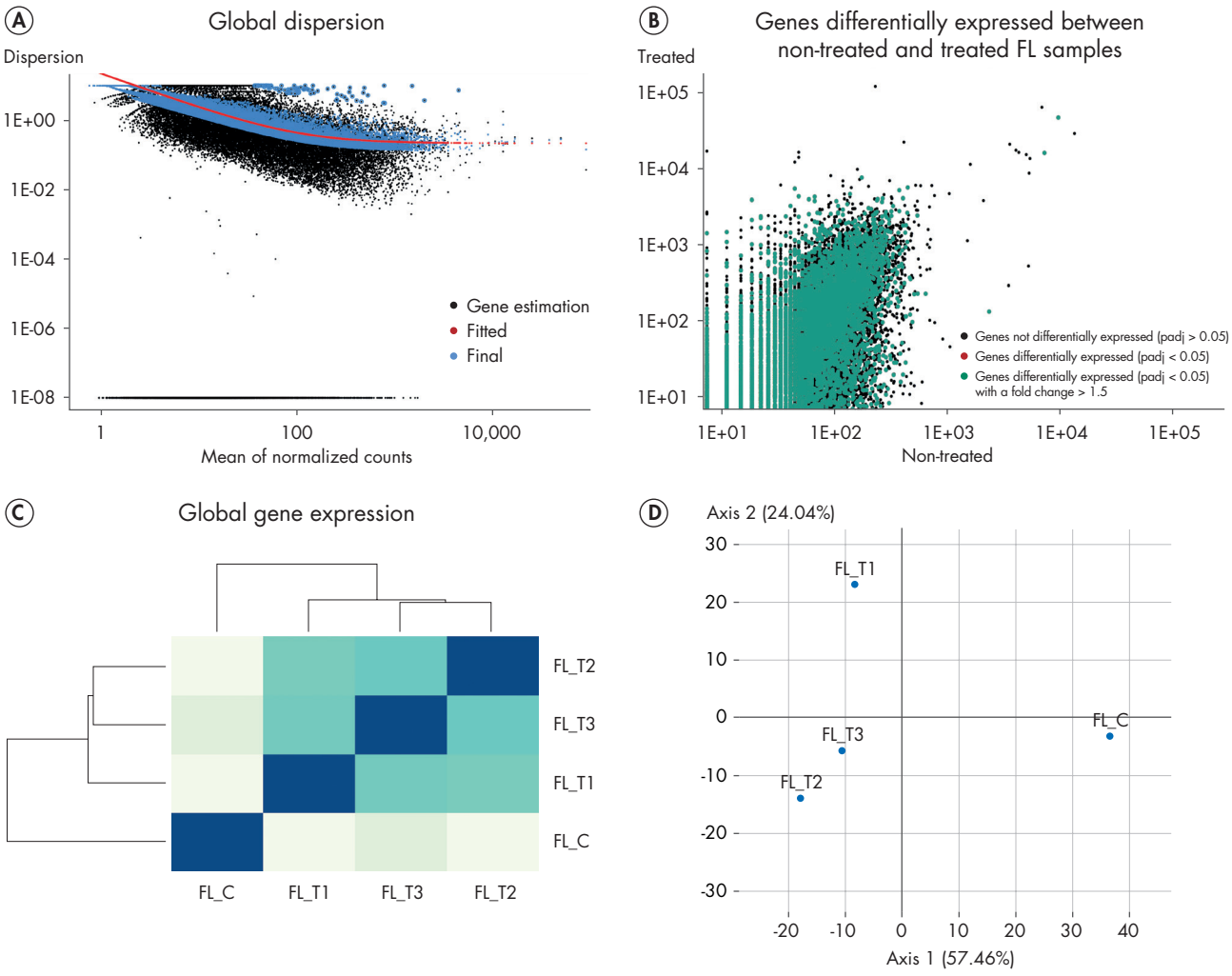


Figure 4. DESeq2 results on FFPE samples. [A] Dispersion plot showing a typical dispersion estimate; [B] Genes differentially expressed between non-treated and treated FL samples; [C] Heatmap of the distance matrix giving an overview of similarities and dissimilarities between samples; [D] Principal component plot of the samples. FL-FFPE samples at diagnosis and after treatment are noted FL_C and FL_T1, _T2 or _T3 (3 different patients), respectively.

Known and novel transcriptomic signatures identified

DESeq2 method was used to compare differential gene expression between HDACi-treated (FL_T1, _T2 and _T3) and non-treated FFPE samples (FL_C), using the number of sequence fragments assigned to each gene. A large number of genes were differentially expressed with at least a 1.5-fold change (adjusted $p<0.05$) (Figures 4A and 4B). Interestingly, the heatmap of the distance matrix between samples indicates that treated samples are more similar in global gene expression than non-treated samples (Figure 4C). The x-axis of the principal component plot shows that transcriptomes from patients are modified upon HDACi treatment (Figure 4D). Each patient showing a different

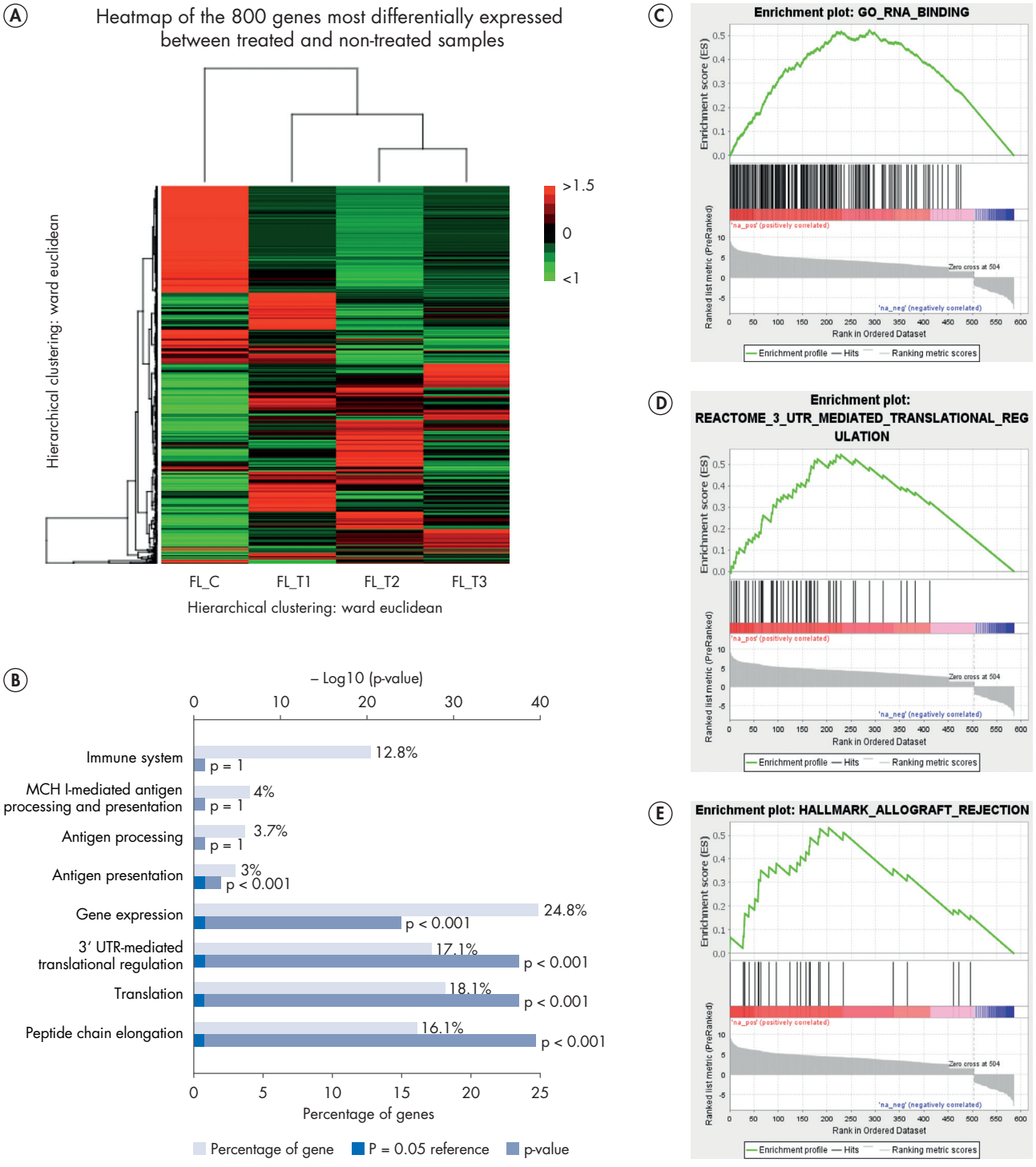


Figure 5. Results of global gene expression profiling. [A] Unsupervised hierarchical clustering; [B] Major pathway enrichment of genes differentially expressed using FunRich tool; [C-E] Major gene set enrichment found with a GSEA analysis. FL-FFPE samples at diagnosis and after treatment are noted FL_C and FL_T1, _T2 or _T3 (3 different patients), respectively.

expression pattern (as shown in the y-axis of the principal component plot) is in accordance with heterogeneous drug responses usually seen in FL patients.

Global gene expression signatures were also compared between HDACi-treated and non-treated samples using unsupervised hierarchical clustering (Figure 5A). This signature corresponds to the global influence of the HDACi treatment on the transcriptome of patients. Profile differences among treated patients involve distinct pathways, which in this case are most likely related to their specific and heterogeneous response to the drug. As expected, the functional enrichment pathways corresponding to the differentially expressed genes are involved in lymphoproliferation, B-cell regulation (Ag processing and presentation) or immune response (Figure 5B). The novel immune system signature discovered indicates that HDACi also regulates crosstalk between immune cells and B-cells. Not surprisingly, we also detected epigenetic modification-related signatures (Figure 5B, C, D), such as gene expression, RNA-binding and 3′UTR-mediated translation pathway signatures. We also identified enrichment in allograft rejection (Figure 5E), which could be attributed to the drug's effect in enhancing the expression of genes that increases suppressive functions [5].

Conclusion

We successfully performed RNA-seq using novel QIAGEN RNA-seq technologies on poor-quality FFPE RNA samples from FL patients, which were originally rejected by two NGS service providers for library preparation and sequencing. We demonstrated in this study that the QIAseq library preparation process is well suited for RNA samples of poor quality and low yield, such as FFPE samples and sorted cultured cells. We also demonstrated that the QIAseq FastSelect –rRNA HMR Kit led to near-complete rRNA removal in FFPE samples. The RNA-seq analysis identified a new transcriptomic signature in patients receiving an epigenetic treatment. In addition to confirming known targets such as some very important genes that are re-expressed again in patients receiving treatment [6], we identified new target genes that could explain differences in drug response. These findings are highly relevant in the context of HDAC inhibitors being explored as promising drugs for follicular lymphoma treatment.

Acknowledgments

Sequencing was performed in the Biogenouest Genomics/ Human & Environmental Genomics core facility of Rennes (Biosit/OSUR). Patient samples were obtained from the Centre de Ressources Biologiques (CRB)-Santé of Rennes (BB-0033-00056, **www.crbsante-rennes.com**). INSERM-UMR1236's bioinformatics team developed a data analysis pipeline to process RNA-seq data.

References

1. Bass, P.B., Engel, K.B., Greytak, S.R. and Moore, H.M. (2014) A review of preanalytical factors affecting molecular, protein and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: How well do you know your FFPE specimen? Arch. Pathol. Lab Med. **138**, 1520–1530.
2. Morschhauser, F. et al. (2015) Phase 1 study of the oral histone deacetylase inhibitor abexinostat in patients with Hodgkin lymphoma, non-Hodgkin lymphoma, or chronic lymphocytic leukemia. Invest. New Drugs **33**, 423–43.
3. Dobin, A. et al. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England) **29**, 15–21.
4. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics **30**, 923–930.
5. Hull, E.E., Montgomery, M.R. and Leyva, K.J. (2016) HDAC inhibitors as epigenetic regulators of the immune system: Impacts on cancer therapy and inflammatory diseases. Biomed. Res. Int. 1–15.
6. Desmots, F. et al. (2019) Pan-HDAC inhibitors restore PRDM1 response to IL21 in CREBBP-mutated follicular lymphoma. Clinical Cancer Research **25**, 735–746.

Ordering Information

Product	Contents	Cat. no.
AllPrep DNA/RNA Mini Kit (50)	For 50 minipreps: AllPrep DNA Spin Columns, RNeasy® Mini Spin Columns, Collection Tubes, RNase-Free Water and Buffers	80204
QIAseq Stranded Total RNA Lib Kit (24)	For strand-specific RNA-seq library preparation of 24 samples: Kits include magnetized reaction cleanup beads and sample index adapters for Illumina NGS platforms	180743
QIAseq Stranded Total RNA Lib Kit (96)	For strand-specific RNA-seq library preparation of 96 samples: Kits include magnetized reaction cleanup beads and sample index adapters for Illumina NGS platforms	180745
QIAseq FastSelect –rRNA HMR Kit (24)	Includes 3 tubes of QIAseq FastSelect reagent for rRNA removal; sufficient for 24 reactions from human, mouse and rat samples	334386
QIAseq FastSelect –rRNA HMR Kit (96)	Includes 1 tube of QIAseq FastSelect reagent for rRNA removal; sufficient for 96 reactions from human, mouse and rat samples	334387
QIAseq FastSelect –rRNA HMR Kit (384)	Includes 4 tubes of QIAseq FastSelect reagent for rRNA removal; sufficient for 384 reactions from human, mouse and rat samples	334388

Learn more at **www.QIAGEN.com**

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual. QIAGEN kit handbooks and user manuals are available at **www.qiagen.com** or can be requested from QIAGEN Technical Services or your local distributor.