Gene expression profiling on formalin-fixed, paraffin-embedded (FFPE) canine tumour tissue – How do Lexogen's QuantSeq 3′ and NanoString's nCounter® compare?

Alexander F.H. Haake[1], Alina K. Langenhagen[1], Sandro Andreotti[2], Achim D. Gruber[1]
[1] Institute of Veterinary Pathology & [2] Bioinformatics Solution Center (BSC), Freie Universität Berlin

Introduction

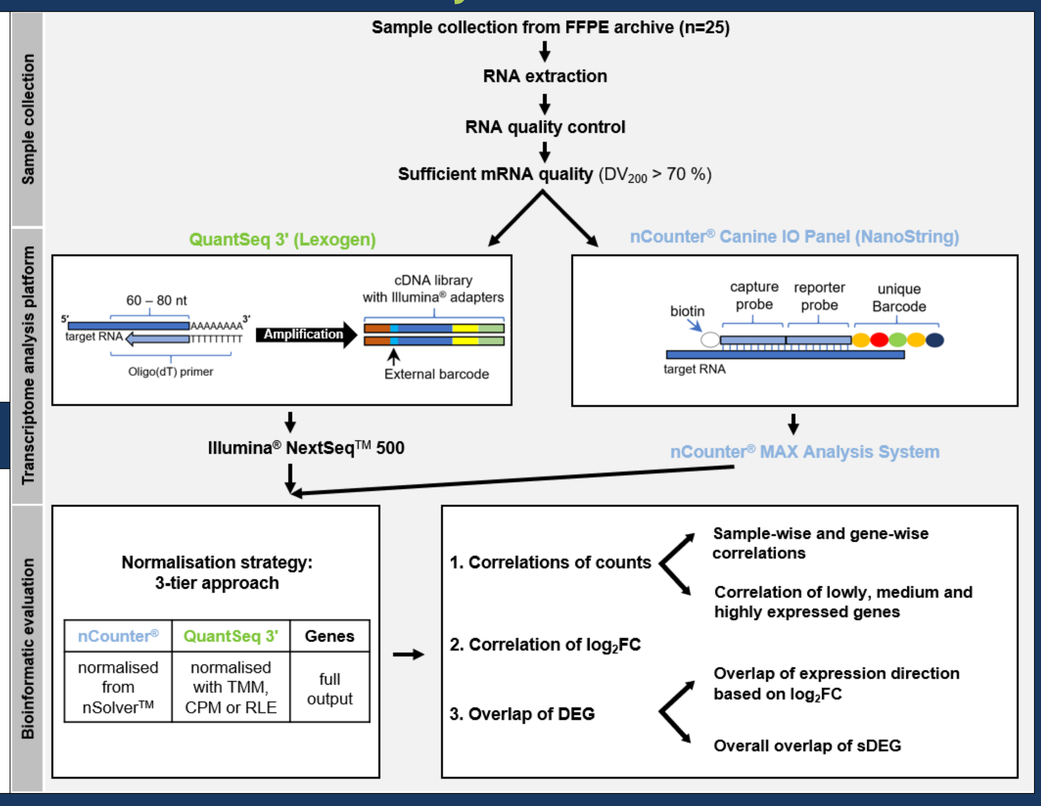
QuantSeq 3′[1] and nCounter®[2] are contemporary high-throughput screening methods for **transcriptome** analyses. **QuantSeq 3′** generates short **whole transcriptome libraries** from the **3′ end** of mRNA molecules which can be sequenced by sequencing by synthesis (SBS). **nCounter®** is an **RNA hybridisation assay** which allows approximately **800 genes** to be analysed via colour-coded molecular barcodes linked to probes. Both technologies enable analyses using short RNA sequences and are thus **well-suited for FFPE material**, where RNA is **fragmented**.

Here, the two methods were compared on multiple levels using two different canine tumours.

Materials & Methods

- **25** archival **FFPE samples** (storage duration: 8 – 0 years)
- RNA extracted from **10 HGA** (hepatoid gland adenomas) & **15 AGASAC** (apocrine gland anal sac adenocarcinomas)
- RNA from same extraction batch analysed with QuantSeq 3′ & nCounter® **Canine IO Panel** with 30 probe Panel Plus
- Normalisation as shown in study workflow & differential gene expression (DGE) analysis
- **Correlation coefficients** calculated for each sample- and gene-wise gene **counts** & **log_2fold change (log_2FC)**
- Overlap of **expression direction** based on log_2FC
- Overlap of **significantly differentially expressed genes (sDEG)**

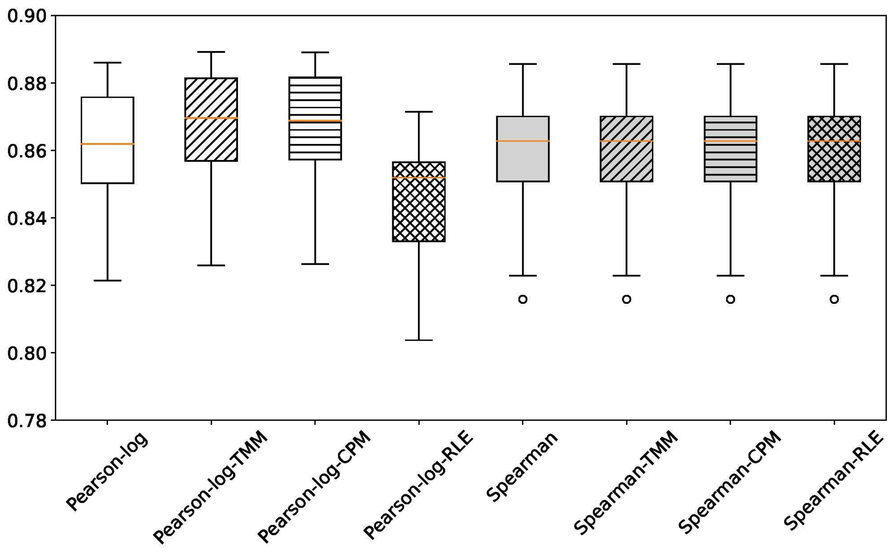
Study workflow



Results

all genes/ sample ⟷ all genes/ sample

Sample-wise count correlations



Of the 830 genes included on the nCounter® Canine IO Panel + Panel Plus, 821 (98.8%) genes were found in the QuantSeq 3′ data.

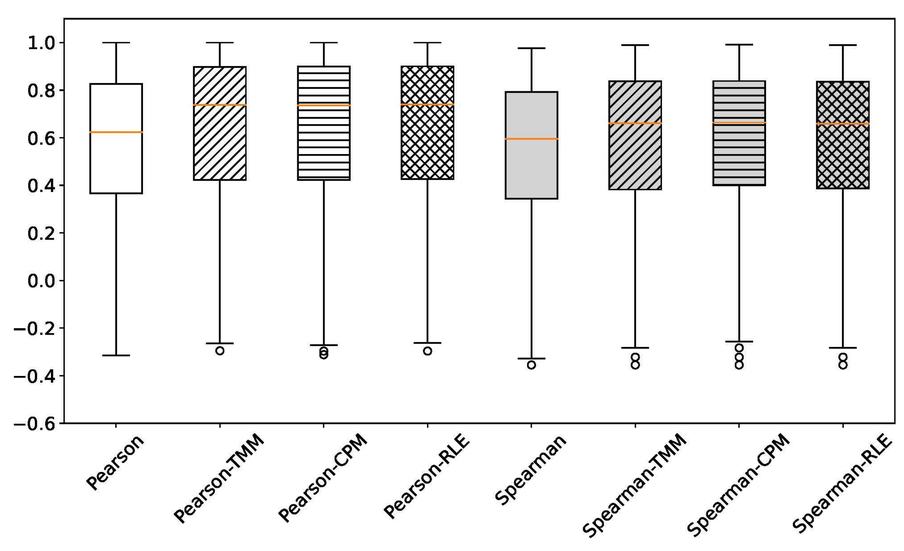
Pearson correlation coefficient
≙ correlation of gene counts
(linear relationship)

Spearman's rank correlation coefficient
≙ correlation of gene rank
(monotonic relationship)

Interpretation of correlation strengths
very strong:
≥ 0.8 to 1
moderately strong:
≥ 0.6 to < 0.8
fair:
≥ 0.3 to < 0.6
poor:
< 0.3

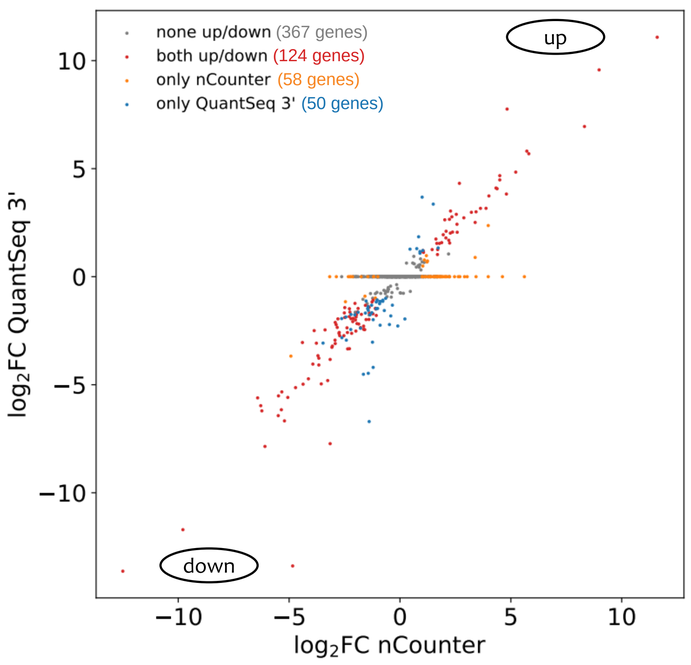
specific gene/ sample ⟷ specific gene/ sample

Gene-wise count correlations

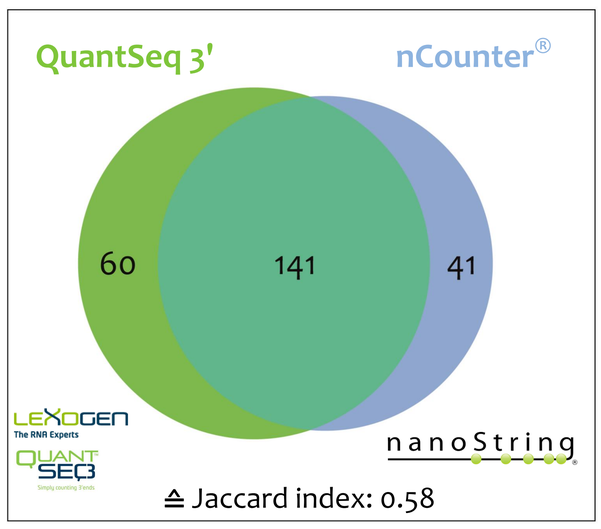


Correlation of counts from QuantSeq 3′ and the nCounter® Canine IO Panel. The box plots show the **sample-wise** and **gene-wise** correlations of counts from both methods. The Pearson-log (for sample-wise) or Pearson (for gene-wise) correlation (background white), indicates the correlation at the count level, while the Spearman correlation (background grey) reflects the values at the gene rank level. The calculations were performed without normalisation (patternless) and with 3 different normalisation methods: Trimmed Means of M-values (**TMM**) - routinely used in the **edgeR** package (obliquely striped), counts per million (**CPM**) (horizontally striped), and relative log expression (**RLE**) - routinely used in the **DESeq2** package (reticulated). Orange horizontal lines represent the medians. Whiskers indicate standard deviation. Circles indicate outliers.

Overall correlation of log_2FC:
Pearson: 0.848 / Spearman: 0.841



Overlap of expression direction based on log_2FC. The scatter plot shows the log_2FC of a given gene from the QuantSeq 3′ data (y-axis) plotted against the log_2FC of the corresponding gene from the nCounter® data (x-axis). The dots are coloured according to the correspondence of expression direction: significantly ($p_{adj} \leq 0.05$ and $log_2FC \leq -1$ or ≥ 1) highly/lowly expressed in both methods, significantly differentially expressed in only one method (nCounter®, QuantSeq 3′), or not significantly differentially expressed in both methods. In total, 599 genes are mapped.



Overlap of significantly differentially expressed genes (sDEG). The Venn diagram depicts 201 sDEG for QuantSeq 3′ and 182 sDEG for nCounter®, with an overlap of 141 sDEG. This corresponds to a Jaccard index of 0.58. This can be assigned to a moderate similarity between the two datasets.

Conclusions

- Both methods generated overall **similar findings** when comparing the overlapping subset of investigated genes.
- The use of 3 different normalisation methods (TMM, CPM, RLE) generated very similar results.
- Despite the two platforms' technological differences, the **strong correlations** on the different levels investigated show that the data can be used for **reciprocal validation of transcriptome results** from **FFPE samples**.

	QuantSeq 3′	nCounter®
Output	whole transcriptome (dependant on reading depth)	800 preselected target genes (+6 – 55 user-defined genes), commercial gene panels or custom panels
Principle of method	strand-specific next-generation sequencing (NGS) libraries generated close to 3′ end of polyadenlyated RNA	mRNA transcripts are directly measured with specific gene probes
Approach suitability	**hypothesis-generating**	**hypothesis-driven** (Canine IO Panel: immuno-oncological landscape)
Required consumables & reagents	QuantSeq 3′ mRNA-Seq Library Kit FDW for Illumina® or Ion Torrent™	nCounter® CodeSets, Primer Pool, Master Kit (incl. cartridges)
Necessary equipment	sequencer (Illumina®, Ion Torrent™)	nCounter® Analysis System (SPRINT, Pro, MAX/FLEX)
Know-how /services required	bioinformatics (raw data processing, alignment to genome, software for differential gene expression analysis)	none (intuitive nSolver™ Analysis Software)

Contact

alexander.haake@fu-berlin.de

References

1. Corley, S. M., Troy, N. M., Bosco, A. & Wilkins, M. R. QuantSeq. 3′ Sequencing combined with Salmon provides a fast, reliable approach for high throughput RNA expression analysis. *Scientific Reports* **9**, 18895 (2019). doi.org:10.1038/s41598-019-55434-x
2. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology* **26**, 317-325 (2008). doi.org:10.1038/nbt1385