PMI RESEARCH & DEVELOPMENT

Evaluation of a novel miRNA-specific normalization method and comparison between Affymetrix and Exiqon platforms

Sewer A[1], Kogel U[2], Gubian S[1], Hengstermann A[2], Peitsch M[1], Hoeng J[1]

[1]Philip Morris International R&D, Philip Morris Products S.A., Neuchâtel, Switzerland
[2]Philip Morris International R&D, Philip Morris Research Laboratories GmbH, Cologne, Germany

Outline

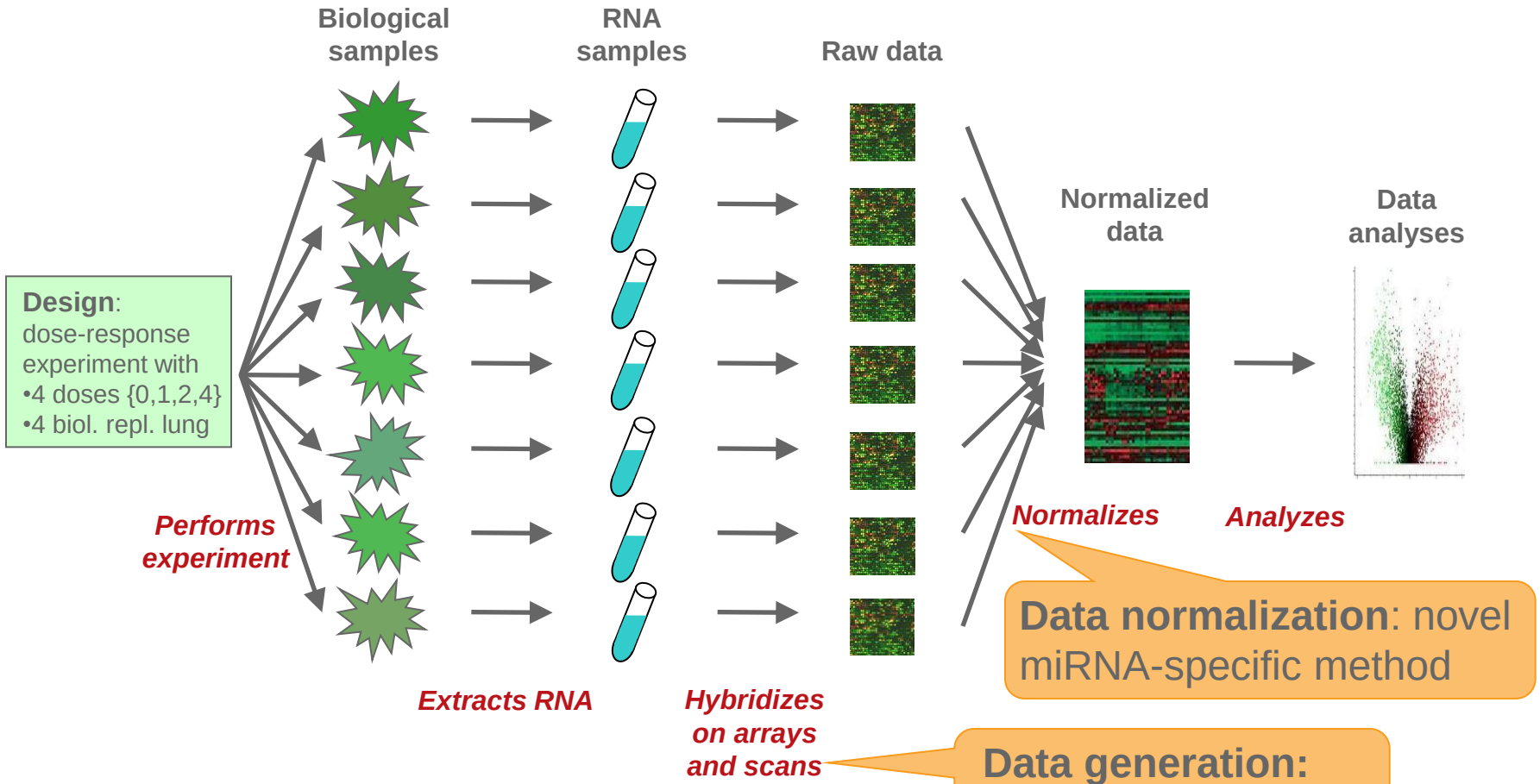
- Introduction
- Raw data comparison (metrics on quality controls)
- Normalized data comparison (metrics based on the biological response)
- Conclusions

Objectives of this study

- ***Data generation***: compare the “new” *Affymetrix® GeneChip® miRNA array* with the “established” *Exiqon miRCURY LNA™* platform
  (by developing objective comparison criteria for raw data quality)

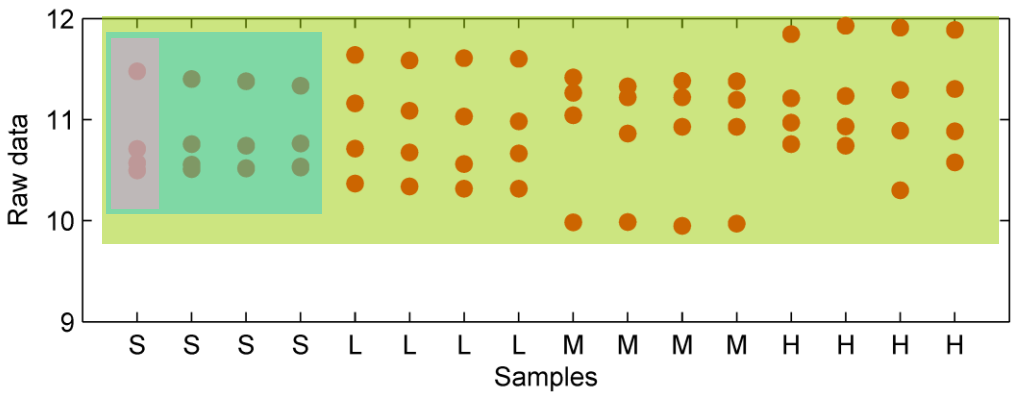
- ***Data pre-processing***: evaluate a novel miRNA-specific normalization method
  (by comparing the biological response of other normalization methods)

Generic workflow for miRNA expression studies



PMI RESEARCH & DEVELOPMENT

Sources of variability in array raw data



Treatment = CS dose

Main contributions to the raw data variance:

within-array variability
→ technical
→ platform

between-array variability
→ technical & biological
→ normalization

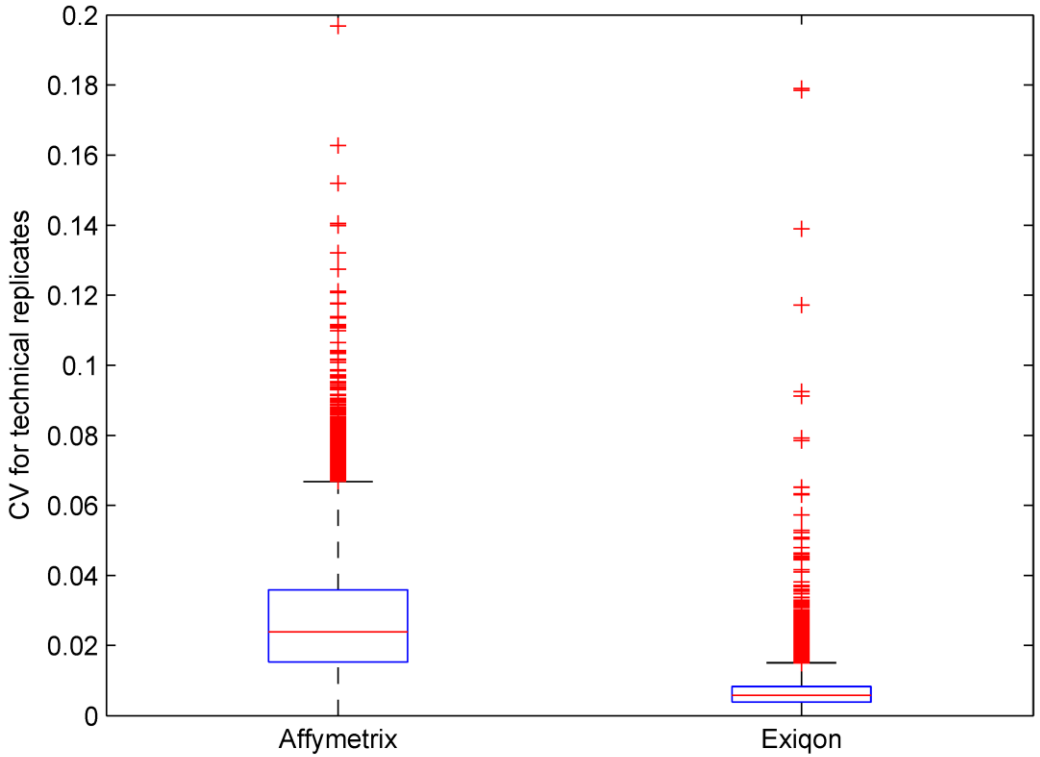
treatment-induced variance
→biological
→analysis
(& normalization)

Outline

- Introduction
- Raw data comparison (metrics on quality controls)
- Normalized data comparison (metrics based on the biological response)
- Conclusions

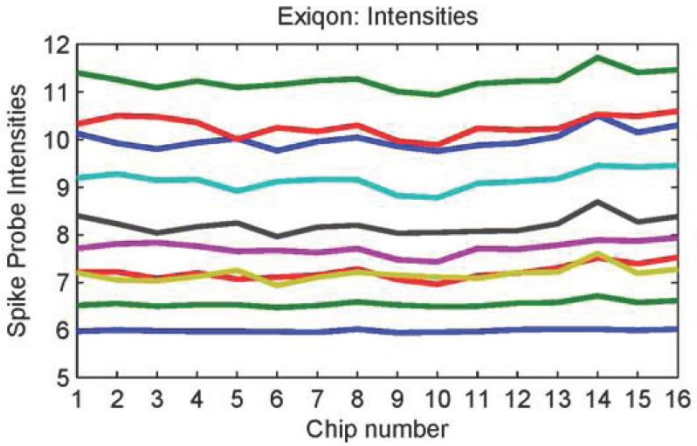
Within-array variability: CV_{within} (1)

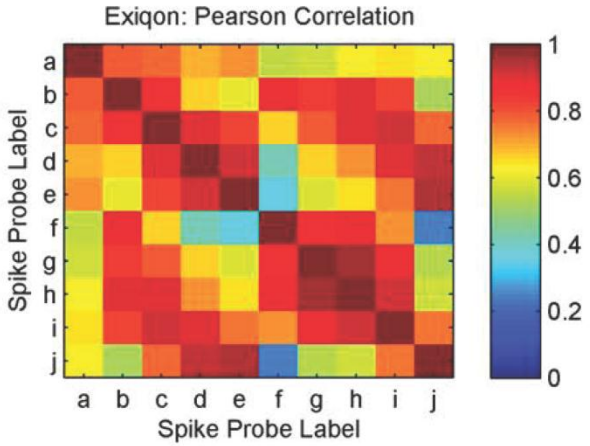
$$CVwithin = \frac{\text{standard deviation}(\text{probes} \in \text{probeset})}{\text{mean}(\text{probes} \in \text{probeset})}$$



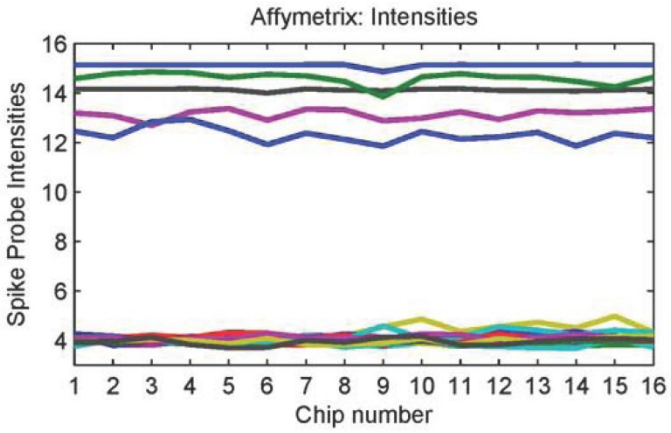
Between-array technical variability: spike-in controls

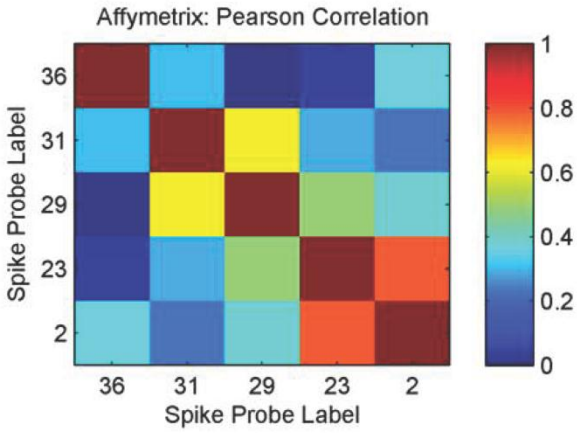
Exiqon





Affymetrix





PMI RESEARCH & DEVELOPMENT

Outline

- Introduction
- Raw data comparison (metrics on quality controls)
- Normalized data comparison (metrics based on the biological response)
- Conclusions

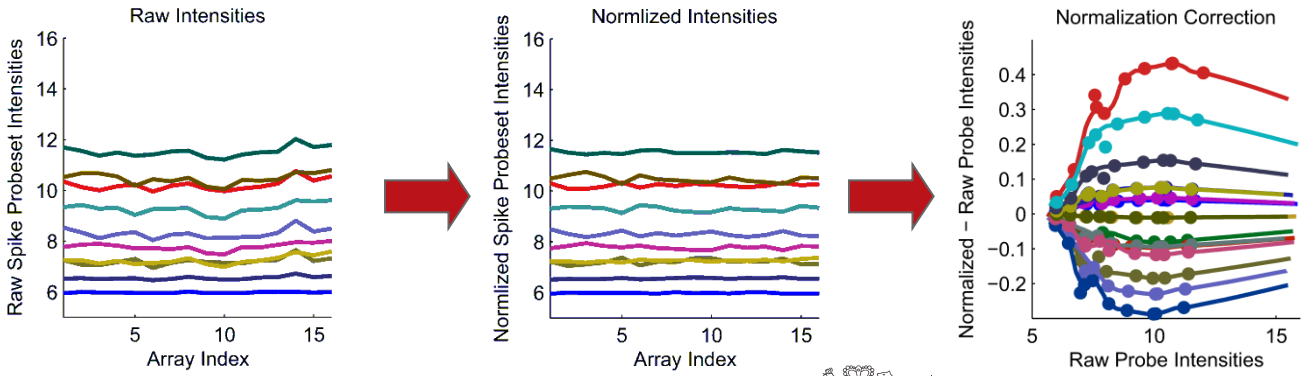
Normalizing miRNA array data

- **Raw data normalization** makes expression data comparable between arrays (by removing the array-specific biases)
- **mRNA arrays**: “house-keeping” genes = reference for intensity calibration (*loess* & *quantile* normalization)
- **miRNA arrays**: **there are no “house-keeping” miRNAs!!!**

  → need other references

  → use the spike-in controls in our normalization method

Rough idea:



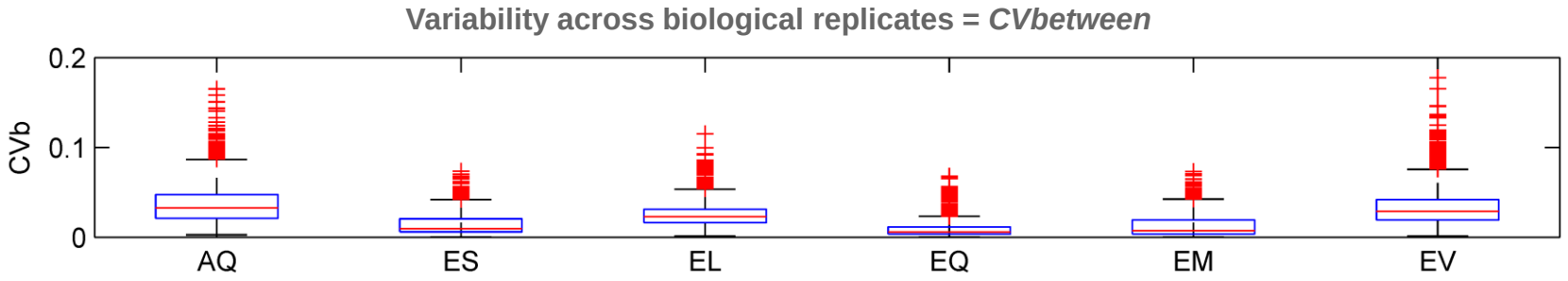
PMI RESEARCH & DEVELOPMENT

Between-array variability: $CV_{between}$

Pipeline = platform + normalization

where

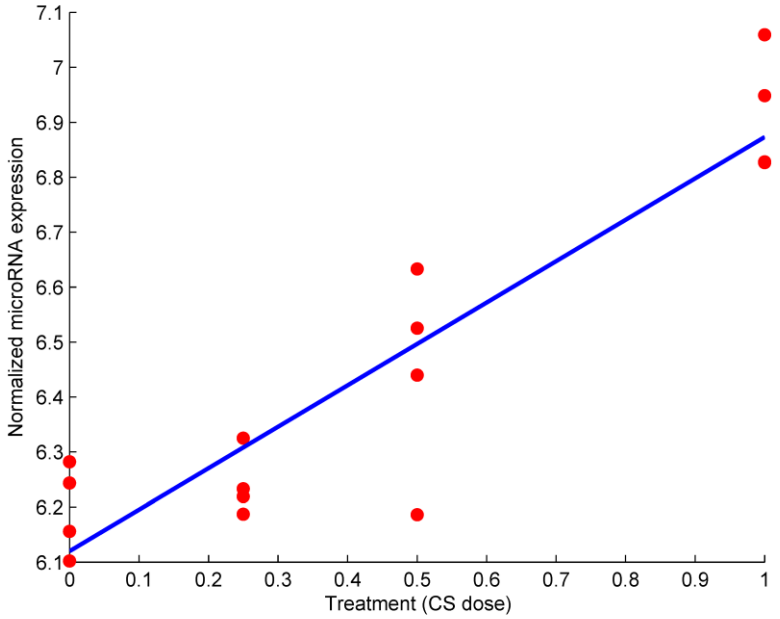
- platform Є {“**A**ffymetrix”,”**E**xiqon”}
- normalization Є {“**S**pike-in based”,”**L**oess”, “**Q**uantile”, “**M**edian”, “**V**SN”}

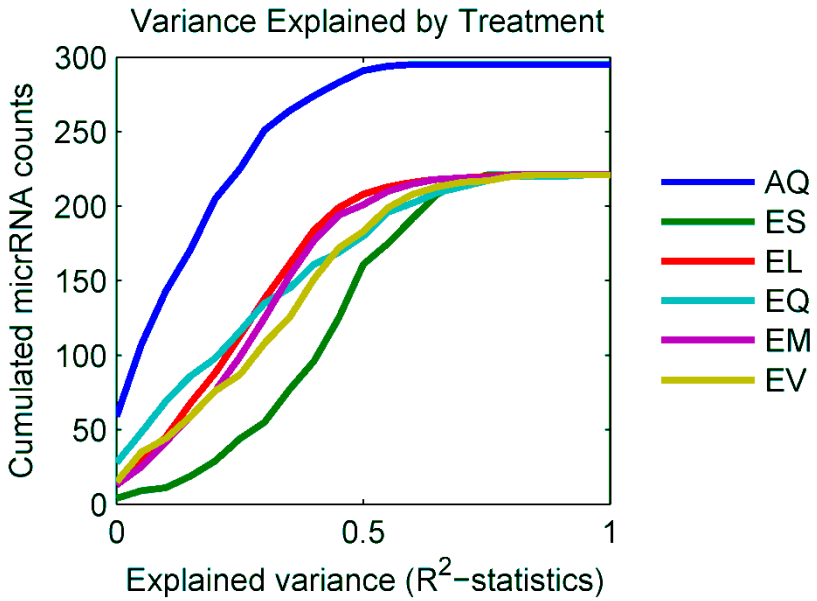


PMI RESEARCH & DEVELOPMENT

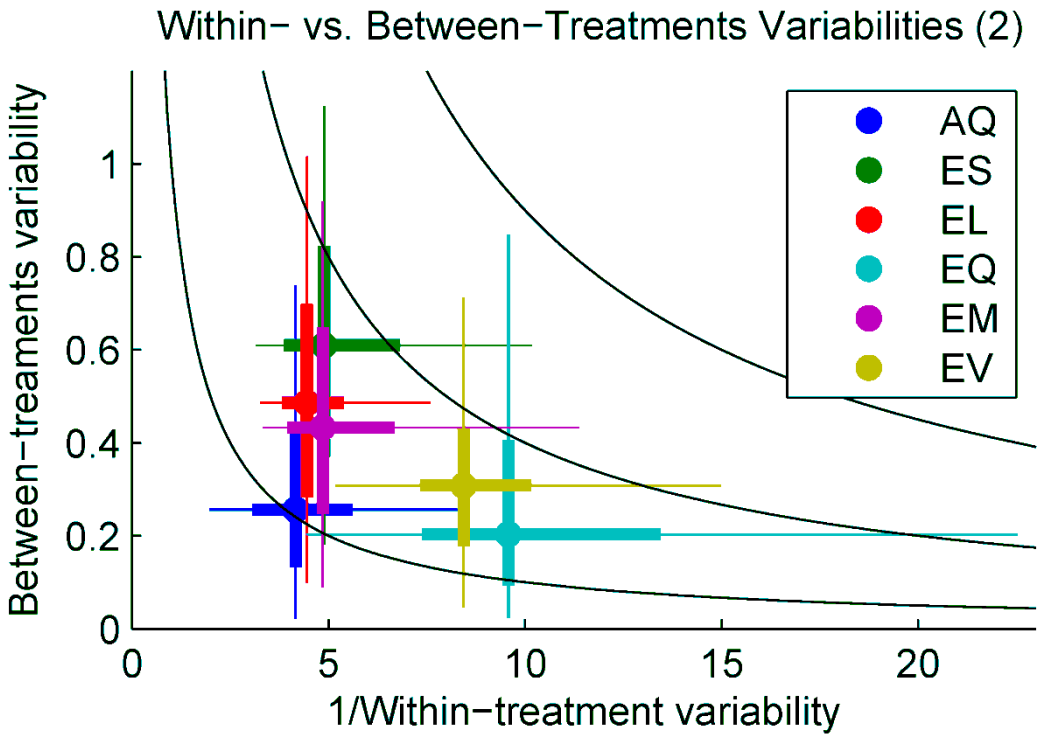
Treatment-induced variability: R^2

Dose-dependent response capture by a linear model (→R^2 statistics)



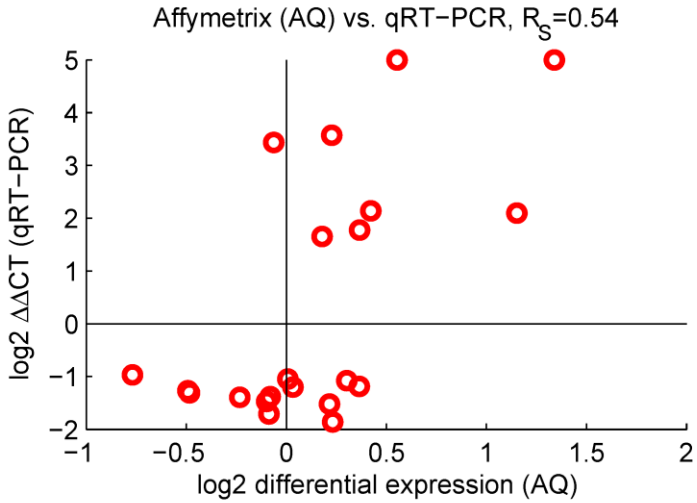


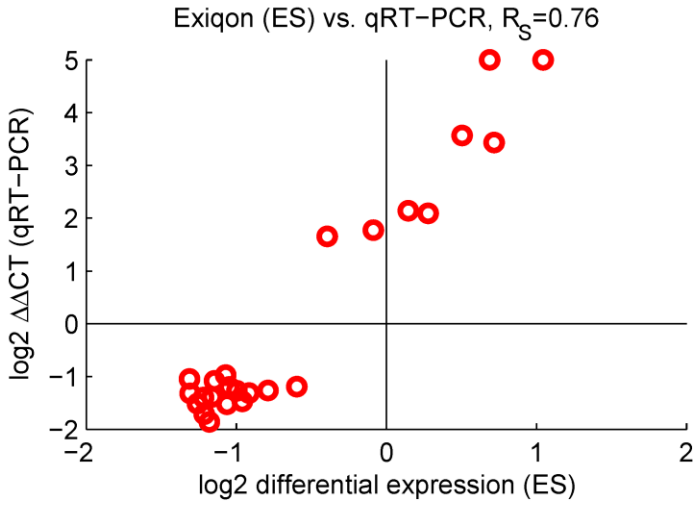
Balance for normalization: low $CV_{between}$ and high R^2

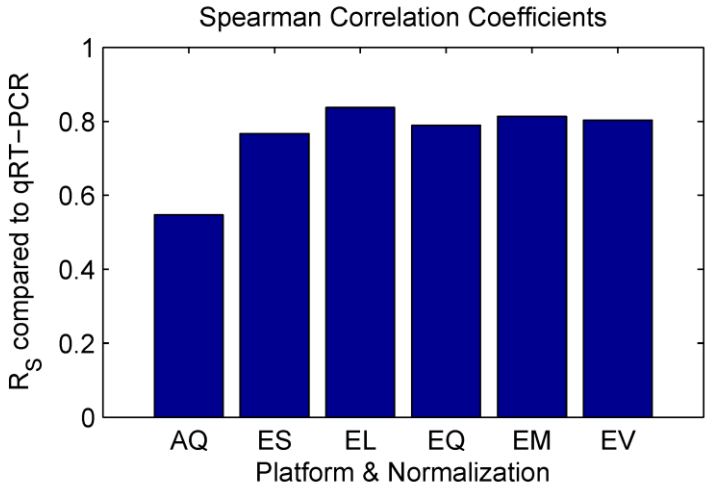


PMI RESEARCH & DEVELOPMENT

qRT-PCR validation on a subset of miRNAs







PMI RESEARCH & DEVELOPMENT

Conclusions

- QC metrics show better results for Exiqon miRCURY LNA™
- The spike-in control-based normalization method performs as good as other methods
- Normalization methods differ in how they reduce the technical and treatment-induced variabilities
- qRT-PCR results show that globally the platform is more important than the normalization

- An *R* package implements the spike-in control-based normalization method (send requests to Sylvain.Gubian@pmintl.com)

Thank you for your attention!