Comparison of endogenous control genes for normalisation of relative quantitative real-time PCR data in a study characterising microRNA expression in human breast cancer tissues

P Davoren, R McNeill, AJ Lowery, N Miller, MJ Kerin

Department of Surgery

University College Hospital Galway.
Breast Cancer

More than 1700 new cases diagnosed in Ireland every year

20% of all malignant cancers in Irish women

Prognostic factors
- Nodal status
- Tumour size
- Histological grade
- Hormone receptor status


Irish cancer society
Gene expression profiling of Breast Cancer

- Patient-tailored treatment

- mRNA expression signatures
  - Oncotype DX™ 21-gene assay (1)

- microRNA expression signatures
  - may be more powerful than mRNA in classifying tumours (2)
  - may aid in prognostication and predicted response to therapy

(2) Lu et al, *Nature*, 2005
microRNAs

- Negatively regulate gene expression at the transcriptome level
- microRNA → multiple targets
- Mutated or abberantly expressed in cancer
- Novel class of tumour suppressor genes and oncogenes

Esquela-Kerscher and Slack, Nature Reviews, 2006
microRNA profiling

- Hybridisation-based technologies
  - Northern-blotting
  - Microarray
  - Bead-based hybridisation

- qPCR
  - Increased specificity
  - Low template requirement
  - Large dynamic range
  - Multiplexing capabilities
Normalization of qPCR data

Raw data

Correct for

Technical variation in clinical samples:

Detect

Gene-specific/biological variation

Commonly normalized using a constitutively-expressed Endogenous Control (EC) gene
Considerations

- No established endogenous controls for microRNA profiling studies
  - Let-7a, 18S rRNA, GMN (1)
  - RNU19, RNU66 (2)

- Template = total RNA or small RNA-enriched RNA

- Ideally
  - Same level of expression
  - Same assay chemistry

(1) Bandres et al, Molecular Cancer, 2006, 5:29
(2) Chen et al, Cancer Res, 2007; 67(3):976-83
## Candidate EC genes

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Accession number</th>
<th>Function of RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>let-7a</td>
<td>MI0000060 *</td>
<td>downregulated in lung cancer, targets RAS oncogenes</td>
</tr>
<tr>
<td>miR-16</td>
<td>MI0000070 *</td>
<td>Negatively regulates the anti-apoptotic protein BCL2 in CLL patients, recommended by AB</td>
</tr>
<tr>
<td>miR-10b</td>
<td>MI0000267 *</td>
<td>expressed in human breast tissue</td>
</tr>
<tr>
<td>miR-21</td>
<td>MI0000077 *</td>
<td>Anti-apoptotic factor; upregulated in various cancers</td>
</tr>
<tr>
<td>miR-26b</td>
<td>MI0000084 *</td>
<td>expressed in human breast tissue, recommended by AB</td>
</tr>
<tr>
<td>RNU48</td>
<td>NR_002745 *</td>
<td>predicted to guide the 2’O-ribose methylation of 28S rRNA</td>
</tr>
<tr>
<td>Z30</td>
<td>AJ007733 *</td>
<td>predicted to guide the methylation of the Am47 residue in U6 snRNA</td>
</tr>
</tbody>
</table>

*miRBase  
**NCBI**
Samples in study

- Fresh-frozen Breast Tissue samples (n=36)
  - Normal (n=5)
  - Primary Breast Tumours (n=26)
    - Metastases Free on presentation
      - Metastases-Free (n=13)
      - Bone metastases (n=7)
      - Bone & Visceral metastases (n=6)
  - Benign (n=5)

- 5-year follow-up
Sample collection

RNA isolation (small RNA-enriched)

RNA analysis

Reverse Transcription

Real-time PCR
qPCR

- Amplification efficiency determined for each gene
  \[ E = (10^{(1/-\text{slope})} - 1) \times 100 \]

- Samples represented in triplicate

- No-template-control for each gene

- Inter-run calibrators for each gene
  - pool of 5 normal breast tissues
  - Standard deviation (0.2-0.3 Ct)
Relative quantification

- **qBase v1.3.5** (Jan Hellemans & Jo Vandesompele)

- **Relative quantification model**
  - Allows use of multiple EC genes
  - Adjusts for efficiency of GOI & EC

- **Output**
  - Relative quantification of GOI using user-defined EC genes
  - Produces data file of RQ values for further analysis of EC genes
**geNorm** (Vandesompele et al, Genome Biology, 2002)

- **Gene stability measure $M$**
  - average pairwise variation of a gene with all other genes

- **Pairwise variation $V$**

<table>
<thead>
<tr>
<th>sample 1</th>
<th>sample 2</th>
<th>sample n</th>
</tr>
</thead>
<tbody>
<tr>
<td>gene A</td>
<td>gene A</td>
<td>gene A</td>
</tr>
<tr>
<td>a1</td>
<td>a2</td>
<td>an</td>
</tr>
<tr>
<td>b1</td>
<td>b2</td>
<td>bn</td>
</tr>
</tbody>
</table>

  $\log_2(a_1/b_1)$, $\log_2(b_1/b_2)$, $\log_2(a_n/b_n) 
  $ Standard deviation = $ V$

- **Output**
  - Ranks genes according to their stability
  - Normalisation factor based on geometric average of input EC genes
NormFinder (Andersen et al, Cancer Research, 2004)

- estimation of both the intra- and the inter-group expression variation

- Combines these to give a stability value for each gene.

**Requirement**
- No prior expectation of expression difference between groups

**Output**
- Ranks genes according to their stability
- Recommends best combination of two genes
Further investigation

- No significant variability in RQ of target gene when additional EC gene is used for normalisation

- Significant variability in RQ of target gene depending on single EC gene used

ANOVA $p < 0.001$
Conclusions

- EC genes need to be validated in context of study
- let-7a is most stably expressed in this cohort
  - Minimum of three genes recommended for reliable normalisation (1)
- sn(o)RNAs may not be suitable for every application
- Microarray $\rightarrow$ additional candidates

(1) Vandesompele et al, Genome Biology, 2002
Work conducted at:
National Breast Cancer Research Institute, University College Hospital, Galway, Ireland

Wed. 28th March, 9:50am, Lecture Hall 14
Evaluation of endogenous control genes for real-time quantitative PCR in breast cancer tissues.
McNeill R.E., Miller, N. and Kerin, M.J.