Author's response to reviews

Title: A new molecular breast cancer subclass defined from a large scale real-time quantitative RT-PCR study

Authors:

Maia Chanrion (mboulfroy@valdorel.fnclcc.fr)
Helene Fontaine (hfontaine@valdorel.fnclcc.fr)
Carmen Rodriguez (crodriguez@valdorel.fnclcc.fr)
Vincent Negre (vnegre@valdorel.fnclcc.fr)
Frederic Bibeau (fbibeau@valdorel.fnclcc.fr)
Charles Theillet (theillet@valdorel.fnclcc.fr)
Alain Henaut (alainhenaut@yahoo.fr)
Jean-Marie Darbon (jmdarbon@valdorel.fnclcc.fr)

Version: 3 Date: 22 December 2006

Author's response to reviews: see over
Please find enclosed our revised manuscript entitled “A large scale real-time quantitative RT-PCR study defines a new molecular breast cancer subclass” by Chanrion et al., as well as answers to the referee comments.

Ethics and consent: according to the French law, using tumor tissues from a tumor library do not need approval from ethics committee. We have added a sentence in the Methods section of the manuscript (page 4, lane 8-9) to claim that informed consent has been obtained from patients.

Language: as suggested, our manuscript has been corrected by a professional copyediting service.

Answers to referee #1

1- Asked for some precision on how tumor subgroups discussed were defined. Referee #2 shared this point.

The 12 subgroups were selected based on the hierarchical clustering analysis of gene expression data (Figure 1). Subgroups were defined by crossing both tumor and gene clusters. To this mean, we did not solely rely on the analysis of dendrograms and took also the heat map patterns into account. However, in doing this, we considered that the subgroups were just a starting ground for further analyses based on expression averaging (Figure S1) and Chi² computation (Table 3). Expression averaging has been successfully performed by others (Sorlie et al.) and we partly based our procedure on this example. Chi² analysis was performed in a second time to assess the robustness of the previously determined subgroups. In our mind, the Chi² test has a number of positive points. Notwithstanding that it is a simple and robust statistical test, the fact it uses discrete variables renders data normality dispensable. A further positive argument is the relatively stringent threshold value we applied (p value ≤ 10⁻⁴). This allowed us to retain 9 of 12 predetermined subgroups. We believe that the existence of 6 of 9 subgroups on an external dataset is a strong indication of the relevance of our approach.

2- judged that instead of using our entire dataset as a training set and use external data for validation, we should have used the 2/3, 1/3 split method on our dataset.
We agree that the split method is commonly used, however, we did not apply it because of it is known to lead to overtraining and resubstitution estimate. Furthermore, like the Chi2 analysis, it is an internal validation. Therefore, we chose to validate our classification on an external validation set using expression data from the Sorlie’s and van’t Veer’s studies. External validation is largely admitted as a robust way to assess the strength of a classification.

3- proposed that, instead of Chi2 analysis, we select genes on the basis of ROC curves and validate them by calculating centroids and classifying the genes on the basis of Pearson and Spearman correlation.

We agree that ROC curves are well adapted for supervised analysis and definition of a classifier. However, we have some doubts on the congruence of ROC with our analysis, since we are not trying to determine and optimize a molecular signature. Our study is indeed a unsupervised analysis, whose goal is to discover new molecular subtypes exhibiting particular bioclinical features. The suggestion of calculating centroids and Pearson-Spearman correlation is interesting indeed, it’s why we used this approach to check the reliability of our clustering division (page 8, lanes 18-23 and Figure S1). To go on the analysis, we decided to use the Chi2 test for the reasons indicated above (point 1).

4- questioned the selection of our gene set.

We agree that selecting an intrinsic gene set based on a greater variability between samples of different patients than between samples of the same patient (Sorlie et al) is a very relevant way to classify tumors. In this work, we attempted an alternative approach in order to eventually define new subtypes. In order to address the point of the intrinsic gene set, we attempted to classify the Sorlie tumor set by means of the 15 genes contained in both Sorlie ‘s signature and our gene set. Doing so, we were unable to identify the subtypes defined using the 500 intrinsic gene set. We have clarified this point, page 13, lanes 7-10.

Minor Revisions

1- Correlation van’t Veer tumors and ours in subgroup 8 (now subgroup 7)…
Whatever the cell-of-origin subtypes of the Sorlie’s and van’t Veer tumors, all tumors attributed to a given subgroup exhibited similar correlation to the signature specifying this subgroup. This correlation was found to be slightly lower than the one found for tumours of the training set (mean value : 70%).

2- Could group 3 and 4 correspond to luminal A and group 1 and 2 to luminal B?

This would be a tempting hypothesis, however, we think it unlikely that luminal A and B distribute so clearly in groups 3/4 and groups 1-2 respectively. There are several reasons to this : (i) recurrences are not more frequent in groups 1-2 than in 3/4 (table 4), (ii) in the van’t Veer cohort, luminal B tumors are predominantly found in groups 3/4 rather than in groups 1-2 (Figure 2), (iii) luminal A tumors are distributed in groups 1 through 4 (as noted in the text, page 10 lane 14). In
conclusion, we do not think that our gene set allows a correct discrimination between luminal A and B tumors.

3- Identification of the samples provided by Dr Katsaros

Dr Katsaros provided 14 tumors, which clustered in all subgroups with a majority in subgroup 3/4 (not surprising as those tumors were ER+). All the patients were treated in Montpellier except those provided by Dr Katsaros, who were treated in Torino along identical standards.

Answers to referee #2

1- was recurrence of the different subtypes biased by the treatment method?

We agree with the referee that the recurrence rate may be partially biased by the therapeutical regimen as well as other parameters including age. This limit is common in a number of retrospective studies, including those of Sorlie et al. and van’t Veer et al. There is a comment in the discussion that addresses this point (page 13, the sentence before the last).

2- criteria used to select 97 and 12 samples in the van’t Veer and Sorlie datasets.

Samples from the van’Veer and Sorlie data were selected on the basis of the availability of data expression concerning our 47-gene set without any other consideration. This point has been made clear in the text (page 7, lane 12).

3- RT-PCR internal controls and primer specificity.

We agree that it is best to use a pool of at least 3 house-keeping genes as internal control. However, our study was initiated before these normalization methods were published and we started using 2 genes: 36B4 and 28S. As the 2 genes exhibited similar behaviour, we dropped 36B4. The specificity of each primer couple was demonstrated by dissociation curve analysis as specified now, page 5, lane 12.

4- we seemed surprised that our gene set was unable to discriminate luminal A and B tumors.

Surprised not really, however, we had to address this question as most genes in our 47 gene set corresponded to ER targets or regulators and 15 of the 47 were shared with the 500-gene set used by Sorlie et al. As stated in point 4 of our reply to referee #1, 15 gene is too small a number to correctly discriminate luminal A/luminal B tumors.

5- benefits of Q-RTPCR vs. microarrays.

The discussion has been modified to take into account this point (page 11, lane 29).

6- About unsupervised analysis based on selected gene set.
It is indeed possible to perform a unsupervised clustering analysis on the basis of selected genes expression. We have clarified the rationale of our gene selection, page 7, paragraph “gene set selection”.

**Minor Revisions**

1- page 9, lane 19 (now lane 2)...level of over-expression.

For each of these genes, we have compared the mean expression value in tumours from subgroups 1 to 4 and tumours from subgroups 5 to 12 and we have performed a t-test:

<table>
<thead>
<tr>
<th>Gene</th>
<th>mean expression in subgroups 1-4 (n=86)</th>
<th>mean expression in subgroups 5-12 (n=120)</th>
<th>t-test</th>
<th>Fold-change</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGF1R</td>
<td>6.0</td>
<td>2.4</td>
<td>p&lt;0.0001</td>
<td>2.5</td>
</tr>
<tr>
<td>KRT19</td>
<td>30.0</td>
<td>17.2</td>
<td>p&lt;0.0001</td>
<td>1.7</td>
</tr>
<tr>
<td>LIV1</td>
<td>36.0</td>
<td>10.5</td>
<td>p&lt;0.0001</td>
<td>3.4</td>
</tr>
<tr>
<td>CCND1</td>
<td>9.1</td>
<td>3.9</td>
<td>p&lt;0.0001</td>
<td>2.3</td>
</tr>
<tr>
<td>ESR1</td>
<td>3.2</td>
<td>1.0</td>
<td>p&lt;0.0001</td>
<td>3.2</td>
</tr>
<tr>
<td>ERBB4</td>
<td>0.3</td>
<td>0.2</td>
<td>p=0.007</td>
<td>1.8</td>
</tr>
<tr>
<td>GATA3</td>
<td>13.3</td>
<td>6.7</td>
<td>p&lt;0.0001</td>
<td>2.0</td>
</tr>
<tr>
<td>IGFBP4</td>
<td>8.0</td>
<td>6.1</td>
<td>p=0.0018</td>
<td>1.3</td>
</tr>
<tr>
<td>PR</td>
<td>0.5</td>
<td>0.2</td>
<td>p=0.0009</td>
<td>2.1</td>
</tr>
<tr>
<td>TFF1/pS2</td>
<td>21.5</td>
<td>6.2</td>
<td>p=0.0013</td>
<td>3.5</td>
</tr>
</tbody>
</table>

2- We have modified the Figure 3 according the referee’s comments.

**Answers to referee #3**

1- We have modified the text to clarify the criteria for selection of our gene set as suggested by the referee.

   …paragraph “gene set selection”, page 7

2- We have modified the discussion to clarify the interest of our findings.