Supporting Information File S1

Statistical analysis of results presented in Fig. 1 B:

In order to confirm that Assay 1 estimates the male/female ratio correctly and that Assay 2 and Assay 3 differ significantly from Assay 1, we tested if the regression curve of Assay 1 is coincident (has the same slope and intercept) with the regression curve of Assay 2 or Assay 3. In a first step the variances $\sigma^2$ of the three different data sets are tested for equality by computing the residual variances using the following formula:

$$ S^2 = \frac{1}{n-2} \sum_{i=1}^{n} \left( y_i - \hat{y}_i \right)^2 $$

with $\hat{y}_i = empirical\ regression\ line\ \hat{a} + \hat{b} \cdot x_i$

Table with data sets of the three assays and the calculated residual variances $S^2$

(see also Fig. 1 B)

<table>
<thead>
<tr>
<th>percentage of male DNA</th>
<th>Assay 1</th>
<th>Assay 2</th>
<th>Assay 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dys14/18S [%]</td>
<td>SRY/c-myc [%]</td>
<td>SRY/18S [%]</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>9.488</td>
<td>16.764</td>
<td>22.517</td>
</tr>
<tr>
<td>10</td>
<td>14.523</td>
<td>17.962</td>
<td>47.437</td>
</tr>
<tr>
<td>20</td>
<td>25.259</td>
<td>84.667</td>
<td>77.489</td>
</tr>
<tr>
<td>30</td>
<td>36.821</td>
<td>92.005</td>
<td>107.761</td>
</tr>
<tr>
<td>50</td>
<td>52.284</td>
<td>141.951</td>
<td>310.609</td>
</tr>
<tr>
<td>$\hat{S}^2$</td>
<td>24.390</td>
<td>212.520</td>
<td>1654.429</td>
</tr>
</tbody>
</table>

For verifying the null hypothesis that the residual variance of the Assay 1 data set is equal to the residual variance of the Assay 2 or Assay 3 data sets ($H_0: \sigma_1^2 = \sigma_2^2$ or $H_0: \sigma_1^2 = \sigma_3^2$), the two-tailed $F$-Test is used. The test can be carried out by dividing the larger residual variance $\hat{S}^2_1$ by the smaller residual variance $\hat{S}^2_2$:

$$ f = \frac{\hat{S}^2_1}{\hat{S}^2_2} \quad \text{with} \quad (S_1 \geq S_2) $$
Based on the 1-\(\alpha/2\) percentage point in the F-distribution table with \(\alpha = 0.05\) (i.e., corresponding to a confidence level of 95%) and \(m_1 = n_1 - 2\) and \(m_2 = n_2 - 2\) degrees of freedom, the null hypothesis is rejected for \(f \geq F_{m_1,m_2;1-\alpha/2}\). In this case: \(F_{m_1,m_2;1-\alpha/2} = 9.6\). The null hypothesis is accepted when \(f \leq F_{m_1,m_2;1-\alpha/2}\). The residual variances of the data set from Assay 1 and those from Assay 3 differ significantly because \(f \geq F_{m_1,m_2;1-\alpha/2} = 67.81 \geq 9.6\), whereas the null hypothesis for the comparison of Assay 1 and Assay 2 can be accepted: \(f \leq F_{m_1,m_2;1-\alpha/2} = 8.71 \leq 9.6\).

As Assay 1 and Assay 2 do not differ significantly in terms of their respective residual variances, we tested if the regression coefficients (i.e., slopes) \(\hat{b}_1\) and \(\hat{b}_2\) of these 2 assays are significantly different. To verify the null hypothesis that \(\hat{b}_1\) and \(\hat{b}_2\) are equal the following formulas can be used:

\[
I_r^{(b)} = \frac{\hat{b}_{(1)} - \hat{b}_{(2)}}{S \times \sqrt{\frac{1}{(n_1 - 1) \times s^2_{x_1}} + \frac{1}{(n_2 - 1) \times s^2_{x_2}}}}
\]

with \(S = \sqrt{\frac{s^2_1 + (n_2 - 2) \times s^2_2}{n_1 + n_2 - 4}}\)

Table with calculated values for variances \(S^2\), \(S^*\), \(s^2_x\) and slope \(b\)

<table>
<thead>
<tr>
<th>values</th>
<th>Assay 1</th>
<th>Assay 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{s}^2)</td>
<td>24.390</td>
<td>212.520</td>
</tr>
<tr>
<td>(\hat{b})</td>
<td>1.0254</td>
<td>2.9299</td>
</tr>
<tr>
<td>(\hat{s}^*)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(S)</td>
<td>10.88</td>
<td></td>
</tr>
<tr>
<td>(s^2_x)</td>
<td>344.167</td>
<td>344.167</td>
</tr>
</tbody>
</table>

Referring to the 1-\(\alpha/2\) percentage point (95% level of confidence), based on \(m_1 = n_1 - 2\) and \(m_2 = n_2 - 2\) degrees of freedom in the tabulated T distribution, the results (slope) of Assay 2 differ significantly from the results (slope) of Assay 1. The t-value for the comparison of
Assay 1 and Assay 2 is about $t_r^{(b)} = 5.13$ whereas the $t_{m;1-\alpha/2}$ determined for $m = n_1 + n_2 - 4$ degrees of freedom and a 95% confidence level (1-α/2 percentage point) is about 2.3.

Therefore is $t_r^{(b)} \geq t_{m;1-\alpha/2}$ and the null hypothesis is rejected.

Based on this analysis, it is determined that the results obtained by Assay 2 and Assay 3 differ significantly from the results obtained by Assay 1.