Equations for calculating the expected gametic and allelic frequencies under Model 1

The expected frequencies $F_c$ of the different 1n-2c configurations can be derived from the recombination frequencies between the three $S_1$ genes and between the $S_1$ complex locus and the centromere (see schematic representation 1 below), and are given in Table S3-1.

The expected frequencies of each gamete that belong to a specific 1n-2c configuration are equal to $F_c/4$, where $F_c$ is the expected frequency of the specific 1n-2c configuration.

Let $D$ be the sum of the frequencies of the viable gametes. Under model 1, $D$ simplifies to:

$$D = \left(1 - 2r_1\right)\left(\frac{r_3}{2} + 1\right) + 2r_1\left[\frac{1}{2}r_2 + \frac{r_1}{2} - r_1r_2\right]$$

The expected frequencies of viable gametes are obtained by dividing their corresponding expected frequency by $D$.

Thus, the final frequency of gametes bearing the $S_1^+$ allele, $f\left(S_1^+\right)$, is obtained by summing the individual frequencies of viable gamete that bear this allele. Under Model 1, this simplifies to:

$$f\left(S_1^+\right) = \left[\frac{r_3}{2} + r_1\left(r_1 - r_2 - 2r_1r_2\right)\right]/D$$

Similar equations can be derived for Models 2 to 4, with numerical examples based on observed data given in File “S1 - Genetic models.xlsx”.

Under all models, the estimated for recombination fractions $r_1$ and $r_2$ are naturally given by dividing the observed recombination fractions, $r_{1\text{max}}$ and $r_{2\text{max}}$, by $2D$.

$$\hat{r}_1 = \frac{r_{1\text{max}}}{2D}; \quad \hat{r}_2 = \frac{r_{2\text{max}}}{2D}.$$

In the case of the presence of the additional factor $S_1C$ (see schematic representation 2 below), the effect of $S_1C$ is applied after the initial selection due to the $S_1A$ - $S_1^-$ - $S_1B$ locus action.

As $S_1C$ only affects the $S_1^-$ gametes when a recombination event occurs between $S_1B$ and $S_1C$, the expected frequencies of these gametes are obtained by multiplying them by $(1 - r_3)$, where $r_3$ is the recombination fraction between $S_1B$ and $S_1C$, and dividing them by the relative sum of all gamete frequencies, $1 - r_3f\left(S_1^+\right)$. The frequencies of other viable gametes are obtained by dividing them by the relative sum of all gamete frequencies, $1 - r_3f\left(S_1^+\right)$. 
TABLE S3-1

Expected frequencies $F_c$ of the different 1n-2c configuration expressed in function of the recombination frequencies between the three $S_i$ genes, and associated gametic frequencies under no selection and Model 1 selection

<table>
<thead>
<tr>
<th>1n-2c configuration</th>
<th>Expected configuration frequency without selection ($F_c$)</th>
<th>Expected gamete frequency without selection ($F_g$)</th>
<th>Survival under Model 1</th>
<th>Expected gamete frequency under selection (Model 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{iA}$ $S_i$ $S_{iB}$</td>
<td>$F_c^1 = \frac{1}{4}$</td>
<td>Viable $F_g/D$</td>
<td>Aborted</td>
<td>$F_g/D$</td>
</tr>
<tr>
<td>$S_{iB}$ $S_{iA}$ $S_i$</td>
<td>$F_c^2 = \frac{1}{4}$</td>
<td>Viable $F_g/D$</td>
<td>Aborted</td>
<td>$F_g/D$</td>
</tr>
<tr>
<td>$S_{iB}$ $S_{iA}$ $S_i$</td>
<td>$F_c^3 = \frac{1}{4}$</td>
<td>Viable $F_g/D$</td>
<td>Aborted</td>
<td>$F_g/D$</td>
</tr>
<tr>
<td>$S_{iB}$ $S_{iA}$ $S_i$</td>
<td>$F_c^4 = \frac{1}{4}$</td>
<td>Aborted 0</td>
<td>Viable $F_g/D$</td>
<td>$F_g/D$</td>
</tr>
<tr>
<td>$S_{iB}$ $S_{iA}$ $S_i$</td>
<td>$F_c^5 = \frac{1}{4}$</td>
<td>Viable $F_g/D$</td>
<td>Aborted 0</td>
<td>$F_g/D$</td>
</tr>
</tbody>
</table>

$F_c^1 = \left(1 - 2r_3\right)\left[1 - 2(r_1 + r_2 - 2r_1r_2)\right]$

$F_c^2 = 2r_3\left(1 - 2r_1\right)(1 - 2r_3)$

$F_c^3 = 2r_3\left(1 - 2r_2\right)(1 - 2r_3)$

$F_c^4 = 2r_3\left(1 - 2r_2\right)(1 - 2r_3)$

$F_c^5 = 4r_1r_2(1 - 2r_3)$
The sum of the frequencies of the viable gametes thus simplifies to:

\[ D' = D' \left(1 - f(S_1)\right) \delta_s \]

while the final frequency of gametes bearing the \( S_1 \) allele converts to:

\[ f(S_1) = \left(1 - \hat{r}_1\right) f(S_1) \delta_s / \left(1 - f(S_1) r_1\right) \]

and the estimates for \( r_1 \) and \( r_2 \) convert to:

\[ \hat{r}_1 = \frac{r_1 \max}{2D'}; \hat{r}_2 = \frac{r_2 \max}{2D'} \]
**Schematic representation 1:** positions of the three $S_1$ genes on rice chromosome 6, expressed as recombination fractions between each other and between the $S_1$ locus and the centromere.

![Schematic representation 1](image)

**Schematic representation 2:** positions of the three $S_1$ genes and the $S_1C$ gene on rice chromosome 6, expressed as recombination fractions between each other and between the $S_1$ locus and the centromere.

![Schematic representation 2](image)