Where do all the gene copies go?

Resolving a complex GMO with Oxford Nanopore sequencing technology

14 March 2019
Wouter de Bonte

Final symposium ‘Porelab’
Proteins
Chemicals

DNA molecules leading to proteins through a chemical process.
# Illumina sequencing vs Southern blot

<table>
<thead>
<tr>
<th>strain</th>
<th>copy number</th>
<th>integration sites</th>
<th>coverage Illumina data</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5-6</td>
<td>1</td>
<td>7.5</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>2</td>
<td>4.3</td>
</tr>
<tr>
<td>C</td>
<td>6-8</td>
<td>3</td>
<td>7.3</td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>3</td>
<td>14.3</td>
</tr>
<tr>
<td>E</td>
<td>&gt;9</td>
<td>2</td>
<td>9.6</td>
</tr>
<tr>
<td>F</td>
<td>4-6</td>
<td>2-3</td>
<td>14.3</td>
</tr>
</tbody>
</table>
Oxford Nanopore sequencing vs Illumina sequencing
Porelab experiment

Complex *Aspergillus niger* strain, CitB#99 (Hossain et al., 2016)

Producer of itaconic acid

4 different genes inserted (with their selection markers)

Copy number determined by Southern blot
Sequencing of CitB#99 with Oxford Nanopore sequencing

DNA isolation → Sequencing → Trimming and filtering → *de novo* assembly → Analysis of integrated genes
Copy number of inserted genes in CitB#99

<table>
<thead>
<tr>
<th></th>
<th>cadA</th>
<th>mfsA</th>
<th>mttA</th>
<th>citB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern blot</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Nanopore data</td>
<td>9</td>
<td>2</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>
Integration sites in CitB#99

Isolated on contig or at end of contig

7 out of 25 contigs < 25 kb

Example: all 6 copies of *citB* are isolated on the small contigs

Possible solutions:
- Different way of assembling
- Generate longer reads
Considerations

Starting material is crucial

Characterization of GMOs is important for registration (EFSA/FDA)

Oxford Nanopore data needs state-of-the-art bioinformatics analysis to be valuable

Large structural genomic variations (translocations, deletions)