Unique Probe Mapping met behulp van PQ-trees

Leiden, 6 april 2006
**Physical Mapping**

Cut the DNA in each YAC clone and clone into overlapping cosmid clones.

Select a subset of cosmid clones of minimum total length that covers the YAC DNA.

Duplicate the cosmid and then cut the copies randomly. Select and sequence short fragments and then reassemble them into a deduced cosmid string.

**C**: Full DNA

Cut C and clone into overlapping YAC clones.
using a map of the genome
hybridization mapping

Probes

Clones

y
x
z
w
unique probe mapping

clones 1,2,…,6
probes A,B,…,G

Matrix representation

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
reordering of probes

probes

clone

clones contain consecutive probes

order
### Interval Graphs

#### Probe

<table>
<thead>
<tr>
<th>E</th>
<th>B</th>
<th>F</th>
<th>C</th>
<th>A</th>
<th>G</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Clone

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Matrix Representation

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
expressie  code  binaire zoek

trie  syntax  2,3 boom
PQ-trees

representation for permutations

\{ 123, 132, 213, 231, 312, 321 \}  \quad \{ 123, 321 \}
PQ-trees

**datastructure** to represent all possibilities

P permutation

Q linear order

PQ trees represent possible reorderings (permutations of probes)
example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

clones  { A, C, D }   { A, B, C, E }

1
D AC
BE

2

D AC BE

EB CA D

D CA BE

EB AC D

D AC EB

BE CA D

D CA EB

BE AC D
PQ-trees

equivalent representations

{ 123, 132, 213, 231, 312, 321 }
**PQ-tree algorithm**

\[ \text{reduce}(T,S) \]

- **T** PQ tree \( \sim \) set of permutations
- **S** new clone \( \sim \) set of (consecutive) probes

add requirement **S** to tree **T**
- ‘keep **S** together’

- colour leaves in **S**
- apply **transformations**
  - to get consecutive leaves
- apply **replacement rules**
  - to add new restriction to tree

\[ \begin{align*}
  \text{P} & \quad \text{all leaves in **S**} \\
  \text{Q} & \quad \text{segment in **S**}
\end{align*} \]
S = \{A,C,E\}

'root'

no coloured nodes outside this subtree
replacement rules

1. Original
   
2. Replaced
   
3. Original
   
4. Replaced
   
5. Original
   
6. Replaced
   
7. Original
   
8. Replaced
   
9. Original
   

replacement rules (2,3)

root

⇒

(2)

non-root

⇒

(3)
replacement rules (4,5)

root
⇒
(4)

non-root
⇒
(5)
also 7th STOC, 1975.
challenges

• find the right model (simplification)

  ▪ noise, errors,
    too much / not enough data
  ▪ heuristics & AI approach
  ▪ is it data mining?

• interdisciplinary