The Long Road to JATS

Paul Donohoe – Senior XML Developer
Jenny Sherman – XML Developer
Ashwin Mistry - XML Developer

Macmillan Science and Education
In the beginning ...

- over half a million articles (increased to a million by the time the project started)
- 143 journals (increased to 180)
- three in-house DTDs
- five typesetters
- numerous workflows and production systems
- teams based across the world
Mapping process
Mapping to JATS – the value of examining the content

<!ELEMENT artpubdt (#PCDATA)>
<!-- Date of initial online article publication under "continuous publishing" program. -->

Feasibility report mapping

<artpubdt>19991125</artpubdt>

Maps to:

<pub-date pub-type="epreprint" year="1999" month="11" day="25"/>
Mapping to JATS – the value of examining the content

<!ELEMENT artpubdt (#PCDATA)>
<!-- Date of initial online article publication under "continuous publishing" program. -->

Content-based mapping

<artpubdt>19991125</artpubdt>

Maps to:

Nothing. Used for only 10 articles in one supplement.
## Standardization to JATS

### List type mapping from JATS to AJ or NPG

<table>
<thead>
<tr>
<th></th>
<th>AJ</th>
<th>JATS</th>
<th>NPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers</td>
<td>1 2 3</td>
<td>1 2 3</td>
<td>1 2 3</td>
</tr>
<tr>
<td>Numbers with parens</td>
<td>1) 2) 3)</td>
<td>1) 2) 3)</td>
<td>1) 2) 3)</td>
</tr>
<tr>
<td>Letters</td>
<td>a b c</td>
<td>a b c</td>
<td>a b c</td>
</tr>
<tr>
<td>Letters with parens</td>
<td>a) b) c)</td>
<td>a) b) c)</td>
<td>a) b) c)</td>
</tr>
<tr>
<td>Uppercase letters</td>
<td></td>
<td>A B C</td>
<td>A B C</td>
</tr>
<tr>
<td>Roman numerals</td>
<td>i ii iii</td>
<td>i ii iii</td>
<td>i ii iii</td>
</tr>
<tr>
<td>Roman numerals with parens</td>
<td>i) ii) iii)</td>
<td>i) ii) iii)</td>
<td>i) ii) iii)</td>
</tr>
<tr>
<td>Uppercase Roman numerals</td>
<td>I II III</td>
<td>I II III</td>
<td>I II III</td>
</tr>
<tr>
<td>Bullets</td>
<td>• • •</td>
<td>• • •</td>
<td>• • •</td>
</tr>
</tbody>
</table>
Project kickoff
RNA regulons in *Hox* 5′ UTRs confer ribosome specificity to gene regulation

Shifeng Xue, Siqi Tian, Kotaro Fujii, Wipapat Kladwang, Rhiju Das & Maria Barna

**Abstract**

Emerging evidence suggests that the ribosome has a regulatory function in directing how the genome is translated in time and space. However, how this regulation is encoded in the messenger RNA sequence remains largely unknown. Here we uncover unique RNA regulons embedded in homeobox (Hox) 5′ untranslated regions (UTRs) that confer ribosome-mediated control of gene expression. These structured RNA elements, resembling viral internal ribosome entry sites, might be responsible for a cell- and tissue-specific ribosome selection process involved in regulating Hox gene expression.
Internal needs
Existing workflow
Existing to final desired workflow
Suggested interim workflow
Current hybrid workflow
Opportunities
Opportunities

\[ PC_x = \sum_{i=1}^{n} m_i \times l_{x,i} \]
Scrum Development Process
(Re)Write a test

Does the test fail?

Test Fails

(Re)Write production code

Tests Fail

Run all tests

All Tests Succeed

Clean up code

Test Succeeds

Repeat

Test Fails

(Re)Write production code
Unit Testing
Content Hub Architecture
Core Ontology
Plant hormones: On-the-spot reporting

The development of a new ...  

The toolbox of plant hormone reporters.

Expression-based reporters ...
The Big Picture
New Publishing Platform
New Search
Other benefits
Continuous Modelling
<?xml version="1.0" encoding="UTF-8"?>

<meta:corresponding-authors>
  <meta:author corresp="yes">
    <meta:full>Yasushi Ishida</meta:full>
    <meta:email>yaishida2009@yahoo.co.jp</meta:email>
  </meta:author>
</meta:corresponding-authors>
Next Steps

• Single environment publishing
• Data sends for third parties
• Archive conversion
Lessons Learned

• Don’t start from here!
• Do not rely on DTD to DTD mappings
• Agile development methodology for tool development