Coordination and divergence of cell-specific transcription and translation of genes in Arabidopsis root cells

Dhivyaa Rajasundaram
Paris, 19/05/2014
Technology driven opportunities and challenges

Heterogenous data sources

- Annotation data (Standard biological annotation terms such as GO and KEGG)
- Metabolite data (Metabolites and their intermediate products)
- Interaction data (Protein-protein, Protein-DNA, and genetic interactions)
- Imaging studies (Spectroscopic techniques, confocal microscope experiments)
- Nucleic acid sequence data (Genomic DNA sequences and several subgroups of RNA sequences)
- Gene transcript expression data (Microarray and RNA Seq experiments)
- Proteomic studies (Protein level expression, structural analysis of proteins)
Decoding heterogenous data in an integrative way

- Integration of multiple layers of information to understand the functional principles and total dynamics of cellular systems.

How do cells of different types arise from a homogeneous cell pool during development of an organism?

Connections must be made between genes to gene expression and regulation/ interactions to determine how the components work together as a system.

- Challenging task that can be addressed by multivariate statistical approaches.

- Integrative data analysis - more precision, better accuracy, and greater statistical power than any individual dataset would provide.
Flow of cellular information

Is central dogma a global property of cellular information flow?

The Central Dogma of Molecular Biology

Modified from Tobias et al., 2009
Background

- Integrating and comparing transcriptome vs proteome – numerous works in biology.
- Data coverage of proteome is still limited.
- Additional efforts have been made to infer relationship between genes/transcripts and functional elements (metabolites).
Translational control – an interface to mRNA and proteins

- Abundance of each given mRNA in the transcriptome does not necessarily mirror the expression of the encoded proteins within the proteome.

- Only stress dependent coordination studies in yeast and mammalian cells integrating these two system levels.

- Arabidopsis root cells - tractable model for such a study.
Gene expression datasets of *Arabidopsis* root transcriptome and translatatome

- Cell type specific transcriptional signatures.
- Radial dataset – 19 promoters to drive the expression of 14 non-overlapping cell types.
- Translatome atlas of root cell specific populations.
- 10 promoters used to study specific populations from seedling roots.

Brady et al., 2007

Mustroph et al., 2009
List of cell types common to both the transcriptome and the translatome dataset along with the promoters used

<table>
<thead>
<tr>
<th>Cell-type</th>
<th>Transcriptome</th>
<th>Translatome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phloem companion cells</td>
<td><em>SUC2</em>, <em>APL</em></td>
<td><em>SUC2</em>, SULTR2</td>
<td>(Brady et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Lee et al. 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Mustroph et al. 2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Nawy et al. 2005)</td>
</tr>
<tr>
<td>Root vasculature</td>
<td><em>WOL</em></td>
<td><em>WOL</em>, SHR</td>
<td>(Brady et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Mustroph et al. 2009)</td>
</tr>
<tr>
<td>Quiescent centre</td>
<td>AGL42, J0571, <em>SCR</em></td>
<td><em>SCR</em></td>
<td>(Birnbaum et al. 2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Brady et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Mustroph et al. 2009)</td>
</tr>
<tr>
<td>Cortex</td>
<td>CORTEX</td>
<td>CO2, PEP (based on whether it is meristematic, elongation or maturation zones)</td>
<td>(Brady et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Lee et al. 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Mustroph et al. 2009)</td>
</tr>
<tr>
<td>Non-hair cells /Root atrichoblast epidermis</td>
<td><em>GL2</em></td>
<td><em>GL2</em></td>
<td>(Brady et al. 2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Mustroph et al. 2009)</td>
</tr>
</tbody>
</table>
Analysis of similarity of cell-specific mRNA levels on the level of transcriptome and translatome in Arabidopsis root cells

\[
Z = \frac{PCC - \mu}{\sigma}
\]
Coupled total and polysome associated mRNA levels (high PCC and a z-score of $\geq 1.96$)

Uncoupled mRNA levels (low PCC and a z-score of $\leq 1.96$)
Transcription and translation of cell wall-related genes are highly correlated

<table>
<thead>
<tr>
<th>Genes</th>
<th>Identical promoters</th>
<th>Common promoters</th>
</tr>
</thead>
<tbody>
<tr>
<td>COUPLED</td>
<td>851</td>
<td>790</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genes</th>
<th>Identical promoters</th>
<th>Common promoters</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNCOUPLED</td>
<td>494</td>
<td>373</td>
</tr>
</tbody>
</table>

- Regulation of transcription
- Post-translational modification
- Root tissue formation processes
- Cell wall related processes
  - cell wall modification
  - secondary cell wall biogenesis
  - xylan biosynthetic process

- Cell growth
- Root and meristem development
- Protein glycosylation
- Cytoskeletal organization
Identification of altered regulation of gene expression on both system levels
Co-expressed relationships at the transcriptional level are generally not preserved at the translational level.

<table>
<thead>
<tr>
<th>Genes</th>
<th>Identical promoters</th>
<th>Common promoters</th>
</tr>
</thead>
<tbody>
<tr>
<td>COUPLED</td>
<td>71</td>
<td>39</td>
</tr>
<tr>
<td>UNCOUPLED</td>
<td>10051</td>
<td>12681</td>
</tr>
</tbody>
</table>
Genes displaying conserved expression levels (PCC) and co-expression relationships (EC-scores) across root cell-types in translatome and transcriptome.
Gene set enrichment analysis of genes displaying conserved expression and co-expression relationships

- The majority of the identified genes are transcription factors, or contain predicted DNA binding protein domains, e.g., *ATHB-3*, *MYB46*, *VND7* and *WRKY9*.

- Genes are associated with key regulatory roles in roots, either for developmental or response processes.

- *WRKY9* - mediating cell responses to nutrient deprivation
  *MYB46* - developmental program of secondary wall biosynthesis
  *VND7* - transcriptional master switches for plant meta- and protoxylem formation in Arabidopsis
Root cell-type similarity based on transcriptome and translatome

- To elucidate whether transcriptional and/or translational may be conserved across multiple cell types.

- We first identified genes that showed differential expression and translation across the datasets using ANOVA.
Conserved genes of the root translatome and transcriptome (four cell types - identical promoters)
Conserved genes of the root translative and transcriptome (five cell types- common promoters)
Analysis of cell-type specificity of gene expression in transcriptome and translatome of *Arabidopsis* root cells
First prominent motif of the DE genes of the identical promoters

Enrichment analysis using GO-BP terms associated with the genes in the first motif (214 genes) reveals mainly transport processes, as well as responses to sugar stimuli to be overrepresented.
Second prominent motif of the DE genes of the identical promoters

- Genes were enriched for cell wall modification, xylan biosynthetic process and root hair cell differentiation/elongation.
- The GO-BP terms oxidative stress, oxidation-reduction processes and auxin polar transport were also enriched.
Conclusion

- Genome-scale analysis shows difference in variation and conservation across translatome and transcriptome.

- Low overall observed conservation between cell-type specific gene expression on translatome and transcriptome is present (in agreement with the study from Tebaldi et al., 2012).

- Genes showing strongest altered expression conservation (AEC) were obtained by calculating Z-scores.

- Conserved genes of the root translatome and transcriptome exhibiting cell-type specific expression are associated to cell wall biogenesis and other root related processes.
Acknowledgment

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7 2007-2013) under Grant Agreement n°263916.

This presentation reflects the author’s views only. The European Community is not liable for any use that may be made of the information contained herein.

More information about the WallTraC project at www.walltrac-itn.eu.