Understanding Gene Model Maps

Michael N. Lombardo, Faisal Z. Qureshi

Faculty of Science, University of Ontario Institute of Technology
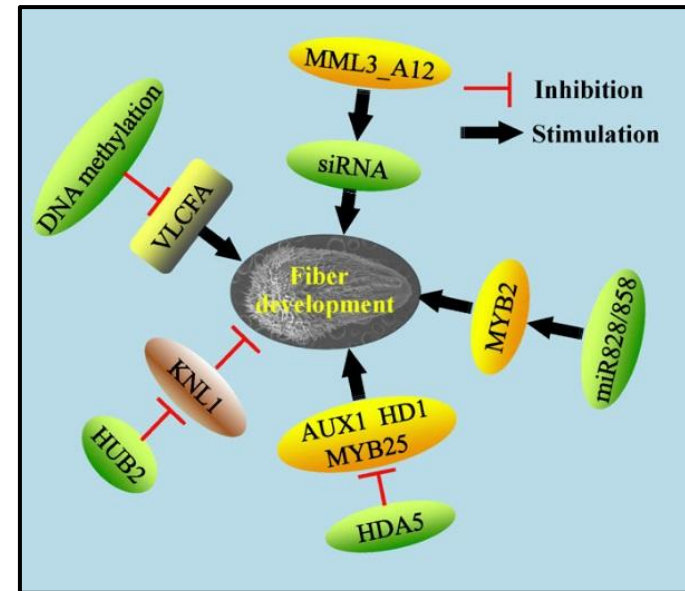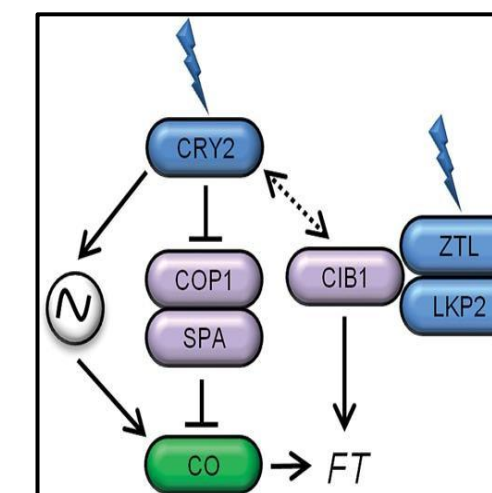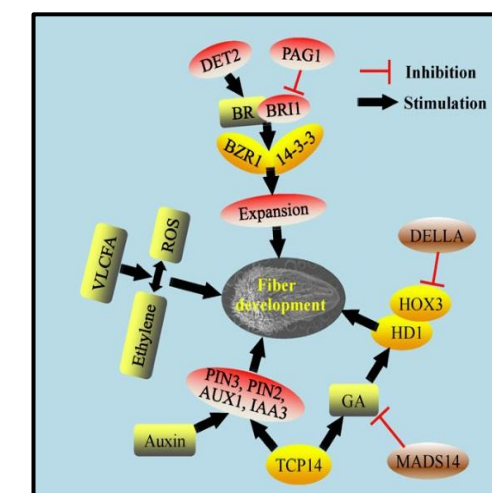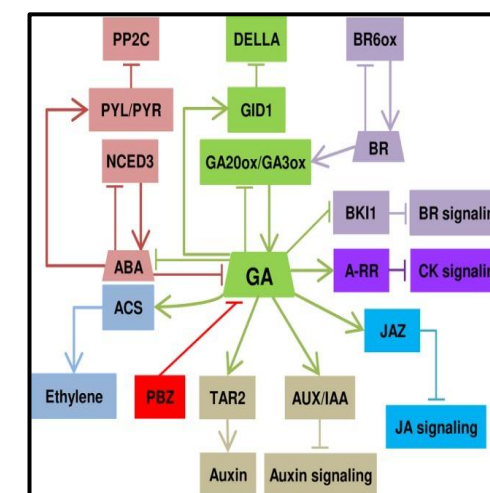
datlombardo.ca

vclab.ca

Gene Model Map

Desired Output

- DNA methylation inhibits VLCFA
- MML3_A12 stimulates siRNA
- miR828/858 stimulates MYB2
- HDA5 inhibits AUX1 HD1 MYB25
- HUB2 inhibits KNL1
- VLCFA stimulates Fiber Development
- siRNA stimulates Fiber Development
- MYB2 stimulates Fiber Development
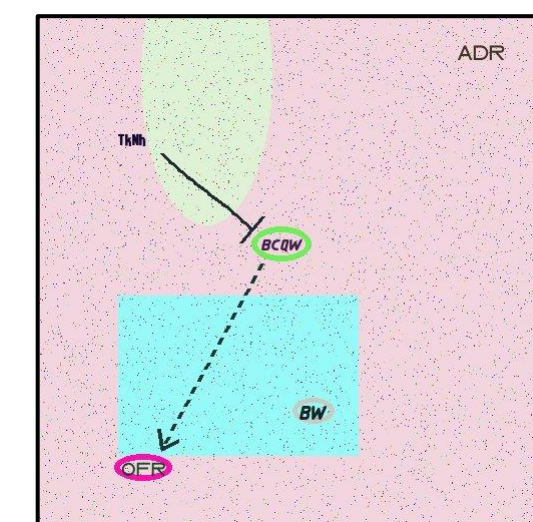- AUX1 HD1 MYB25 stimulates Fiber development
- KNL1 inhibits Fiber Development

BAR Toronto

- University of Toronto's BAR website attracts 1 million visits per month by biological researchers.

Gene Expression and Protein Tools

View expression patterns as electronic fluorescent pictographs or heatmaps, explore promoters, identify protein-protein interactions and more.

ePlants

BAR's Visualization Selection - bar.utoronto.ca

AT3G24650 - AP2/B3-like transcriptional factor family protein. AT3G24650 has 1 splice variant. The highest subcellular localization of AT3G24650 is the nucleus. AT3G24650 has been shown experimentally to interact with 94 proteins, including At2g25890 and At4g17730.

Recent publications have shown the following conclusions: 1) Interestingly, PIL5 directly regulates many transcription regulators of hormone signaling and indirectly regulates many genes involved in hormone metabolism. 2) Breakdown of Chl is a proactive process during the late stage of seed maturation, as well as during leaf senescence and fruit ripening. 3) The interaction of BES1 with ABI5 significantly suppressed the binding of ABI5 to the promoter regions of downstream genes, which resulted in their reduced expression and consequently facilitated seed germination.

Related genes:
AT2G20180
AT3G36270
AT2G22640
AT1G29170
AT1G14920

Query result of ABI3 on BAR's GAIA.
bar.utoronto.ca/gaia/search?query=ABI3

Datasets

GeneNet

Name	# Models	# Non-Models
GeneNet-98	55	43
GeneNet-500	208	292

GeneNetSyn

Blob Count: 5 Rela Count: 2

Darknet Labels
0, 0, TkNh, 0.1816, 0.2382, 0.1191, 0.06835
0, 1, oFR, 0.2031, 0.875, 0.1132, 0.0527
0, 2, bcqw, 0.4726, 0.4394, 0.1171, 0.0683
0, 3, BW, 0.5351, 0.7656, 0.0820, 0.0507
0, 4, ADr, 0.8613, 0.0644, 0.1347, 0.0507
1, -1, inhibition, 0.336, 0.3515, 0.2382, 0.211
4, -2, inhib-head, 0.414, 0.4121, 0.125, 0.125
2, -1, activation, 0.332, 0.66, 0.2421, 0.4336
3, -2, act-head, 0.2383, 0.831, 0.0938, 0.0860

Non-Normalized Blob Locations
X Y W H
[93, 122, 61, 35]
[104, 448, 58, 27]
[242, 225, 60, 35]
[274, 392, 42, 26]
[441, 33, 69, 26]

Relationship Information
0, TkNh, 2, bcqw, 1 → tkNh inhibits bcqw
2, bcqw, 1, oFR, 2 → bcqw activates oFR

Bounding Box Edges
[56, 101], [130, 101], [56, 143], [130, 143]
[69, 432], [139, 432], [69, 464], [139, 464]
[206, 204], [278, 204], [206, 246], [278, 246]
[249, 376], [299, 376], [249, 408], [299, 408]
[400, 17], [482, 17], [400, 49], [482, 49]
[1, [[111, 126], [233, 126], [111, 234], [233, 234]]
[2, [[108, 227], [232, 227], [108, 449], [232, 449]]]

Ground Truth

- JEB inhibits pjsWZ
- pjsWZ activates BBBRG
- BBBRG activates pQbT
- pQbT activates UyE

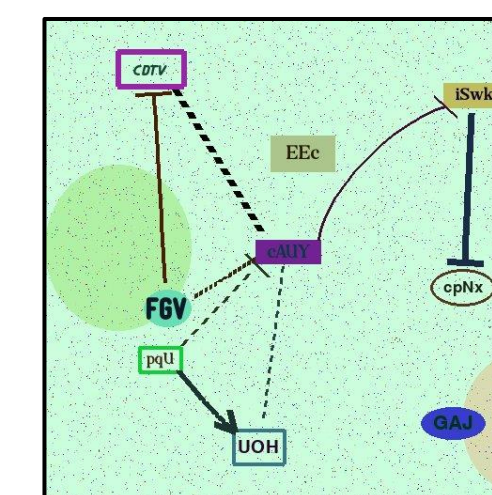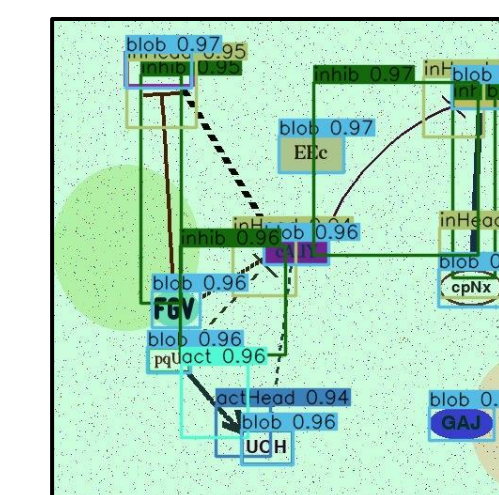## Acknowledgements

Region Detection

- **Problem**: Identify the location of all regions of interest which include both constituents and relationships. Given a relationship we need to determine where the head is located, and it's type.
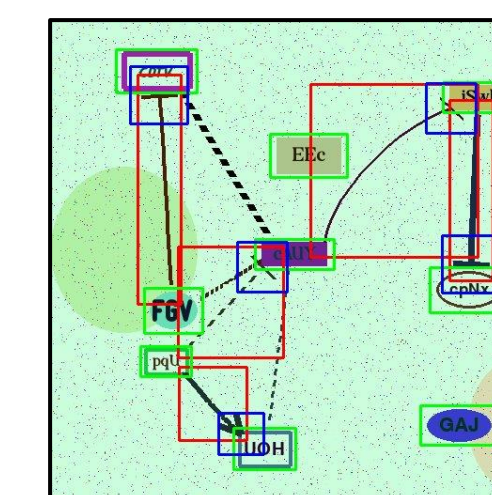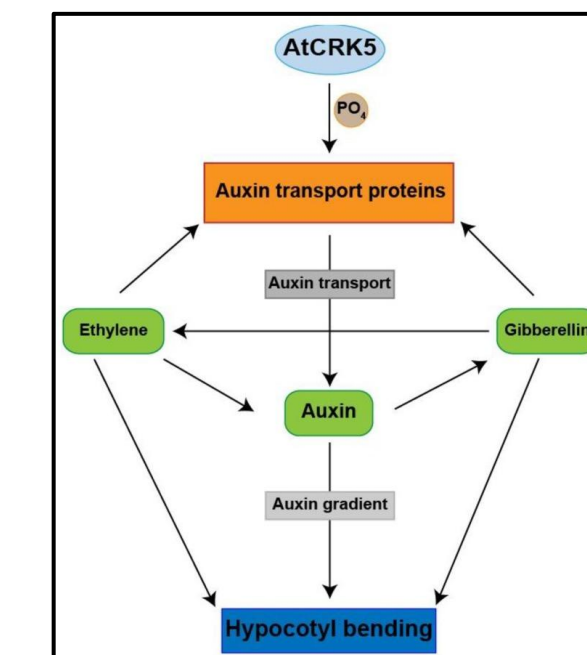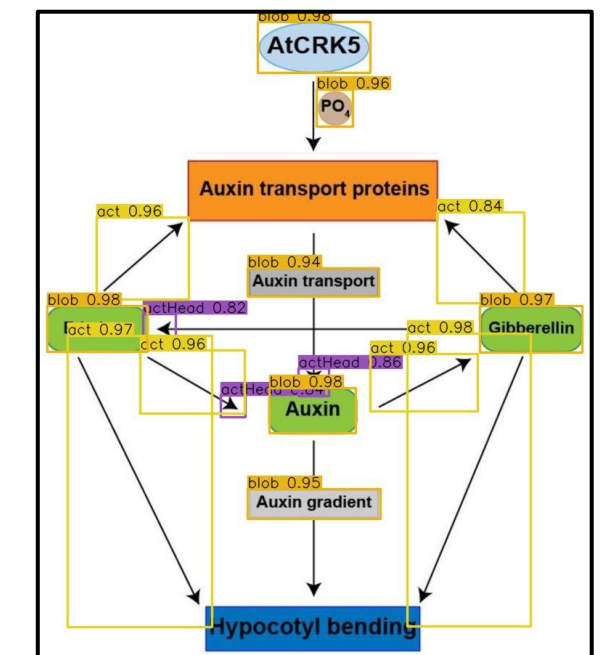
GeneNetSyn

Input Image   Prediction   Ground Truth
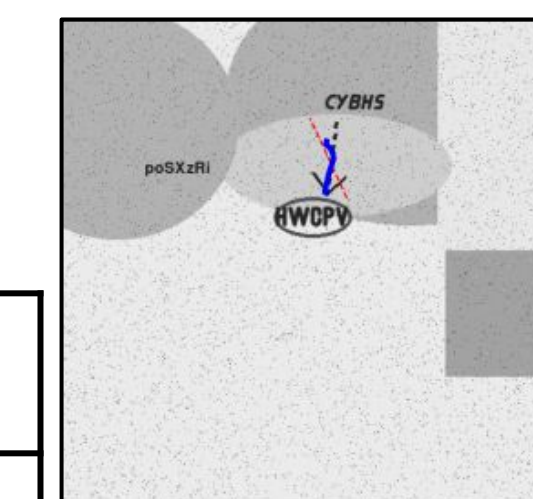
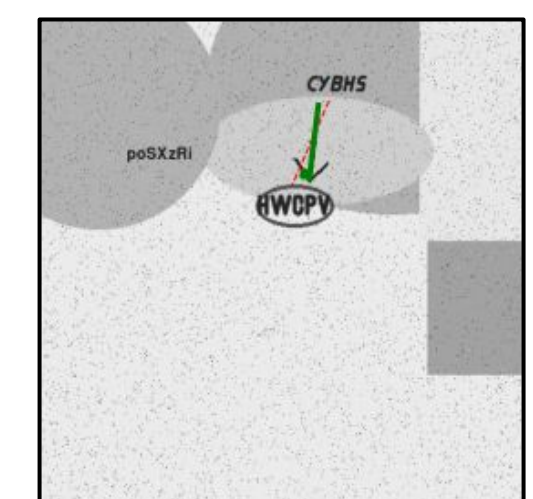### GeneNet

Input Image   Prediction

Parsing Relationships

- **Problem**: Parse the output of the region detection step, for each relationship find the two constituents which are connected on both edges. Develop a textual description such that *A Relationship B*.

Active Contour

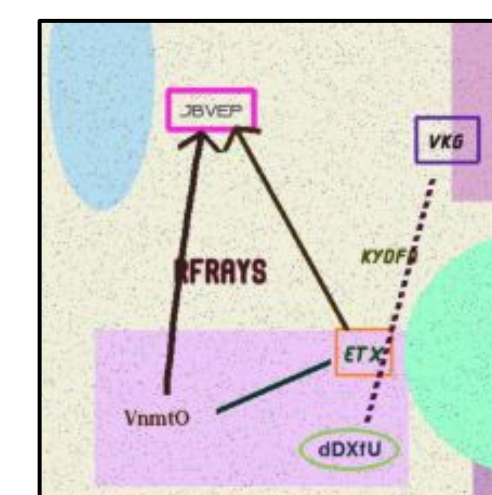| Confidence	Relationship Prediction	Blobs Detected	Relationships Detected
80%	80.22%	99.99%	98.62%
90%	77.23%	92.92%	92.90%

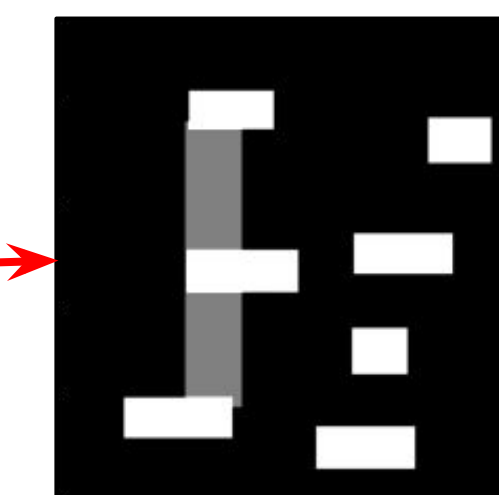Left to Right Contour Result

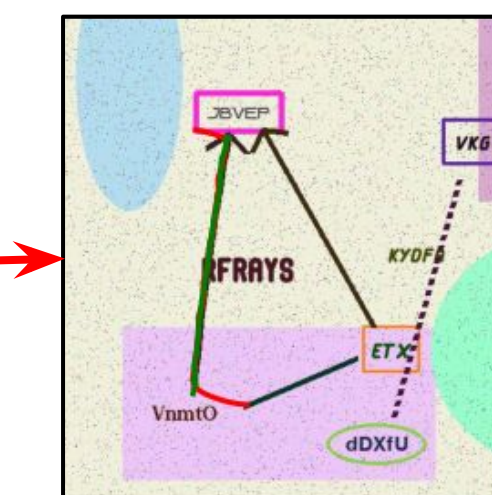Right to Left Contour Result

Throughput of a Single Relationship from GeneNetSyn
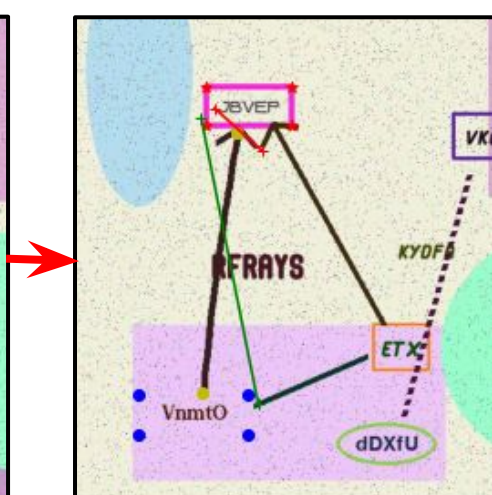
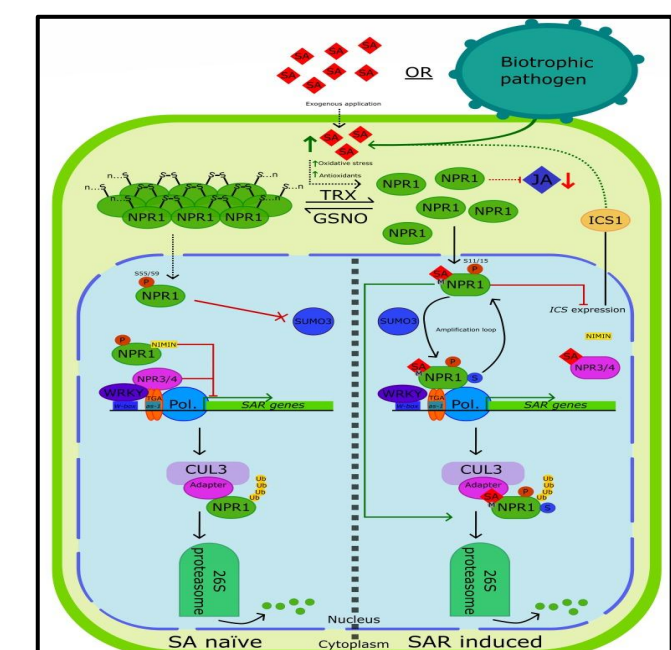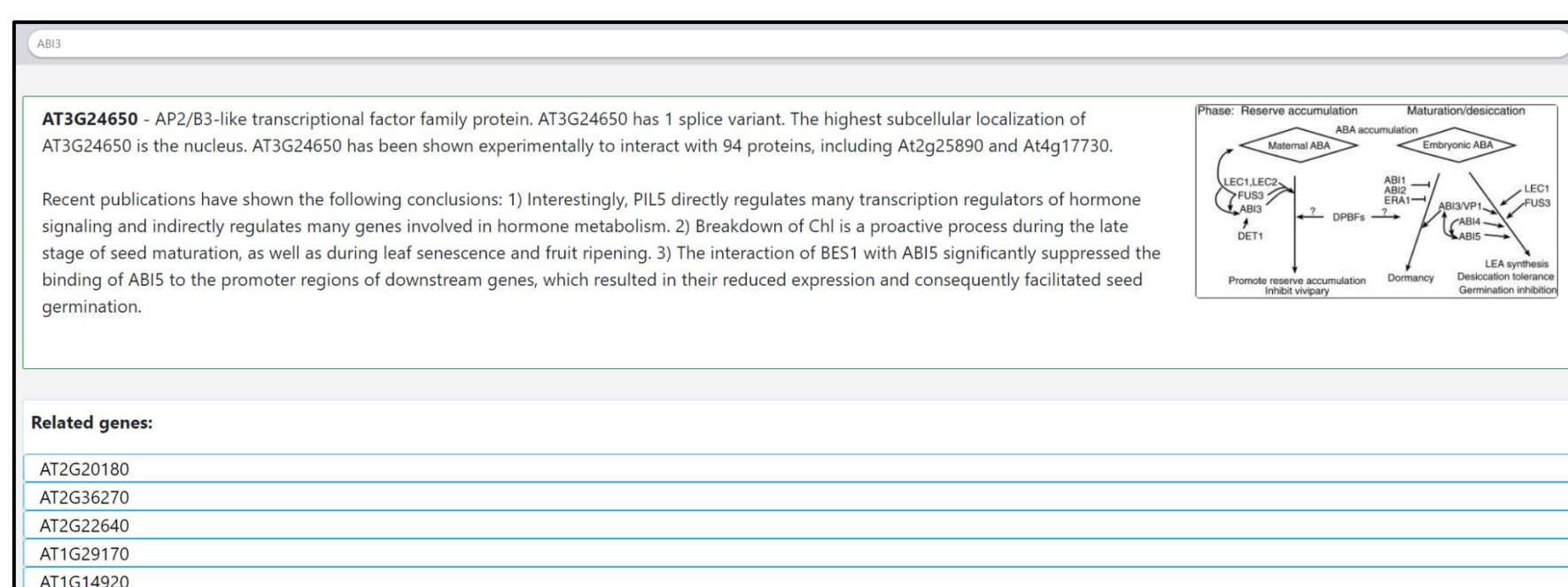Input Image   Region Detection Output   Active Contour Result   Parsed Result

Final Result

VnmtO activates JBVEP

Future Work

- Tune approach for the most complex biological diagrams
- Integrate system into BAR's GAIA web browser
- Expand search space into new sets of genes, introducing wider spread of diagram understanding adoption.