**ToxCast Embryonic Stem Cell (H9) Profiling for Predicting Developmental Toxicity**

Thomas B. Knudsen¹ Todd J. Zurlinden¹, Katerine S. Saill¹, and Nancy C. Baker¹,²

¹National Center for Computational Toxicology and ²Leidos – Research Triangle Park, NC 27711

---

**Introduction**

The protocol commonly used to test for prenatal developmental toxicity (i.e., OECD TG 414) is based on observation of fetal malformations, usually in pregnant rats and/or rabbits. EPA is evaluating new approach methodologies (NAMs) that can be used to quickly evaluate the human toxicity potential of chemicals with less reliance on animal testing.

ToxCast generates in vitro data on thousands of chemicals utilizing high-throughput screening (HTS) methods. To increase the assay space for predicting human developmental toxicity, we profiled ToxCast chemicals in the devTOXqP assay [1]. The assay measures a critical drop in the ornithine/cystine ratio H9 hESC culture medium. Data for 1062 chemicals was pipelined in ToxCast to generate the STM dataset (pending release); a positive signal for developmental toxicity was elicited by 183 (17%) [2].

Here, we assessed the predictive value of the STM data with a subset of 432 ToxCast chemicals tested for prenatal developmental toxicity.

**1. Binary classification**

<table>
<thead>
<tr>
<th>Condition</th>
<th>STM</th>
<th>BR2</th>
<th>BR1</th>
<th>BR4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>44</td>
<td>30</td>
<td>43</td>
<td>45</td>
</tr>
<tr>
<td>Medium</td>
<td>88</td>
<td>74</td>
<td>90</td>
<td>92</td>
</tr>
<tr>
<td>High</td>
<td>82</td>
<td>71</td>
<td>80</td>
<td>83</td>
</tr>
</tbody>
</table>

**2. Recursive analysis (high-stringency model)**

- Machine-learning with 5-fold cross-validation on train/test split;
- >200 ToxCast features correlated with developmental toxicity;
- STM was the top-weighted feature but others augment predictivity.

**3. Biological plausibility**

- Cluster analysis of STM (hESCs) and CREB/NRF2/PXR (Hep2a) hit-calls show a generally inverted correlation (n=127 chemicals in the high-stringency model);
- Assuming STM measures an adverse response, the inverted correlation infers an adaptive response for the following pathways:
  - CREB→↑ xenobiotic metabolism and export (XME pathway);
  - sets up a hierarchical rules-based workflow for building a decision workflow:
    - Rule 1: STM (+) & CREB3 (+) predicts TN condition in 19 of 22 cases (86.4%);
    - Rule 2: CREB3/NRF2/PXR (+) predicts TN condition and overrides STM (+) call;
    - Rule 3: STM (+) & PXR (+) or NRF2 (+) predicts TN condition in 63 of 69 cases (91.3%);
    - Rule 4: STM (+) & CREB3/NRF2/PXR (+) condition predicts TN in 25 of 30 cases (83.3%).

**4. Binary classification refined (work in progress)**

**5. Conclusions**

The STM biomarker surfaced as the top weighted feature for developmental toxicity in the ToxCast in vitro portfolio. Balanced accuracy (BAC) reached 78% with well-supported evidence for developmental toxicity but depreciated to 62% with when criteria for developmental toxicity was relaxed. Augmenting the STM response with data from several ToxCast assays that measure adaptive cellular responses (UPR, ARE, XME pathways) improved positive predictive value to BAC ~90%.

---

**Specific Aims**

- Binary classification to correlate STM responses with animal models of human prenatal developmental toxicity captured from ToxRefDB and other sources of in vivo data; and
- Recursive partitioning: mining in vitro bioactivity profiles of >800 ToxCast in vitro assays to assess the value of the STM dataset for predicting human developmental toxicity.

---

**ToxCast Embryonic Stem Cell (H9) Profiling for Predicting Developmental Toxicity**

Thomas B. Knudsen¹ Todd J. Zurlinden¹, Katerine S. Saill¹, and Nancy C. Baker¹,²

¹National Center for Computational Toxicology and ²Leidos – Research Triangle Park, NC 27711

---

**Introduction**

The protocol commonly used to test for prenatal developmental toxicity (i.e., OECD TG 414) is based on observation of fetal malformations, usually in pregnant rats and/or rabbits. EPA is evaluating new approach methodologies (NAMs) that can be used to quickly evaluate the human toxicity potential of chemicals with less reliance on animal testing.

ToxCast generates in vitro data on thousands of chemicals utilizing high-throughput screening (HTS) methods. To increase the assay space for predicting human developmental toxicity, we profiled ToxCast chemicals in the devTOXqP assay [1]. The assay measures a critical drop in the ornithine/cystine ratio H9 hESC culture medium. Data for 1062 chemicals was pipelined in ToxCast to generate the STM dataset (pending release); a positive signal for developmental toxicity was elicited by 183 (17%) [2].

Here, we assessed the predictive value of the STM data with a subset of 432 ToxCast chemicals tested for prenatal developmental toxicity.

**1. Binary classification**

<table>
<thead>
<tr>
<th>Condition</th>
<th>STM</th>
<th>BR2</th>
<th>BR1</th>
<th>BR4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>44</td>
<td>30</td>
<td>43</td>
<td>45</td>
</tr>
<tr>
<td>Medium</td>
<td>88</td>
<td>74</td>
<td>90</td>
<td>92</td>
</tr>
<tr>
<td>High</td>
<td>82</td>
<td>71</td>
<td>80</td>
<td>83</td>
</tr>
</tbody>
</table>

**2. Recursive analysis (high-stringency model)**

- Machine-learning with 5-fold cross-validation on train/test split;
- >200 ToxCast features correlated with developmental toxicity;
- STM was the top-weighted feature but others augment predictivity.

**3. Biological plausibility**

- Cluster analysis of STM (hESCs) and CREB/NRF2/PXR (Hep2a) hit-calls show a generally inverted correlation (n=127 chemicals in the high-stringency model);
- Assuming STM measures an adverse response, the inverted correlation infers an adaptive response for the following pathways:
  - CREB→↑ xenobiotic metabolism and export (XME pathway);
  - sets up a hierarchical rules-based workflow for building a decision workflow:
    - Rule 1: STM (+) & CREB3 (+) predicts TN condition in 19 of 22 cases (86.4%);
    - Rule 2: CREB3/NRF2/PXR (+) predicts TN condition and overrides STM (+) call;
    - Rule 3: STM (+) & PXR (+) or NRF2 (+) predicts TN condition in 63 of 69 cases (91.3%);
    - Rule 4: STM (+) & CREB3/NRF2/PXR (+) condition predicts TN in 25 of 30 cases (83.3%).

**4. Binary classification refined (work in progress)**

**5. Conclusions**

The STM biomarker surfaced as the top weighted feature for developmental toxicity in the ToxCast in vitro portfolio. Balanced accuracy (BAC) reached 78% with well-supported evidence for developmental toxicity but depreciated to 62% with when criteria for developmental toxicity was relaxed. Augmenting the STM response with data from several ToxCast assays that measure adaptive cellular responses (UPR, ARE, XME pathways) improved positive predictive value to BAC ~90%.