Connecting small-molecule structure to biological activity

Predictive models of small-molecule transcription factor binding

Isaac Joseph
Why find structure activity relationship (SAR)?

• Drug and probe development
  – Efficient enhancement through structure optimization
How is SAR currently found?

• Visual inspection of compound structures
• Limited computational techniques
  – One biological target, or
  – Constrained library diversity

What is our new approach for finding SAR?

Compound substructures

Influence binding with

Transcription factors

Substructural importance model

Importance

Predicting

Revealing

Compound promiscuity

Binds n TFs

Used in
How do we find SAR?

1. Describe structure and activity
2. Create substructure-based predictive model
3. Interpret results
How do we represent structure?

The chemical compound on the left has these substructures, which are represented in the fingerprint matrix on the right. The matrix shows which substructural features are present in the compounds.
How do we represent activity?

H₂N\(\text{O} \quad \text{O} \quad \text{C}(\text{H})\text{I}

binds with

promiscuity vector

3 compounds
How do we find SAR?

1. Describe structure and activity
2. Create substructure-based predictive model
3. Interpret results
How does our model work?

fingerprint matrix
promiscuity vector

substructural features
How does our model work?

- fingerprint matrix
- substructure importance
- promiscuity vector

$\approx$

substructural features

$\times$

$\approx$
How do we handle different interactions?

fingerprint matrix

contains  

lacks  

...
How do we find SAR?

1. Describe structure and activity
2. Create structure-based predictive model
3. Interpret results
What substructures influence promiscuity in certain cases?
How do we find SAR?

1. Describe structure and activity
2. Create structure-based predictive model
3. Interpret results
What will we do in the future?

• Further investigate chosen substructures
• Different representations of promiscuity
• Generate larger substructures
• Additional sets of compounds
Acknowledgements

Analysis
Paul A. Clemons (mentor)
Joshua C. Gilbert
Hyman A. Carrinksi
Sandrine Muller
Nicole E. Bodycombe
Adam A. Margolin

Chemistry
Angela N. Koehler

Administration
Eboney Smith
Bruce Birren
How do we find SAR?

1. Describe structure and activity
2. Create structure-based predictive model
3. Interpret results