

# Using networks to predict the molecular basis of toxicity

Gordana Apic<sup>1</sup>, Frank Bonner<sup>1</sup>, Bojana Cosovic<sup>1</sup>, Robert B. Russell<sup>2</sup>

1. Cambridge Cell Networks Ltd., St John's Innovation Centre, Cowley Road, Cambridge, CB4 0WS, UK  
2. EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany

A common problem for toxicologists is finding a molecular rationale for unexpected toxic endpoints. In this case study we attempt to solve that problem using ToxWiz™, a comprehensive software and database system designed to help make decisions in toxicology and many other areas of drug discovery.

## Introduction

A common problem for toxicologists is finding a molecular rationale for unexpected toxic endpoints.

At a conference in 2006, Boyer<sup>1</sup> described an unexpected observation of testis degeneration in rats, when studying a series of compounds consisting of a "lipophilic core connected to a nicotinic acid". Subsequent work suggested that undesirable interactions with retinoic acid receptors (RARs) might be the molecular basis of this effect.

Although Boyer was not able to disclose full details of the compounds because they are proprietary, he nevertheless issued a challenge to those present. Is it possible to integrate data in such a way that diverse sources of information such as gene-knockouts in model organisms, chemistry and pathology data can be combined to reveal the likely mechanism of this (and other) unexpected toxic endpoints?

Here we attempt to answer that challenge using ToxWiz™, a comprehensive software and database system designed to help

make decisions in toxicology and many other areas of drug discovery.

The system works by exploiting a network of thousands of interacting genes, proteins, chemicals, pathways and biological effects, including toxic endpoints and diseases. The neighbourhood of a molecule in the system captures everything that has been observed about it in the past, either in the public domain or in company proprietary datasets.










## Chemical search for nicotine

Our first step is to perform a substructure search of the ToxWiz database for compounds linked to biology known to involve a nicotinic group. This search yields 19 compounds containing the nicotinic group (Figure 1).

These 19 are then inspected to identify those that contain lipophilic substituents, like the compounds described by Boyer.

Removing compounds that do not contain true nicotinate groups (for example, the drug Amlexenox, in which the nicotinate group is part of a larger ring system) and those without

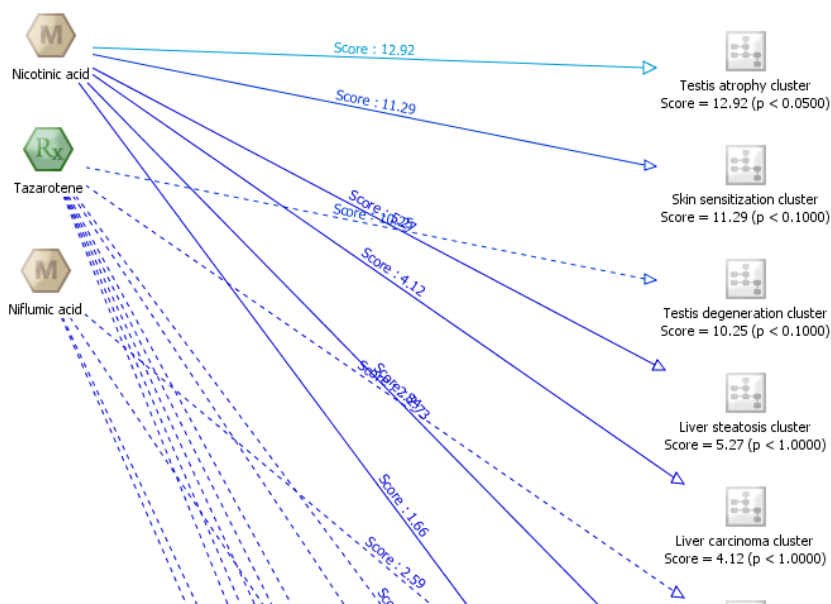
lipophilic substituents, leaves us with between 10 and 12 compounds, depending on how we define the lipophilic groups.

Structure	Molecule Name	Molecule ID	Number of atoms ma
	Nicotinic acid		930 matches 9 of 9 1.0
	Nicotinate		937 matches 9 of 9 1.0
	Quinolnic acid		1066 matches 9 of 9 1.0
	Cidonic acid		68703 matches 9 of 9 1.0
	Imazapyr		54738 matches 9 of 9 1.0
	Nifluric acid		4488 matches 9 of 9 1.0
	Nicotinate mononucleotide		941 matches 9 of 9 1.0
	Amlexenox		2161 matches 9 of 9 1.0
	ZINC04117070		290443 matches 9 of 9 1.0

**Figure 1** Some of the 19 compounds returned using a chemical substructure search with nicotinic acid (highlighted)

## Mapping the compound set onto toxic endpoints

This set of molecules is then mapped onto over 400 toxic endpoint clusters in the ToxWiz database (Figure 2).



**Figure 2** Mapping compounds containing nicotinic acid without significant polar group substituents

Each cluster contains chemicals, genes and proteins that have previously been associated with each endpoint. The mapping process considers whether each molecule in the set is inside a given cluster, or if it interacts with molecules inside a cluster. The clusters are then ranked according to a statistical significance which accounts for the abundance of proteins in biological processes and the size of pathways. For interactions it also considers the strength, or confidence, of the interaction, for example giving a greater weight to manually curated interactions than those determined from text-mining. The clusters are then listed in order of significance on the right of the map.

Testis atrophy and testis degeneration are among the highest ranked clusters (first and third, respectively). You can also see that only three of the compounds map with any significance to endpoints. Inspection shows that several of the compounds are from high-throughput assays in PubChem and lack any literature evidence about their biological function.

The power of the network is illustrated by the map in Figure 2.

Solid lines indicate the direct presence of molecules inside a cluster, implying that there has been at least one previous observation of a direct relationship between the molecule and the endpoint.

By contrast, dashed lines indicate an indirect relationship, in which the molecule interacts with something (usually a protein) inside the endpoint cluster. Such compounds might have a secondary effect.

### Previously observed xenobiotics that cause testis problems

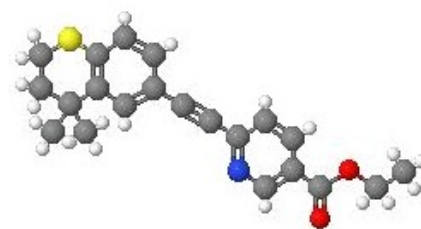
The association of nicotinic acid to problems in the testis is most likely due to processes downstream from, or related to, toxicity induced by phthalates and other similar molecules<sup>2</sup>. Phthalates are known to influence RAR in the course of causing testis developmental problems<sup>3,4</sup>. Searches within the system also identify key problems.

### Nicotinates with lipophilic substituents acting on RAR $\alpha$

One of the compounds containing lipophilic attachments is the drug Tazarotene (Figure 3), an optical

retinoid used for the treatment of psoriasis. This compound is more active against RAR beta or gamma than alpha, but it does show activity against RAR alpha<sup>5</sup>.

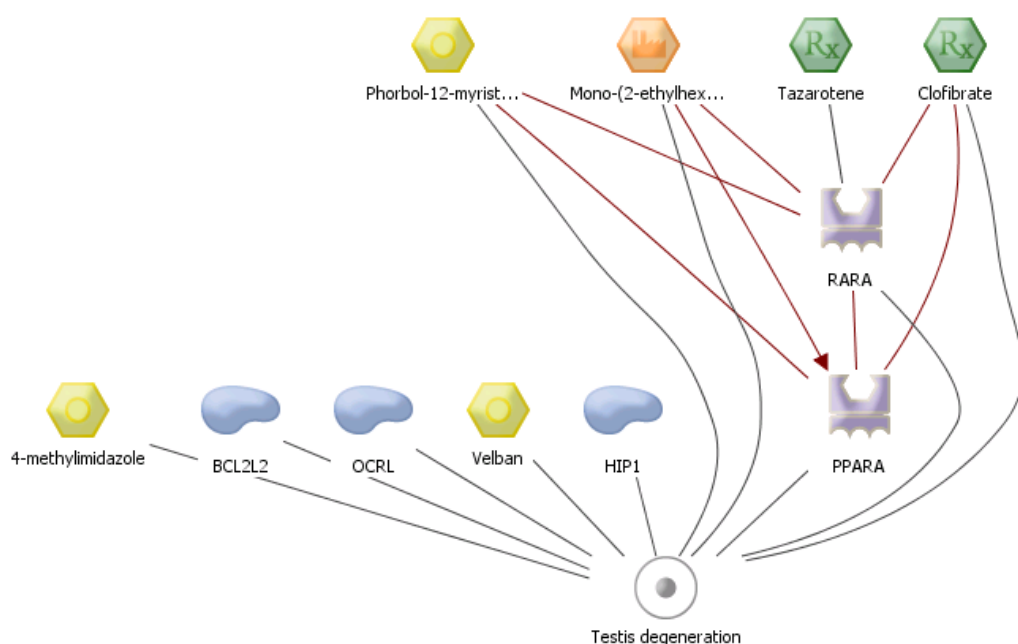
Tazarotene acts by affecting activities of RARs. The compound itself is not associated with testis problems, only with the receptor (hence the dashed lines in Figure 2). Indeed, no evidence of testis problems has been found for this compound. Although, as it acts by targeting this receptor that is, perhaps, unsurprising. But because it conforms to the pattern mentioned by Boyer in his presentation, the proprietary compounds could be acting in a similar way.



**Figure 3** Tazarotene

### The molecular details of the prediction

Displaying the cluster for testis degeneration and then adding to it the compounds from the search, creates a pathway-like diagram that might explain how the compounds are eliciting their effect (Figure 4). In particular, relationships with both RAR $\alpha$  and peroxisome proliferative activated receptor alpha (PPAR $\alpha$ ) are apparent among the compounds known to cause developmental problems in the testis.



**Figure 4** The testis degeneration cluster, showing connections between proteins/chemicals, and with Tazarotene added (and connected to RAR $\alpha$ )

## Summary

This retrospective analysis demonstrates the great predictive potential of ToxWiz. Moreover, its ability to deal with the concept of a group of compounds rather than just one compound can be a great advantage when studying a series of molecules, or when considering the best substituents to try during lead optimization studies. Knowledge of previous studies is always a great benefit in designing new ones, and ToxWiz can be an effective way of finding this information quickly.

## References

1. Boyer, S *In Silico* Toxicity Predictions – Present and Future Tools in Systems Biology *International Society for the Study of Xenobiotics (ISSX)*, Manchester, UK, June 2006.
2. Fukuwatari et al (2003) The effects of phthalate esters on the tryptophan-niacin metabolism. *Adv Exp Med Biol* **527**:659-664.

3. Dufour et al (2003) Peroxisome proliferators disrupt retinoic acid receptor alpha signaling in the testis. *Biol Reprod* **68**:1215-1224.

4. Lufkin et al (1993) High postnatal lethality and testis degeneration in retinoic acid receptor alpha mutant mice. *Proc Natl Acad Sci USA* **90**:7225-7229.

5. Bianchi et al (2004) Topical treatment of basal cell carcinoma with tazarotene: a clinicopathological study on a large series of cases. *Br J Dermatol* **151**:148-156.

## About Cambridge Cell Networks

Cambridge Cell Networks (CCNet) based in Cambridge, UK supplies a range of industry-leading content on biological pathways, chemistry and toxicology, combined with an integrated pathway visualisation and exploration tools to the pharmaceutical and biotechnology industries. Using cutting-edge biological and computational methods combined with knowledge management techniques, CCNet

offers a novel approach to pathway analysis, providing effective target validation and predictive toxicology data, which will ensure the production of safer drugs.

The Company was founded in 2002 by eminent scientific figures from the European Bioinformatics Institute in Cambridge, UK, the University of Cambridge and the European Molecular Biology Laboratory in Heidelberg, Germany. CCNet has facilities in three countries and is staffed by a team of expert biochemists, pharmacologists, bioinformaticians, chemists and industrial toxicologists.