# CRACK IT

## Challenge 26: DARTpaths

# DARTpaths: Mapping developmental and reproductive toxicity (DART) genes and pathways for cross-species comparison of toxic compound effects

## Sponsors:



Launch Meeting

7 September 2017

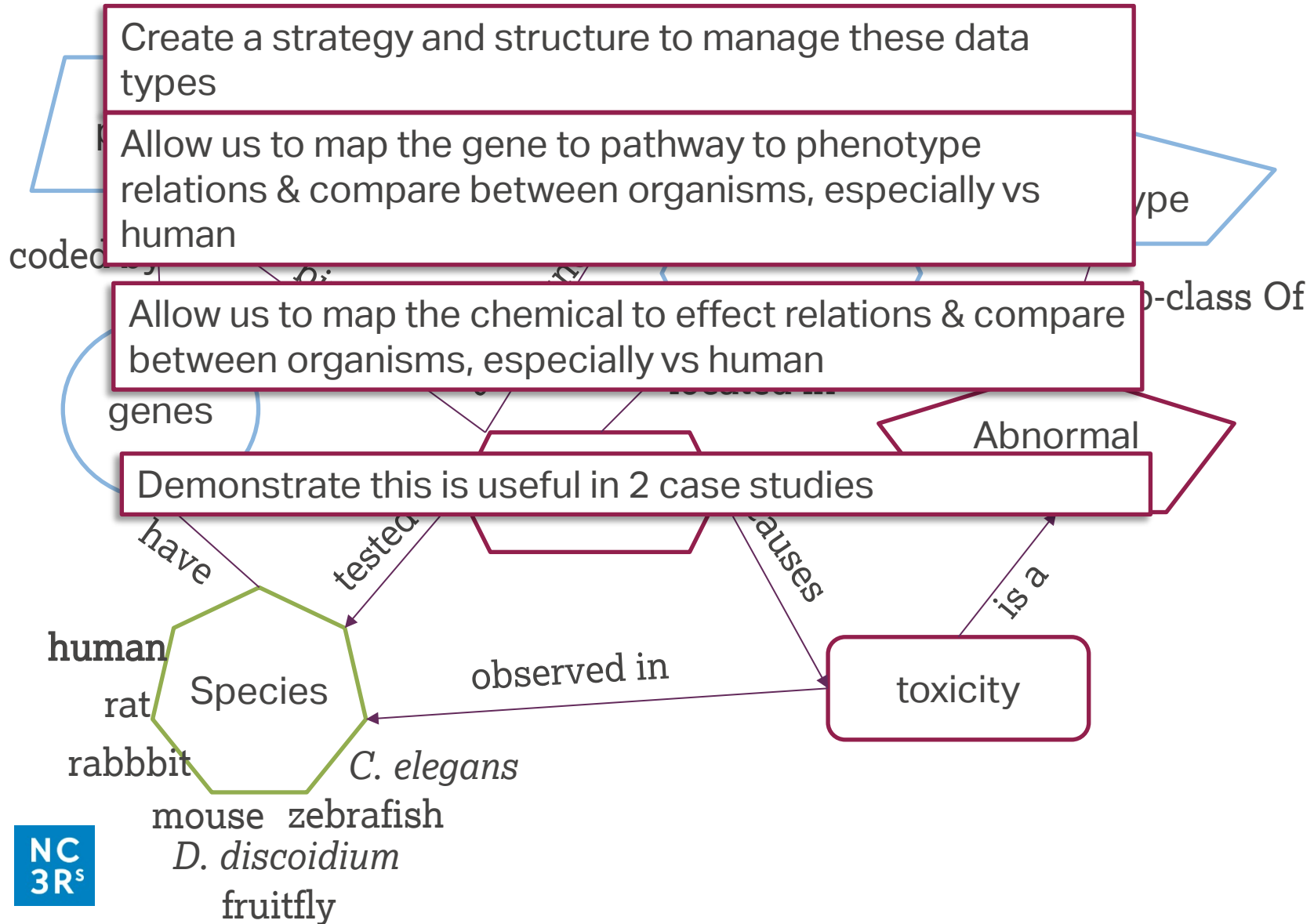M. Rooseboom (Shell) & R. Currie (Syngenta)

# The Challenge:

Create a strategy and structure to manage these data types

Allow us to map the gene to pathway to phenotype relations & compare between organisms, especially vs human

Allow us to map the chemical to effect relations & compare between organisms, especially vs human

Demonstrate this is useful in 2 case studies

coded by

genes

ype

-class Of

Abnormal

have

tested

auses

is a

human

rat

Species

observed in

toxicity

rabbbit

C. elegans

mouse  zebrafish

D. discoidium

fruitfly

NC 3Rs

# Competency questions

For a **particular chemical** that requires testing on reproductive toxicology endpoints, which **alternative species is most representative for man**?

**Which alternative species possess known reproductive toxicology pathways**, and **how do these compare across the alternative species**?

When **a DART pathway** has been identified, **which alternative species has this pathway** and **most resembles relevant pathways in man**? Are some species better suited than others to address specific parts of the reproductive cycle e.g. fertility (male/female), prenatal development, multigenerations

For a particular chemical that requires testing on reproductive toxicology endpoints, for which tests have examples in the **same chemical space been tested**?

Is it overly conservative looking at **omics patterns and/or biomarkers in comparison with phenotypic effects** for the alternative species?

**NC 3Rˢ**

# We aim for

A system to answer the competency questions for at least the chemical spaces and DART pathways that are relevant to us

## and

It should be user friendly and contain the species mentioned

## to enable this

DART targets (MIE) and pathways and available data and existing knowledge on developmental biology and (chemical-induced) developmental defects are to be included

NC3Rˢ

# Why was this Challenge developed?

Scientific

permit the integration of different data related to DART effects
deal with the complexity of the problem
create a tool to provide evidence that supports the use of particular
test

systems in a DART IATA

Business

need to ~~deal~~ predict DART effects, more rapidly and cost effectively
with      no decrease in certainty of prediction

3Rs

enables the best use of non-animal or model species tests
longer term opportunity to replace DART animal studies

# Current situation on DART in regulatory domain

- Use large numbers of animals (primarily rats and rabbits)

  - OECD 414 prenatal development test ~800 rats or ~900 rabbits. An OECD 443 extended one generation test ~1400 rats for a minimal basic design, i.e. no cohorts and extension to F2 generations.

- Use a large quantity of test material

- Provide limited insight to aid chemical design

- Are expensive and time consuming

- And the relevance to humans of effects seen in these studies is not always clear

NC 3Rs

# Current state of the art on alternatives

- Current state of the art is based on test developers creating tests and assessing their validity using a confusion matrix assessment vs a "gold standard"

- It isn't clear when a test system is most appropriate to use

- It isn't clear why some alternative tests work, or whether they are applicable to humans. A move to mechanistic toxicology should address this

- Gathering evidence and interpretation is a manual process using a lot of distributed data/bioinformatics resources

- Uncertainty on the best way to identify orthologs in multigene families

# Current situation in DART science

- In recent years increasing interest in using cell-based, non-mammalian and computational approaches to improve our ability to predict chemicals with DART potential

- However, the current understanding of DART AOPs is incomplete and, potentially, many mechanisms could result in DART effects (National Research Council, 2000)

- Developmental biology is characterised by a complex interplay between a multitude of processes at the molecular, cellular, tissue and organism level, which change continuously with time in development and location in the conceptus. (ie an adaptive self organizing complex system)

- A clear need for integration of available information on the relation between particular genes and particular physiology in the various model organisms and the mapping of those relations between model organisms.

# Deliverables (1)

- Development of a **data management structure that is stable and allows reliable updates** of the content to provide a visualisation of the gene-pathway data and assessment of the validity of the routes and links by comparing expression data for various organisms that have been exposed to the same compound, to **enable a side-by-side comparison of pathway perturbation in the different species**.

- Must be able to:
  -**Link orthologous genes between model organisms** (human, mouse, rat, rabbit, zebrafish, fruitfly, nemotode and slime mould), with a focus on genes involved in developmental pathways. Need a method that provides appropriate linking of genes from large paralog families.
  -**Links orthologous pathways** between organisms and genes to these pathways
  -**Captures the "chemical space"** of chemicals that have been tested in each model system, to define the chemical applicability domain for the test

# Deliverables (2) Phase 1

- Identification of further **key competency questions** (in addition to those in appendix 3) that the toxicology community has regarding models for DART prediction.

- **Identification and prioritisation of key DART pathways & MIEs** (based on the chemical spaces of interest) to use in the data model to demonstrate a cross-organism mapping strategy.

- An **initial data model** (describing the concepts and relationships between them that will be) containing the species of interest and a prioritised pathway as an exemplar.

- A **plan to establish which mechanistic data** from *in vitro* cell models and alternative non-mammalian model organisms in the chemical spaces being mentioned must be generated within the project.

- A plan to fully develop and implement this to meet the key deliverables as defined above.

- A proposal for **long term accessibility and sustainability** of the data model and tools.

- **Robust plans to deliver Phase 2** of the Challenge including commercialisation and dissemination.

# Deliverables (3) Phase 2

- Extend to develop a cross-organism mapping strategy for **multiple DART pathways**.

- **Generation of mechanistic data** from *in vitro* cell models and alternative non-mammalian model organisms for chemical classes where data is missing, and integrate such data into the model.

- Develop a user friendly **graphical user interface** (GUI) that permits the users to interact with the data to answer the competency questions.

- Demonstration that the framework works for the **two PoCs** and that additional requirements to improve the model are identified.

- Demonstration that appropriate **dissemination activities/plans** are in place to maximise uptake of the developed tool.

- CRACK IT Challenges competition is designed to support the development of new 3Rs technologies and approaches, which will improve business processes and/or lead to new marketable products. **The application must include a plan to commercialise the results into a product or service**. This should be taken into consideration when completing your application.

**NC 3Rs**

# The two Proof of Concepts (POCs)

**Identify key data gaps in mechanistic understanding and demonstrate that newly generated mechanistic data for classes of substances with these gaps (e.g. UVCBs\*) can be integrated;**

**To evaluate whether this framework and mechanistic data could have predicted a developmental toxicity *de novo* in a novel chemical class (e.g. of a crop protection chemical).**

\*For UVCB (Chemical Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials) substances (e.g. petrochemicals/petroleum products), which comprise about a quarter of the chemicals on the European market by volume, limited mechanistic data is available while there is a great amount of guidance-driven conventional *in vivo* data present and could be shared by the sponsors

# What we don't want

A pack of great publications but not a usable software tool

Focus only on irrelevant or least relevant pathways for the chemical spaces mentioned

We are NOT after new DART assays/ replacement tests, although the ability to extend this to new tests in the future is needed

Extension into other species is possible but not at the expense of the key species mentioned

NC 3Rs

CRACK IT

# Sponsor in-kind contributions

Know-how on DART in the chemical industry

Input into prioritisation of DART pathways

Software user requirements

Where available, information gathered in the PREDART CRACK IT Challenge.

Where available, *in vitro* and *in vivo* phenotypic and omics data to support the PoCs.

Connection to networks of toxicologists and other expertise relevant to the Challenge.

# Examples of data sources

Currently available information on developmental biology processes in various species and chemical DART data is enormous but greatly scattered
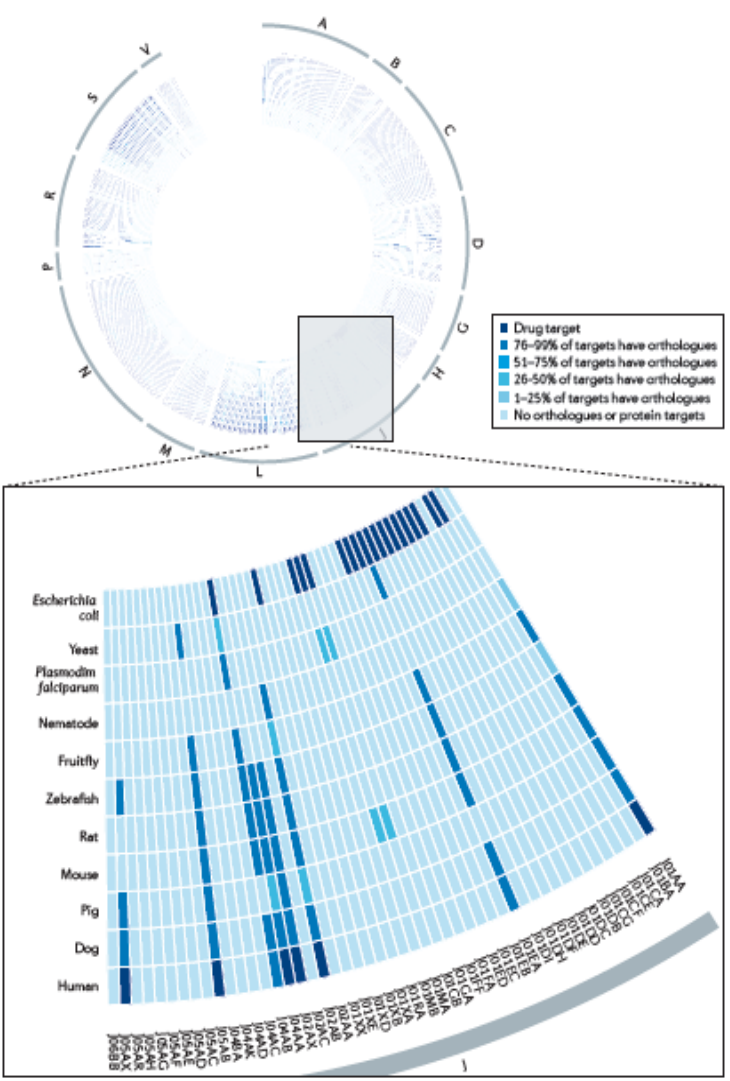
ToxCast, Connectivity Map, OECD QSAR Toolbox, DevTox (Germany), US EPA ToxRefDB, OECD eChemPortal, OECD AOP, ECHA DB, TOXNET, PUBMED, diXa & CTDBase (Toxicogenomics), NC3Rs PREDART CRACK IT, industry data

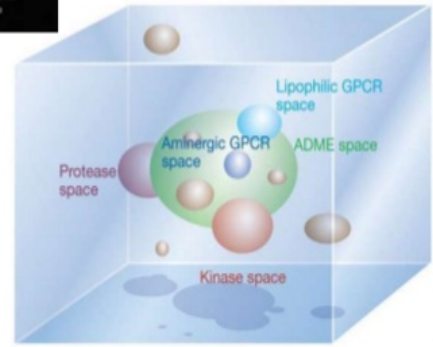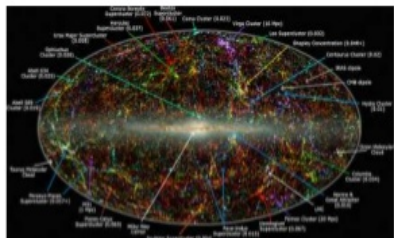This work should be complimentary to other working the field, e.g. http://www.ecetoc.org/publication/special-report-no-19-building-prenatal-developmental-toxicity-ontology/.

There may also be insufficient mechanistic data in some chemical classes so the generation of this lacking information from *in vitro* cell models and alternative non-mammalian model organisms may be needed as part of this project.

# Just a simple example of a usable visual: for DARTpaths by analogy a list of DART targets and chemical space



Legend:
- Drug target
- 76–99% of targets have orthologues
- 51–75% of targets have orthologues
- 26–50% of targets have orthologues
- 1–25% of targets have orthologues
- No orthologues or protein targets

Species (rows):
Escherichia coli
Yeast
Plasmodium falciparum
Nematode
Fruitfly
Zebrafish
Rat
Mouse
Pig
Dog
Human

Chemical space - $10^{60}$



Lipophilic GPCR space
Aminergic GPCR space
ADME space
Protease space
Kinase space

https://www.slideshare.net/valerytkachenko16/ompol-visualisation-of-large-chemical-spaces

Santos *et al.*, A comprehensive map of molecular drug targets.
*Nature Rev Drug Discovery*, 2017, **16**, 19-34

NC 3Rs

# Thank you for listening

The Sponsors are happy to discuss the challenge and potential applications with people in the run up to the submission deadline

Contacts are:

**Shell**
- Dr Martjin Rooseboom          Martijn.Rooseboom@shell.com

**Syngenta**
- Dr Richard Currie          richard.currie@syngenta.com