Comparison of cross-species homologous drug targets to support Target safety Assessments

N. J Coltman[1], F. Tenant[1], R. Roberts[1,2] and J. Sidaway[1]

[1]ApconiX Ltd., Alderley Edge, United Kingdom and [2]University of Birmingham, United Kingdom

Introduction

- Safety issues associated with drug targets significantly contribute to drug discovery project failures and hence conducting target safety assessments early in this process is important for project success.
- Selection of the correct animal species for preclinical toxicity assessment is crucial for translating toxicological risks to humans. However, not all drug discovery targets are homologues between typical preclinical animal species.
- The aims of this work were:
  - Identify homologous gene targets shared between humans, mouse, rat, dog and cynomolgus monkey.
  - Extract human druggable targets from the Open Targets platform with either small molecule clinical or discovery precedence, and curate a dataset of homologous 'druggable' gene targets between selected preclinical animal species.
- This proposed workflow could be used to support target safety assessment to evaluate the expression of drug targets within preclinical species for translational toxicology predictions and identify potential species of proposed drug targets.

Methods

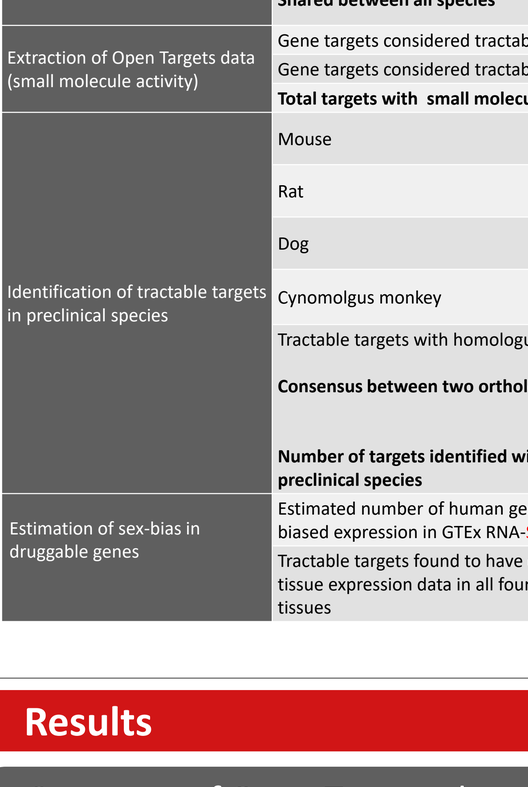
- We systemically identified homologous gene targets shared between humans, and four common preclinical toxicology species using a Biomart-Ensembl pipeline and the Orthologous Matrix (OMA) tool to derive a consensus set of orthologous targets tractable between preclinical toxicology species
- R' scripture was used to extract human target genes from the Open Targets Platform with either clinical or discovery precedence for small molecule druggability which were filtered against the identified preclinical species orthologues
- BASH and R' languages were used to compile RNA-Seq expression data from across 12 different tissues from the Human Genotype-Tissue Expression dataset (GTEx) and matched datasets found in the literature from the selected preclinical species, using expression data derived from Naqvi et al. (2018)
- RNA-Seq expression datasets were filtered against the Open Targets dataset to identify tractable small molecule drug targets to identify missing orthologues

Table 1. A summary of the key data extracted and analysed in bioinformatics pipeline.

Category	Subcategory	Number of genes
Cross species homologous genes	Mouse-Human orthologues	19,214, Biomart; 16,110, OMA
	Rat-Human orthologues	17,320, Biomart; 15,415, OMA
	Dog-Human orthologues	16,305, Biomart; 15,278, OMA
	Cynomolgus monkey-Human orthologues	18,887, Biomart; 16,340, OMA
	Shared between all species	**13,096, OMA; 12,950, Biomart**
Extraction of Open Targets data (small molecule activity)	Gene targets considered tractable from a clinical perspective	1049
	Gene targets considered tractable from a discovery perspective	3361
	Total targets with small molecule 'druggability' precedence	**4410**
Identification of tractable targets in preclinical species	Mouse	4234, Biomart; 4180, OMA
	Rat	4182, Biomart; 4104, OMA
	Dog	4097, Biomart; 4024, OMA
	Cynomolgus monkey	3999, Biomart; 4124, OMA
	Tractable targets with homologues observed in all preclinical species	3634, Biomart; 3688, OMA)
	Consensus between two orthology prediction methods	**4160 tractable genes with homologues in all four species**
	Number of targets identified with an ortholog absent in at least one of the preclinical species	**194**
Estimation of sex-bias in druggable genes	Estimated number of human genes identified by Oliva et al. (2020) to have sex-biased expression in GTEx RNA-Seq human tissue datasets	13,294
	Tractable targets found to have conserved sex-bias in Naqvi et al. (2018) RNA-seq tissue expression data in all four preclinical species in at least 1/12 mammalian tissues	2259

Results

Summary of Open Targets dataset



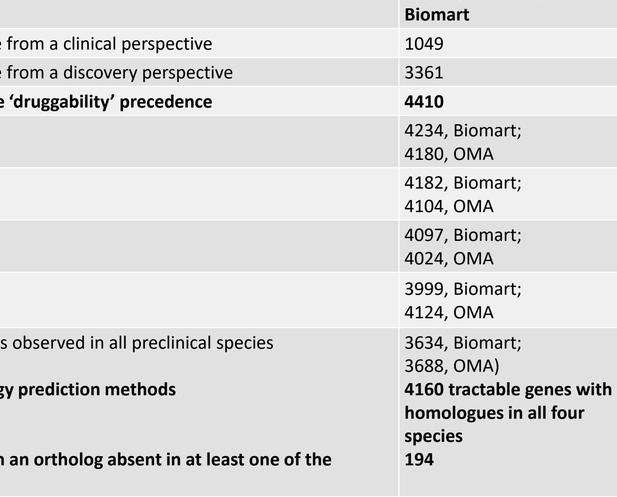


Figure 1. Comparison and overview of human genes with small molecule tractability status in Open Targets dataset. were (A) Overview of Open Targets, drug target tractability dataset used in analyses. (B) Consensus between Biomart and OMA pipelines for the distribution of orthologous tractable drug targets in preclinical species, present in the Open Targets dataset.

Absent orthologous genes with small molecule tractability between species

IDG family	Target symbol and name	M	R	D	C	Example drugs
Enzyme	CETP: Cholesteryl ester transfer protein	✗	✗	✓	✓	ANACETRAPIB, EVACETRAPIB, DALCETRAPIB, OBICETRAPIB, ROCACETRAPIB, TORCETRAPIB
	PARP10: Protein mono-ADP-ribosyltransferase	✓	✓	✗	✓	*
	PARP15: Protein mono-ADP-ribosyltransferase	✗	✗	✓	✓	*
	PLA2G2A: Phospholipase A2, membrane associated	✓	✓	✗	✓	VARESPLADIB
	PLA2G2E: Group IIE secretory phospholipase A2	✓	✗	✓	✓	*
GPCR	CXCR2: C-X-C chemokine receptor type 2	✗	✗	✓	✓	LADARIXIN, REPARIXIN, NAVARIXIN, ELUBRIXIN, DANIRIXIN, SX-682
	GPR52: G-protein coupled receptor 52	✓	✓	✓	✗	*
	HTR1E: 5-hydroxytryptamine receptor 1E	✗	✗	✓	✓	DEXFENFLURAMINE , ZIMELDINE , AMISULPRIDE
	P2RY11: P2Y purinoceptor 11	✗	✗	✓	✓	*
	P2RY4: P2Y purinoceptor 4	✓	✓	✗	✓	DIQUAFOSOL TETRASODIUM
	TAS2R5: Taste receptor type 2 member 5	✗	✗	✓	✓	1,10-Phenanthroline-5,6-dione, 4,7-dimethyl-1,10-phenanthroline
Ion channel	AQP10: Aquaporin-10	✗	✓	✓	✓	*
	HTR3C: 5-hydroxytryptamine receptor 3C	✗	✗	✓	✓	DEXFENFLURAMINE , TEDATIOXETINE, ZIMELDINE , AMISULPRIDE, CLOTHIAPINE, RAMOSETRON , RG3487, TROPISETRON
	HTR3D: 5-hydroxytryptamine receptor 3D	✗	✗	✓	✓	
	HTR3E: 5-hydroxytryptamine receptor 3E	✗	✗	✓	✓	
	KCNA4: Potassium voltage-gated channel subfamily A member 4	✓	✓	✓	✗	
	KCNC3: Potassium voltage-gated channel subfamily C member 3	✓	✓	✓	✗	GUANIDINE , NERISPIRDINE, TEDISAMIL, DALFAMPRIDINE, AMIFAMPRIDINE PHOSPHATE,
	KCNG2: Potassium voltage-gated channel subfamily G member 2	✓	✓	✗	✓	
Transporter	SLC22A11: Solute carrier family 22 member 11	✗	✗	✓	✓	PROBENECID
	SLC47A2: Multidrug and toxin extrusion protein 2	✗	✗	✓	✓	*
Other	BCL2L10: Bcl-2-like protein 10	✓	✓	✗	✓	OBATOCLAX MESYLATE
	BCL2L2: Bcl-2-like protein 2	✓	✓	✓	✗	NAVITOCLAX
	CXCL8: Interleukin-8	✗	✗	✓	✓	*; ABX-IL8#, HUMAX-IL8#
	GYPA: Glycophorin-A	✓	✗	✓	✓	*
	HLA-B: HLA class I histocompatibility antigen B	✗	✗	✗	✗	*

**Presence of binding pocket, Med-High ligand binding predicted*

 # Large molecule

Figure 2. Cross species comparison of absent orthologous genes with small molecule tractability in the Open Targets dataset. M = mouse, R = rat, D = dog, C = cynomolgus monkey.

Expression bias of druggable target genes in preclinical species

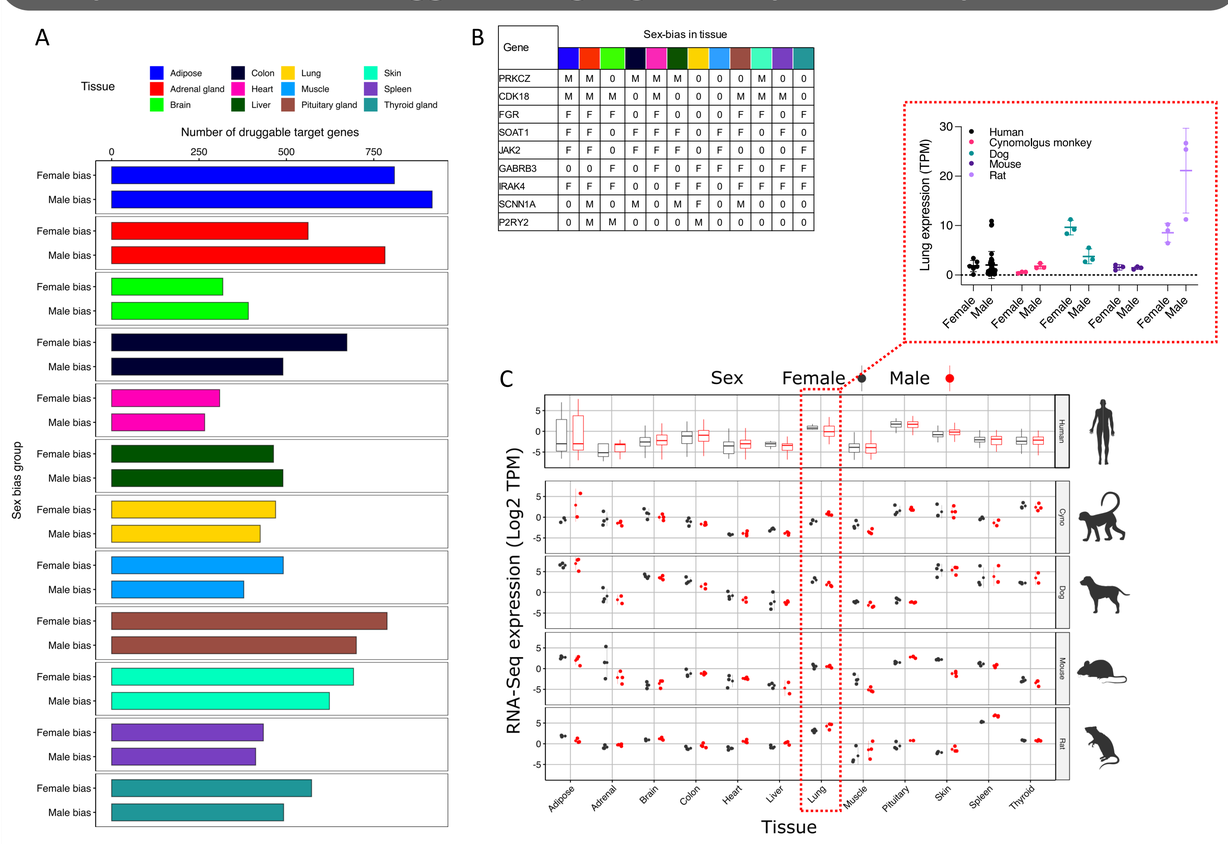


Figure 4. Analysis of druggable target genes with conserved sex-bias and species differences at a tissue-specific level according to the Naqvi *et al.* dataset. (A) Number of target genes with estimated conserved sex-bias in 12 mammalian tissues. (B) Examples of genes with conserved sex-bias in tissues (M = male, F = female). (C) Cross-species comparison of RNA-seq tissue expression of ALOX15, a gene with predicted species and sex differences across preclinical species, with inset expression profile in the lung.

Conclusion

- Progressing novel therapeutics through preclinical studies to clinical trials is a challenge which can be undermined when direct translation of target biology is inadequate between preclinical species and humans.
- Understanding target tractability and expression in preclinical species is a key component of the Target Safety Assessments and helps guides guide selection of the optimal species for the toxicology program.
- Our workflow has shown that most small molecule druggable targets are found to have at least one orthologue in the four of the most routinely deployed preclinical toxicology species.
- Crucially, up to 5 % of currently identified druggable targets in the Open Targets platform, may not be tractable in all four preclinical species which may pose issues with translation of proposed biology and safety.
- This proposed workflow could be used to support target safety assessment to evaluate the expression of drug targets within preclinical species for translational toxicology predictions and identify potential species of proposed drug targets.

References and datasets:

Ensembl-BioMart - Cunningham, F, et al. 2022. 'Ensembl 2022'.: https://doi.org/10.1093/nar/gkab1049. *Version 106, Apr 2022 was used.*

Naqvi, S et al. 2019. 'Conservation, Acquisition, and Functional Impact of Sex-Biased Gene Expression in Mammals': https://doi.org/10.1126/science.aaw7317. *Datasets fetched from GSE125483 on 4 Feb 2022*

Oliva, M, et al. 2020. 'The Impact of Sex on Gene Expression across Human Tissues': https://doi.org/10.1126/science.aba3066.

OMA - Altenhoff, A, et al. 2021. 'OMA Orthology in 2021': https://doi.org/10.1093/nar/gkaa1007. *Apr 2021 release was used.*

Open Targets - Brown, K, et al. 2018. 'Approaches to Target Tractability Assessment – a Practical Perspective': https://doi.org/10.1039/C7MD00633K. *Dataset was downloaded on 1 Apr 2022.*

'The Genotype-Tissue Expression (GTEx) Project'. 2013: https://doi.org/10.1038/ng.2653. *V8 dataset was downloaded 29 Jan 2022*