

## REVIEW

# Modelling human genetic disorders in *Xenopus tropicalis*

Helen Rankin Willsey<sup>1,2,\*</sup>, Eleanor G. Seaby<sup>3</sup>, Annie Godwin<sup>4</sup>, Sarah Ennis<sup>3</sup>, Matthew Guille<sup>4</sup> and Robert M. Grainger<sup>5</sup>

## ABSTRACT

Recent progress in human disease genetics is leading to rapid advances in understanding pathobiological mechanisms. However, the sheer number of risk-conveying genetic variants being identified demands *in vivo* model systems that are amenable to functional analyses at scale. Here we provide a practical guide for using the diploid frog species *Xenopus tropicalis* to study many genes and variants to uncover conserved mechanisms of pathobiology relevant to human disease. We discuss key considerations in modelling human genetic disorders: genetic architecture, conservation, phenotyping strategy and rigour, as well as more complex topics, such as penetrance, expressivity, sex differences and current challenges in the field. As the patient-driven gene discovery field expands significantly, the cost-effective, rapid and higher throughput nature of *Xenopus* make it an essential member of the model organism armamentarium for understanding gene function in development and in relation to disease.

**KEY WORDS:** Disease, Genetics, CRISPR, *Xenopus tropicalis*

## Introduction

Immense progress in exome- and genome-wide DNA analyses, coupled with the generation of large-scale genetic and phenotypic databases, has led to the curation of long lists of high-confidence disease-causing genes for a multitude of human genetic diseases affecting all tissues (Bamshad et al., 2019). However, the molecular, cellular, developmental and homeostatic functions of most of these genes are either unknown or poorly understood. Adding to this complexity, most genes are pleiotropic, so even when the function of a gene in one tissue at one developmental time is known, this knowledge may not translate to other tissues, times or contexts (Cummings et al., 2020; Panagiotakos and Pasca, 2022; Sestan and State, 2018; Willsey et al., 2018a, 2022b). This challenge of interpreting disease-risk gene lists can be readily met by using the high-throughput vertebrate model system *Xenopus tropicalis*, a diploid frog species (Bhattacharya et al., 2015; Blackburn and Miller, 2019; Deniz et al., 2018; Edwards et al., 2021; Exner and Willsey, 2021; Fainsod and Moody, 2022; Getwan and Lienkamp, 2017; Grainger, 2012; Griffin et al., 2020; Harland and

Grainger, 2011; Hwang et al., 2019; Nakayama et al., 2015; Nenni et al., 2019; Willsey et al., 2022a). Like its closely related but allotetraploid species *Xenopus laevis*, *X. tropicalis* (hereafter also referred to as *Xenopus*) is amenable to fast genetic analysis of multiple organ systems simultaneously, affordably, and at scale (Abu-Daya and Godwin, 2023). Here we describe the unique advantages of the powerful model organism *X. tropicalis* for the diagnosis of human disorders and as a tool for understanding the molecular mechanisms underlying them. We also discuss the crucial considerations that are needed to promote rigorous and informative work when using *X. tropicalis*.

## Experimental advantages of *Xenopus* for modelling human genetic disorders

*X. tropicalis* has unique experimental advantages that render it well-suited for modelling human genetic disorders (Exner and Willsey, 2021). First, *X. tropicalis* has a diploid genome that is highly conserved between frogs and humans (Hellsten et al., 2010); its high level of synteny makes the identification of orthologous genes more straightforward than in animal models with less conservation and/or a duplicated genome. Second, the model organism database Xenbase (<https://www.xenbase.org>; RRID:SCR\_003280) provides user-friendly access to the accurate, annotated reference genome, with excellent tools to facilitate genetic analysis and interpretation (Fisher et al., 2023; Fortriede et al., 2020). Additionally, methods for maintaining a thriving laboratory colony are well-established and costs significantly less than maintaining rodents (Nakayama and Grainger, 2023; Shaidani et al., 2020). *X. tropicalis* can be induced to mate year-round, with a single pair producing 4000+ embryos in a day by natural mating or by *in vitro* fertilisation (Lane and Khokha, 2022a,b). Embryos develop quickly; by day 4 they have developed organ systems, such as the central and peripheral nervous system, sensory organs, kidneys, skeletal muscle and cardiovascular systems, but not lungs or limbs. Within 10 days they exhibit robust quantitative behaviours that can serve as phenotypic readouts (Ismail et al., 2022).

Genetic perturbation protocols for *X. tropicalis*, including CRISPR/Cas9 mutagenesis (Bhattacharya et al., 2015; Nakayama et al., 2013), are well-established and cost-effective. Conveniently, mutagenesis of one of the two cells at 2-cell stage embryo yields a unilateral mutant, with one half of the animal carrying a homozygous mutation of interest while the other half serves as a within-animal control (DeLay et al., 2018; Macken et al., 2021; Willsey et al., 2018b). This is unique to *Xenopus* and makes generating thousands of mutant embryos per day feasible, thereby enabling truly parallelized analysis of dozens of risk genes by using a within-animal control. Indeed, this CRISPR-based strategy has a proven track record of exceptional success in studying autism spectrum disorders (ASD) risk genes (Willsey et al., 2021; Rosenthal et al., 2021), congenital anomalies (Kariminejad et al., 2019; Barbosa et al., 2020; Macken et al., 2021) cancer (Dimitrakopoulou et al., 2019), congenital heart disease (Deniz et al., 2018, 2023), kidney disease (DeLay et al., 2018; Getwan and Lienkamp, 2017) and many more (Edwards et al., 2021; Grand et al.,

<sup>1</sup>Department of Psychiatry and Behavioral Sciences, Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA 94158, USA. <sup>2</sup>Chan Zuckerberg Biohub - San Francisco, San Francisco, CA 94518, USA. <sup>3</sup>Genomic Informatics Group, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK. <sup>4</sup>European Xenopus Resource Centre (EXRC), School of Biological Sciences, University of Portsmouth, Portsmouth PO1 2DY, UK. <sup>5</sup>Department of Biology, University of Virginia, Charlottesville, VA 22904, USA.

\*Author for correspondence (helen.willsey@ucsf.edu)

 H.R.W., 0000-0001-8404-3291; E.G.S., 0000-0002-6814-8648; A.G., 0000-0001-5264-7824; S.E., 0000-0003-2648-0869; M.G., 0000-0001-6865-4519; R.M.G., 0000-0003-4539-5280

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

2023; Sega et al., 2019). Similarly, the developmental fate map of the early embryo has been described in detail (Nakamura and Kishiyama, 1971; Hirose and Jacobson, 1979; Jacobson and Hirose, 1981; Dale and Slack, 1987; Moody, 1987a,b; Masho, 1988; Masho and Kubota, 1986; Moody and Kline, 1990), allowing researchers to target genetic perturbations to specific tissues without using complex and time-consuming genetic methods.

Many transgenic and mutant lines are readily available, and there are protocols for generating new ones quickly (Aslan et al., 2017; Bhattacharya et al., 2015; Noble et al., 2022; Ogino et al., 2006; Shibata et al., 2022). Conveniently, The National *Xenopus* Resource (NXR; RRID: RRID:SCR\_013731) and the European *Xenopus* Resource Centre (EXRC; RRID: SCR\_007164) offer services for generating germline mutant lines for the community. Transgenic lines are proven tools for isolating tissues for single-cell analyses (Aztekin et al., 2019; Kakebeen et al., 2020), and protocols for isolating nuclei from mutant lines allow investigators to efficiently perform single-cell gene expression analyses comparing mutant and wild-type embryos (Nakayama et al., 2022a). The generation of mutant lines has recently been expedited by using gynogenesis protocols, in which the paternal genome is not inherited and researchers are, therefore, able to skip a generation in making homozygous mutants (Nakayama et al., 2022b). Similarly, a long history of genetic tool development and ease of phenotyping exists for *Xenopus*. Below, we expand on these and their most appropriate applications in this context. Another powerful feature of frog embryos and tadpoles is that they absorb small molecules from their surrounding culture medium, facilitating large-scale drug screening (Gull et al., 2023; Wheeler and Liu, 2012). In the search for targeted therapeutics for rare genetic disorders, this advantage is paramount.

Moreover, studying human disease genes in *Xenopus* has a history of producing phenotypes that often more closely recapitulate the human condition than even rodent models. For example, mutations in the essential eye transcription factor gene *pax6* in *Xenopus* result in a phenotype very similar to that of patients with the rare disease congenital aniridia (Nakayama et al., 2015), while mutations of this gene in mouse lead to a dissimilar ‘small-eye’ phenotype (Hill et al., 1991). In another example, genetic variants in the scaffold protein-encoding gene *USH1C* cause the rare disorder Usher syndrome 1C, and lead to both deafness and blindness in patients. Both the characteristic human eye and ear abnormalities are recapitulated in the *ush1c*-mutant of *X. tropicalis*, but the structure responsible for the eye deficit is absent in rodent photoreceptor cells and, thus, these animals are not optimal models for Usher syndrome 1C (Schietroma et al., 2017). The key point here is that mammalian models, despite their relative evolutionary proximity, do not always mimic all the features of a human disorder, and that non-mammalian models, like *Xenopus*, may offer advantages – either in phenotype as noted here or even when the *Xenopus* phenotype is not identical to that shown by humans. This is because the frog has advantages regarding the unique experimental strategies that are possible in this system. Finally, with respect to the responsible use of animals in research – i.e. the 3Rs: Replacement, Reduction and Refinement of animal experimentation, initially formulated by William Russell and Rex Burch in 1959 (see Hubrecht and Carter, 2019) – using *Xenopus* as a model rather than mammals is also attractive due to the absence of cerebral and limbic structures responsible for the translation of pain to distress, which are present in humans.

#### Getting started – gene selection considerations

There are many routes to begin functional genetics work in *Xenopus*. For example, one group may begin with a single patient

who presents with several potential disease-causing variants, and need to first prioritise variants for functional investigation by using *Xenopus*. Alternatively, another group may have a long-standing interest in a particular disease and want to model previously unseen variants observed in known disease-causing genes. In another scenario, a group may begin with an organ of interest and want to identify disease-causing genes that affect this organ. Finally, as the number of variants of uncertain significance (VUS) from DNA-sequencing projects continues to increase ever more rapidly, researchers may work in an integrated manner with clinical geneticists and bioinformaticians to test variant-disease links in *Xenopus*. The diverse array of starting points is a testament to the versatility of *Xenopus* functional genomics. Here we will focus on the first scenario, which highlights many of the core considerations that are paramount to any study modelling human disease genetics.

When considering multiple potential disease-causing variants within one patient, a helpful strategy for prioritising variants for modelling in *Xenopus* is to determine whether any loss-of-function (LOF) – i.e. frameshift, essential splice site or nonsense – variants in that gene ever occur in unaffected subjects. If LOF variants of a gene occur in unaffected controls, the gene in question is considered ‘LOF tolerant’ and not to be under evolutionary constraint, so is less likely to be disease causing. To quantify LOF constraint, there is a metric called the LOF observed/expected upper-bound fraction (LOEUF) score (Karczewski et al., 2020). Genes with a low LOEUF score are intolerant to gene perturbation and, thus, provide clues to biological essentiality (Karczewski et al., 2020). Indeed, genes with the lowest LOEUF scores are those most enriched for known disease genes and represent targets for novel disease gene associations (Seaby et al., 2022). Therefore, LOF constraint can help prioritise which genes to study in model systems. LOEUF scores can be found by searching the Genome Aggregation Database (gnomAD), which to date represents the largest population dataset globally from more than 195,000 individuals, curated from over 60 studies (<https://gnomad.broadinstitute.org/>) (Gudmundsson et al., 2022; Karczewski et al., 2020) (Table 1). Importantly, this database is generated from unaffected individuals and, therefore, rare paediatric-onset diseases are underrepresented in it, meaning that monogenic disease-causing variants, many of which will be of interest to *Xenopus* researchers, are largely absent from this database. Rather, this database is a helpful resource for ascertaining the baseline variation that occurs in a gene of interest in an unaffected population. An abundance of LOF variants within a gene of interest in an unaffected population argues against its relevance to disease.

Another consideration when prioritising potentially pathogenic genetic variants for modelling is whether these have already been noticed in patient populations. Databases where this can be explored online exist (Table 1), and they often include the genomic location, gene/protein variant, higher level phenotyping terms, clinical significance and, occasionally, additional incidental findings. One of the most-recognised genetic disease databases is the Online Mendelian Inheritance in Man (OMIM) database (<https://www.omim.org/>), a manually curated open-source website that describes human genes, genetic disorders, inheritance patterns and genotype/phenotype relationships (Table 1) (Hamosh et al., 2005). Another widely used human disease dataset is the DECIPHER (<https://www.deciphergenomics.org/>) online repository of genetic variation with associated phenotypes for >45,000 patients with rare disorders (Table 1) (Firth et al., 2009). The freely available DECIPHER browser enables the user to search genes, regions and phenotypes for patients with matching variants. These databases can also be especially helpful if an investigator is starting with a disorder of

**Table 1. Examples of patient registries and tissue biobanks.**

Registry	Type of data	Availability	Data access	Additional data access
Genome Aggregation Database (gnomAD)	Population genotype data	Open access	<a href="https://gnomad.broadinstitute.org">gnomad.broadinstitute.org</a>	Individual level of data can be applied for through dbGaP <sup>1</sup>
Genotype-Tissue Expression (GTEx) portal	Tissue- and cell-specific gene expression data	Open access	<a href="https://gtexportal.org/home">gtexportal.org/home</a>	Individual level of data can be applied for through dbGaP
Genomics Research to Elucidate the Genetics of Rare diseases (GREGoR) Consortium	Genotype and health data	Requires consortium membership	<a href="https://gregorconsortium.org/data">gregorconsortium.org/data</a>	Some data can be applied for without consortium membership through dbGaP
Genomics England	Genotype and health data	Requires data access approval	<a href="https://www.genomicsengland.co.uk">www.genomicsengland.co.uk</a>	Clinicians can be contacted using a form within the secure Genomics England Research Environment
DECIPHER	Individual pheno- and genotypic data from patient cohorts with neurodevelopmental disorders.	Open access	<a href="https://deciphergenomics.org">deciphergenomics.org</a>	Clinicians can be contacted through an online form on the website
UKBioBank	Biobank data	Requires payment and data-access approval	<a href="https://ukbiobank.ac.uk">ukbiobank.ac.uk</a>	N/A
All of Us	Biobank data	Open access to summary data only	<a href="https://allofus.nih.gov">allofus.nih.gov</a>	Researchers can apply for access to individual data levels
FinnGen (consortium of Finnish public biobanks)	Biobank data	Open access to summary data only	<a href="https://finngen.fi/fi">finngen.fi/fi</a>	Individual level of data available to consortium partners
European Genome-phenome Archive (EGA)	Genetic, phenotypic, and clinical data	Each study requires approval to access	<a href="https://ega-archive.org">ega-archive.org</a>	N/A
ClinVar	Genotype data	Open access	<a href="https://ncbi.nlm.nih.gov/clinvar">ncbi.nlm.nih.gov/clinvar</a>	N/A
Human Gene Mutation Database (HGMD)	Genotype data	Open access to limited datasets following registration	<a href="https://hgmd.cf.ac.uk/ac/index.php">hgmd.cf.ac.uk/ac/index.php</a>	HGMD professional is available commercially

<sup>1</sup>dbGaP, database of Genotypes and Phenotypes

interest and looking for disease-causing genes to study. To preserve patient anonymity, only high-level phenotypes (e.g. abnormality of the cardiovascular system, etc.) are displayed. However, if more detailed clinical information is required, one can contact the clinician or principal investigator of the participating centre where that patient was recruited. This can facilitate access to valuable additional clinical information, such as when the patient was last assessed, updates on disease progression and access to information on co-morbidities or additional damaging variants.

Another very helpful database is ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar>; Landrum et al., 2014), an open-access repository for the standardised classification of variants (Table 1) (Landrum et al., 2014). Anyone using the platform can submit a variant found in a patient sample and assert its clinical significance by applying a standardised set of guidelines used for diagnostic reporting as set out by the American College of Medical Genetics (ACMG) and/or provide supporting data (Richards et al., 2015). Therefore, the database can be queried for variants previously described and validated as pathogenic by ACMG guidelines. ClinVar benefits from an easy-to-use graphical interface enabling researchers to see where the different classes of variants lie across the gene, and links to dbVar, which holds chromosomal information, and to LitVarz, which allows interrogation of the literature based on specific variants and genes.

Once a potentially disease-causing gene and/or variant has been identified, an important strategy when deciding whether they could be productively modelled in *Xenopus* is to align the protein sequences from *X. tropicalis* and human by using the NCBI tool HomoloGene (<https://www.ncbi.nlm.nih.gov/homologene>). This can help determine conservation at protein level – which, in turn,

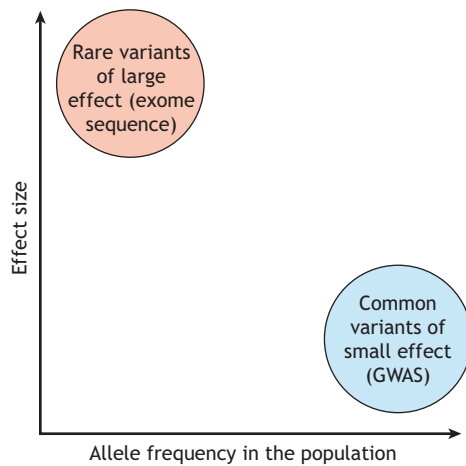
may help selecting reagents, such as antibodies – and whether generating mutant lines with patient-derived variants is possible. Xenbase should also be consulted to assess genomic synteny and to determine which isoforms have been described in *Xenopus*, and when and where in development the gene is expressed (Bowes et al., 2010; Fortriede et al., 2020). This early effort can also preempt any potential issues with genome annotation. Finally, other databases, e.g. ModelMatcher (<https://www.modelmatcher.net/>) and those specific for model organisms, should be checked for existing but unpublished models of the disease gene, either to inform experimental design or to prevent unintentional duplication of effort.

#### Designing strategies that reflect the genetic architecture of the disorder

Once a gene(s) of interest has been selected, the next major consideration is the nature of the patient alleles and the genetic nature of the disorder of interest. For example, one would model a single-gene recessive disorder (Appendix G, Single-Gene Disorders, 2010), such as cystic fibrosis or phenylketonuria, very differently from a complex heterogenous genetic disorder, like autism, or a disorder in which the variants are gain-of-function (GoF), like Noonan syndrome (Roberts et al., 2007). Therefore, familiarity with common genetics terminology and concepts is essential. Here we briefly explore the most critical ones encountered in patient variant modelling.

One essential concept is that of variant effect size and it is critical for modelling genetic disorders in *Xenopus*. Common variants, where population allele frequency exceeds 1%, as identified in genome-wide association studies (GWAS), carry very small individual effect sizes. Any individual variant only increases risk very slightly (Fig. 1) but,





**Fig. 1. Relationship between rare and common variants with respect to allele frequency and effect size.** Rare variants are typically detected in specific, clinically indicated patient trio (i.e. involving testing of the affected individual and both biological parents) studies of exome sequencing, and enriched for *de novo* variants with individually large effects. When perturbed, the genes that carry these variants often produce phenotypes in model organisms. By contrast, common variants are usually identified in much broader genome-wide association studies (GWAS) of large unrelated cohorts. Individually, these common variants carry very small risk and are unlikely to produce a measurable phenotype when recapitulated in model organisms.

because they are common, they carry population-level risk. By contrast, rare variants with population allele frequencies of <1%, as typically identified in patient trio exome sequencing, carry large individual risks (Fig. 1). This way, it is much more likely to observe a phenotype in *Xenopus* when recapitulating rare variants from patient sequencing data than common ones. Common variants, often referred to as single nucleotide polymorphisms (SNPs), are also often in intergenic regions and it can be difficult to reliably assign the affected gene (reviewed by Boyling et al. 2022). Rare variants from patient exome sequencing studies are enriched for protein-coding variants that often directly alter protein structure, while common variants often reside in non-coding regions that can be more difficult to interpret (reviewed by Allou and Mundlos, 2023). This makes patient-derived rare variants more straightforward to model and generally more suited to model organism studies, particularly when LOF variants in the same gene are reported in multiple unrelated affected individuals, since this strengthens the likelihood that variants in that gene are pathogenic.

Copy-number variants (CNVs) often have intermediate effect sizes and allele frequencies (Willsey et al., 2022b), and can be modelled in *Xenopus*, with some caveats. These variants often span large regions of the human genome, so assessing synteny and conservation between *Xenopus* and human is paramount. This is particularly important as intergenic regions usually do have larger evolutionary differences than coding regions. Nevertheless, to model CNVs with CRISPR, one can design two single guide RNAs (gRNAs) that flank the mapped orthologous CNV breakpoints and generate large deletions. One can also inject larger DNA constructs to model duplications. *Xenopus* embryos will tolerate exogenous DNA fairly well, though the inheritance of the construct can be rather mosaic (Cleaver and Krieg, 1999), so it is recommended to include a visible marker within the construct to track expression patterns, or to breed the animals, which is time consuming.

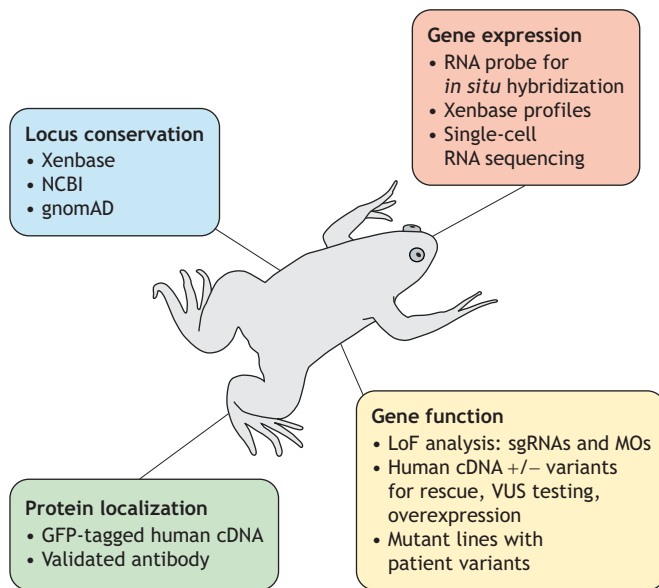
Where LOF variants are implicated in disorder pathobiology, simply perturbing a single gene is the simplest way to model using

CRISPR-based genome editing. One can design an sgRNA that directs Cas9 to cut near the 5' end of the coding sequence to empirically test editing efficiency, and choose sgRNAs of high efficiency and specificity, with the assumption that the cuts will be repaired by non-homologous end-joining, thus, most likely generating deletions and, therefore, frameshift mutations that will result in a non-functional gene product (Willsey et al., 2022a). Unlike LOF variants, the impact of missense variants can be harder to interpret, although there are algorithms that will predict how deleterious a given variant would be to protein structure and, therefore, function. Missense badness, PolyPhen-2 and missense Constraint (collectively forming the MPC) score are some examples of these algorithms (Samocha et al., 2017 preprint). Mimicking missense variants will have a very high impact on the field, since a large proportion of disease variants are missense by nature. Base editing, which changes a single base without cutting the DNA backbone and is, therefore, perfectly suited to missense modelling (Komor et al., 2016), has been shown to work in *Xenopus* (Shi et al., 2019; Park et al., 2017; Naert et al., 2020 and 2021b). *Xenopus* are particularly suited for this type of modelling due to the ease of generating many mutant lines. There is also the potential for doing this work in the generation F<sub>0</sub>, given the high tolerance of injected ribonucleoprotein complexes in the frog.

When modelling a complex genetic disorder where many different genes carry risk, such as most psychiatric disorders, including autism and schizophrenia, there are advantages to working in *Xenopus*. In this case, most of the genes involved are highly pleiotropic, and it is difficult to separate the pleiotropy of a gene from the relevant function in the disorder just by studying individual genes (Willsey et al., 2022b). Instead, it is straightforward in *X. tropicalis* to study many genes and/or variants in parallel because of the speed, ease and cost-effectiveness of CRISPR mutagenesis. In one afternoon, thousands of mutant embryos can be created for many different genes. This can allow for the identification of gene functions and phenotypes that are common for different genes associated with a particular disease of unknown pathobiology. For example, CRISPR perturbations of any of the ten genes with the strongest statistical evidence for association with autism did, individually, cause neurogenesis defects in *Xenopus*, highlighting this process as one potentially relevant to the disorder (Willsey et al., 2021). This 'phenotypic convergence' can also combat the formidable problem of pleiotropy, as studying just one gene at a time may lead to focusing on a phenotype specific to that particular gene and not phenotypes more broadly relevant to other disease-risk genes (Willsey et al., 2021, 2022b). Similarly, performing *in situ* hybridizations (Willsey, 2021) or screening single-cell transcriptomic atlases (Briggs et al., 2018) for many risk genes in parallel can help limit the search space for phenotypes by identifying tissues, cell types and timepoints of co-expression.

#### Building gene-specific tools

Once a gene is selected for investigation, and direction of perturbation and expected effect size have been considered, it is advantageous to begin developing gene-specific tools (Fig. 2). Often, it can be extraordinarily helpful to determine the developmental expression patterns for a gene, as this may point to relevant tissues affected by genetic perturbation and its function. For example, exclusive expression in highly ciliated tissues may suggest a potential function in ciliary biology and, consequently, a role in ciliopathies. This is particularly useful when studying syndromic conditions that affect multiple tissues and require an RNA *in situ* hybridization probe which, in turn, require a DNA template. *Xenopus* plasmid libraries of such



**Fig. 2. Approaches and tools that facilitate human genetic disease modelling in *Xenopus*.** Once a gene of interest is selected, the locus conservation between *X. tropicalis* and humans can be explored in Xenbase, NCBI and gnomAD databases. The expression profile of the gene can be queried by using RNA *in situ* hybridization and on Xenbase. Then the protein can be located in cells and embryos through expression of a tagged human cDNA clone or validated antibody staining. Finally, the function of this gene can be inferred through loss-of-function (LoF) analysis, rescue experiments, overexpression and generating mutant lines with patient-derived variants. This kind of work can quickly point to mechanisms of disease, and tissues and structures for successful studies.

constructs exist, such as the ORFeome (Grant et al., 2015) and the I.M.A.G.E. database (Lennon et al., 1996). Many *Xenopus*-specific constructs, including those donated from laboratory collections, can currently be obtained from horizon (<https://horizondiscovery.com/en>) or the EXRC. If unavailable, plasmids can be generated by PCR from a cDNA library and simple cloning, which is particularly straightforward in *Xenopus* because of the high-quality accessible genome sequence and the ease of which large amounts of stage-specific cDNA can be obtained. It is important to remember that the composition of RNA isoforms may change during development and/or by cell type, and it is advisable to select or engineer the probe that matches the timepoints and tissues of interest, or the one that recognizes a coding sequence that is present in all isoforms.

For studies that require probing the product of the gene in question, investing in a reliable antibody against the *Xenopus* protein of interest can be exceptionally helpful for validating reagents, such as sgRNAs, by confirming that protein levels are altered, and for localising the protein in cells and tissues. However, all antibodies must be rigorously validated for specificity and avidity, ideally through confirming the antibody signal using immunohistochemistry or western blotting, in tissues that express the target and seeing it abolished in mutant tissues where the target is absent.

For LoF analyses, developing multiple orthogonal perturbation reagents is advisable. For CRISPR-based gene knockouts, sgRNAs can be easily designed through online algorithms (Moreno-Mateos et al., 2015; Labun et al., 2019). We also recommend that multiple non-overlapping target sequences are used, as discussed below in the rigour section. While these algorithms will predict cutting efficiencies, the actual efficiency for each sgRNA must be empirically determined in *Xenopus* embryos. One strategy is to use Sanger sequencing of

injected animals followed by sequence-trace deconvolution, and estimation of the insertion and deletion frequencies (Brinkman et al., 2014; Conant et al., 2022). Strategies also exist to maximise phenotypic penetrance of any gene edits in generation  $F_0$  to avoid in-frame mutations (Naert et al., 2020). A helpful orthogonal strategy to CRISPR-based gene knockout is to aim for post-transcriptional knockdown of the target. For this, the first step is to generate and validate a morpholino oligonucleotide, targeting either the start codon – thereby, blocking translation of mRNA, or a splice site – thereby, blocking RNA splicing (Malartre et al., 2006). Morpholinos can be helpful by strongly knocking down expression of the target gene, targeting specific isoforms or depleting maternal RNAs from the zygote. The latter is particularly useful for studying the LOF of maternally expressed mRNAs, since CRISPRs will only target the zygotic genome.

Finally, if the protein of interest is amenable to pharmacological modulation, e.g. by using an inhibitor to block a kinase, this can be a helpful additional path for orthogonal perturbation. Pharmacological inhibitors have the additional advantage of temporal control, as the drug can be added to the embryo culture medium at different times during development, separating pleiotropic roles or bypassing early lethality. Ideally, the targeting specificities of such reagents are controlled by using multiple compounds to target the same protein. For CRISPR experiments, another possible control is to target any potential off-target sites revealed by the algorithm used to choose the sgRNA, to ensure one does not observe the on-target phenotype.

Given the need for relevance to humans, modelling disease-causing variants extends beyond LOF. Expressing the human coding sequence can be very helpful for GoF, localization and rescue experiments, and/or for patient variant testing. It should be noted that transgenic expression of human genes risks disrupting protein-protein interactions by crossing species if these interactions are not evolutionarily conserved, as well as issues with inappropriate stoichiometries and potential toxicity due to high dosage.

### Designing phenotyping assays

Once a disease-causing gene has been selected and the perturbation strategy optimised, developing a phenotyping pipeline for *Xenopus* is straightforward since many tools have been developed and used for embryology studies in the last 60 years. Fortunately, robust transcriptomic, *in situ* hybridization and proteomics data are available on Xenbase to guide the focus of phenotyping assays. Referring to transcriptomic and proteomic data can also inform whether any maternal proteins persisting in the early embryo may mitigate phenotype presentation in early development. During phenotypic experiment design, it is useful to plan and conduct pilot studies to help estimate sample sizes, and to investigate the extent of variability in the phenotype – particularly when analysing mosaic generation  $F_0$  models. The first readout after a genetic perturbation is usually brightfield microscopy of the whole animal to identify gross changes and their relationship to regions of expression or clinical presentation in humans. Aside from being broadly accessible in every laboratory, whole-animal examination is important since disruption of gene function in distinct tissues and at different stages within the whole organism may generate complex phenotype alterations.

Once the affected organ(s) or cell type(s) have been identified, key tissue-level and molecular differences can be assessed in more detail in transgenic *Xenopus* lines or with techniques, such as *in situ* hybridisation and immunohistochemistry. For most human organs an orthologous *Xenopus* organ exists at tadpole stage (~6 days after fertilisation), and these organs can be studied by simple histological staining and microscopy techniques (Fig. 3). More subtle

morphological changes can be captured by state-of-the-art imaging techniques, including MicroCT, light sheet and optical coherence tomography (Date et al., 2019; Deniz et al., 2017; Macken et al., 2021; Naert et al., 2021a,b). Detailed image analysis capacity is being significantly improved by machine-learning algorithms capable of detecting differences between control and mutant samples (Naert et al., 2021a,b).

As advanced research elucidates the nature of more rare human genetic diseases, modelling studies will need to adopt additional sophisticated methods of phenotyping. For example, in response to the large number of neurodevelopmental disorders identified by diagnostic sequencing and analysis, quantitative behavioural assays are now being applied to *Xenopus* disease models. At present, these include analysis of locomotion and working memory (Ismail et al., 2022) as well as epilepsy (Panthi et al., 2023), and mainly use established video-imaging methods and analysis algorithms adapted for *Xenopus* (Ismail et al., 2022; Panthi et al., 2023). These assays have the potential to generate high-throughput data with which to study neurodevelopmental disorders and test repurposed licenced medications (see e.g. the target-agnostic drug discovery platform CogniXense, <https://wyss.harvard.edu/technology/cognixense-for-rare-disease-drug-discovery/>). Work of this nature is being complemented by the historic strengths of the *Xenopus* model, e.g. by using techniques to record neuronal plasticity and action potentials both within intact tadpoles, and in primary cultures of neuronal cells or neuronal/muscle co-cultures derived from embryos with gene variants (Sun and Poo, 1985; Gore et al., 2021; Hiramoto and Cline, 2021).

### Rigour

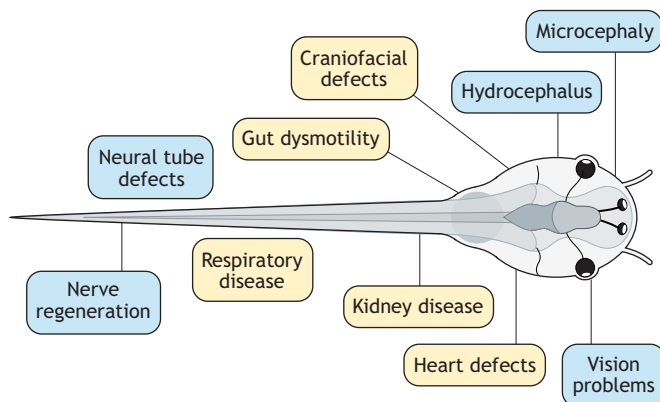
As discussed previously, CRISPR/Cas9-based editing has become a popular approach for the perturbation of gene function when modelling human genetic disease in *Xenopus*. For these experiments, the recommended controls vary greatly, from using non-coding or scrambled sgRNAs to the use of sgRNAs targeting positive-control genes, like pigmentation genes or disease genes in other tissues (Willsey et al., 2021). In addition, investigators also often use multiple sgRNAs targeted to the same gene, as we recommended above, and test whether the edits induced by individual sgRNAs phenocopy each other (Willsey et al., 2018b). The ultimate control, of course, is rescuing the phenotype by expressing the reference allele (Willsey et al., 2020). This rescue

approach also opens up the huge opportunity to test VUS and/or missense variants to infer causality of suspected alleles (Blackburn et al., 2019; Kostiuk and Khokha, 2021). The sequence of the expression construct can be modified with said variants, and if this modified allele is not able to rescue the phenotype, like the reference allele does, this indicates that the missense variant or VUS causes LOF. The rescue allele is most often cloned into an expression plasmid to make synthetic mRNA and is then microinjected. Other vectors include a bacterial artificial chromosome construct (Fish et al., 2014) or knock-in of the desired template into the genome (Nakayama et al., 2020), although these strategies can be technically challenging. Regardless of the rescue allele strategy, a lack of phenotypic rescue does not necessarily mean that the original perturbation was off-target or non-specific. There are many reasons why a rescue experiment may not work, including inadequate gene dosage, cell-type-specific gene expression or temporal dynamics.

Additional experimental controls include genetic perturbation of one cell of a 2-cell-stage embryo of generation F<sub>0</sub>, which allows within-animal comparisons and controls for inter-animal variation, especially when the perturbation is coupled with the expression of a fluorophore for tracing. Phenocopying a LOF perturbation with an additional technique, such as a pharmacological inhibitor or expression of a dominant-negative construct can increase confidence that the observed phenotype is caused by LOF of the gene in question. Generating a germline mutant line that recapitulates the F<sub>0</sub> phenotype can be very powerful confirmation, albeit much more time consuming. While more challenging experimentally, creating a germline benign variant in the gene of interest that does not produce the phenotype in question can be powerful evidence that mutagenesis itself did not create the phenotype. Also, whenever possible, the identity of control and test animals should be masked to the researcher performing the phenotyping to prevent bias. Finally, we advise being cautious when including detailed clinical phenotyping into the design phase of these experiments, since this can bias findings and may also lead to overlooking phenotypes in the frog that are either species-specific or have yet to be identified in the often very small number of human patients with a particular gene variant.

### *Xenopus* as a diagnostic platform

Whilst the ability to sequence DNA has advanced tremendously, the ability to interpret variants significantly lags behind. At present, ~70% of variants captured in the sequencing results of a patient are VUS, which may be in known or newly identified disease-causing genes. VUS are usually coding variants that lack sufficient evidence to determine whether they are pathogenic or benign. Whilst we briefly mentioned expressing VUS in *Xenopus* in the previous section, it is important to emphasise that testing the link between a VUS and disease is significantly different from the LOF studies described above. When attempting to model a VUS in a gene that has not been previously associated with a disease, a simple LOF strategy as outlined in detail above could be a useful first step, although LOF models are insufficient to underpin clinical intervention. However, if a patient-specific VUS is linked to a known disease-causing gene, only precise models of the variant have merit. Even when a novel variant is identified in a gene that, when inactivated, is entirely consistent with the phenotype of a patient, said variant will still be classified as a VUS and often not clinically reported unless functionally validated. Therefore, there is strong demand from genetic disease patients, families and clinicians for cost-effective and rapid *in vivo* models to functionally test these variants and improve diagnostic rates.



**Fig. 3. Phenotyping opportunities in *Xenopus*.** Studying how genetic perturbation disrupts organ function is essential to successful human disease modelling. Various methods are available in *Xenopus* and, due to the functional similarities, can be applied to characterising disorders within organs as diverse as lung, heart, brain, kidney and gut.



*Xenopus* provides an excellent model for this scenario, offering multiple possible approaches. First, the VUS can be recapitulated in the endogenous *Xenopus* gene if this is sufficiently conserved. Whilst this is straightforward for LOF frameshift, deletion, exon skipping and nonsense mutations, single-base changes, however, are more of a challenge and require the improvement of base-editing techniques in the frog (Shi et al., 2019; Park et al., 2017; Naert et al., 2021a,b). Heterozygous, non-mosaic, precise changes, including small insertions, have been achieved by engineering the oocyte genome (Aslan et al., 2017; Martin, 2023) or by injecting long single-stranded DNA templates for insertion into the frog genome via endogenous homology-directed repair (Nakayama et al., 2020, 2022c). These methodologies remain inefficient yet worthwhile.

An alternative approach that mitigates the inefficiencies of engineering precise gene edits is to first disrupt the endogenous *Xenopus* gene and then attempt to rescue its expression, as we discussed above. This can be achieved using a human cDNA expression construct that either carries the reference sequence or one with the VUS (Devotta et al., 2016; Griffin et al., 2018). If the reference sequence can rescue the phenotype but the VUS cannot, that provides excellent evidence that the VUS is deleterious to protein function. The largest issue with this approach is that, in some cases, the human reference sequence cannot rescue the frog phenotype for a variety of possible reasons, including dose, failure to interact properly with *Xenopus* proteins and cell type specificity (discussed by Rana et al., 2006). Alternatively, if the VUS but not the reference allele produces a LOF phenotype, this can be evidence of a dominant-negative mechanism of genetic perturbation. Finally, but perhaps less convincingly, overexpression of the reference sequence can be used to determine the GoF outcome and then this phenotype can be compared with that produced by expression of variants.

In terms of controls for diagnostic work, if the VUS in frog recapitulates the phenotype of the patient, then this already shows the result is consistent in two different species. Coupled with a second experiment that confirms the phenotype in a biochemical or cellular assay, two species sharing a VUS-phenotype link is strong evidence for specificity and relevance in experiments testing the link between a genetic variant and disease. Of course, there is also the potential for the frog to incompletely recapitulate the human presentation due to evolutionary differences. In this case, a complementary model, perhaps human iPSCs or a mammalian model, can be useful to fill in those gaps.

In addition to VUS classifications, *Xenopus* is an excellent platform for modelling diseases with known causal genetic variants but unknown molecular mechanisms of disease (Mascibroda et al., 2022; Wong et al., 2023; Marquez et al., 2021; Colleluori and Khokha, 2023; Sempou et al., 2022; Willsey et al., 2021; Rosenthal et al., 2021; Grand et al., 2023; Getwan et al., 2021; Kulkarni and Khokha, 2018; Boskovski et al., 2013). In this case, specific variants can be studied, and molecular mechanisms identified *in vivo*. As many disease-causing variants are missense and, potentially, not simply LOF, *Xenopus* is extremely valuable because it is cost-effective and simple to create many germline mutant lines to mimic such variants in their native locus, as long as the locus is well-conserved between humans and frogs. Such work identifying new disease mechanisms can open the door to biomarker and therapeutic development.

#### Working with clinicians and ethical considerations

When attempting to model VUS and potential new disease gene(s), there are clear benefits to having direct contact with the primary referring clinician or bioinformatician. Through direct collaborations

with physicians and clinical geneticists, researchers can ascertain the variant-calling methodology and the confidence in the *in silico* predictions of pathogenicity for each variant (Seaby and Ennis, 2020). One challenge associated with this collaboration is that patient genetic and phenotypic data are subject to extensive protections to ensure privacy. Whilst medical privacy regulations vary between countries, most databases generally restrict access to data that may potentially identify the patient. For example, when mining rare disease data from the Genomics England database, it is not possible to export any human phenotype ontology (HPO) terms (Robinson et al., 2008) or a combination of HPO terms that occurs in fewer than five individuals, regardless of genotype data. In the USA, access to clinical data for research purposes often requires Institutional Review Board approval and a data transfer agreement. Such restrictions, even for de-identified and/or fully anonymised data, frequently vary between repositories (Table 1) and researchers need to learn to navigate them prior to planning to model the variants in *Xenopus* and other organisms.

An additional consideration is that there are significant differences in the way that *Xenopus* researchers and clinical geneticists describe changes to the genome. A small, but crucial difference that has reduced cross-disciplinary fluency in the past is that clinical or human geneticists consider the first nucleotide of the translation start codon in the reference sequence as nucleotide 1 of a coding DNA (den Dunnen et al., 2016), while animal researchers normally use the start of the transcript in the annotated genome as nucleotide 1. Similarly, many datasets and clinical reports do not include information on the zygosity of the variant. The short but excellent guide on how variants are annotated in human genetics is thus essential reading (Ogino et al., 2007). Less of a systematic challenge is that the workload of clinicians may not give them time to record the complete phenotypic information about a patient. For this reason, it is often desirable to work with clinicians who are based locally to the model organism laboratory and, ideally, with whose team researchers have close and regular collaboration. This allows researchers to follow up on specific, additional phenotypes that may present in the model but have not been recorded in the patient reports. Comprehensive phenotyping information is often very difficult to find in publicly accessible databases. By working closely with bioinformaticians and clinical geneticists, *Xenopus* researchers can not only enhance both their own understanding of the phenotypes and genotypes associated with each disease, but can also have rapid and powerful impact by testing genotype-phenotype links and helping clinicians support their patients and their families.

Accessing information through regulatory bodies and patient advocacy groups is an important approach, but involves a significant investment to ensure ethics and informed consent are considered properly. However, including these organisations is valuable since they offer opportunities to access patient registries, tissue biobanks and global collaborations, which improves the depth of data available. In particular, these groups often facilitate contacts with those working in different models, whether cells, organoids or other animals. It is increasingly recognised that multi-species comparisons are important for improving our understanding of gene-phenotype relationships and the key molecular findings linked to the perturbation of gene function in the context of human disease modelling and beyond.

To mitigate and anticipate the above issues in data sharing, patient consent, and ethics, the community is developing new data-sharing models. A prime example is the Centers for Mendelian Genomics (CMG) in the USA, now governed by the GREGoR Consortium (Bamshad et al., 2012). At recruitment for GREGoR, patients are

consented for sequencing and for anonymised sharing of their phenotype and genotype data globally with other clinicians and researchers working on the same gene/disease. In GREGoR, an investigator will usually recruit patients locally and patient DNA will be sequenced by one of the CMG sites. Sequenced and processed data are stored on a secure, central cloud solution, keeping one copy of the data in a central repository (Schatz et al., 2022). The CMG analyses the cloud-based genomic data with the investigator and makes it accessible for further study. For example, data from a Swedish cardiovascular disease cohort may be pooled and analysed together with that of a similar cohort from Ecuador. All CMG data are added to the database of Genotypes and Phenotypes (dbGaP) (Tryka et al., 2014) so that external researchers may apply for access. Since each CMG cohort has a named principal investigator, there is a clear line of communication for external researchers to liaise with investigators and clinicians, particularly to address demand for additional clinical information and for new samples for future research. Whilst the use of patient data is heavily regulated in some countries and this adds administrative requirements, it is important to recognise that using patient data from other countries for disease modelling in *Xenopus* may be unethical, depending on the nature of the permissions obtained.

#### Potential complications when working in *Xenopus*

There are obvious differences between *Xenopus* and humans which must be considered when designing human disease modelling experiments and interpreting phenotype data. These include neuroanatomical differences, the development of pronephric kidneys, a three-chambered cardiac structure, the process of metamorphosis, development of septated lungs, storage of fat in fat bodies and the process of skeletal development (Carotenuto et al., 2023). Not all differences should, however, be considered limitations. Despite having only one heart ventricle, *Xenopus* have atrioventricular valves, ventricular trabeculae, asymmetric atrium, and a similar *in vivo* cardiac orientation with left curvature of the outflow tract, which are sufficient anatomical structures to allow modelling of many cardiac abnormalities. Additionally, *Xenopus* tadpoles survive for an extended period with minimal cardiovascular function, enabling researchers to study end-stage cardiac deterioration (Smith et al., 2020) (reviewed by Hoppler and Conlon, 2020). There are many examples of key mechanistic data with human relevance being discovered in *Xenopus* organs that differ from humans (reviewed by Blackburn and Miller, 2019; Exner and Willsey, 2021; Womble et al., 2016).

Whilst *Xenopus* development is beautifully described and understood (Zahn et al., 2022), the focus of research has historically been on pre-metamorphic stages. When modelling human disease, it can be advantageous to directly compare the disease course, physiology, development and ageing processes of the model organisms to that of humans. Although there are many examples of such comparative (patho)physiology in the mouse (reviewed by Cacheiro et al., 2020; Potter et al., 2016), these broader longitudinal studies are currently lacking in *Xenopus* and an aged *Xenopus* model seems difficult to generate due to their prolonged lifespan, which commonly exceeds 15 years in the laboratory.

Bridging between humans and *Xenopus* will invariably raise classic questions of penetrance and expressivity. In human populations, disease-linked variant penetrance and expressivity often varies significantly between individuals (see e.g. Solomon et al., 2009; Mecklenburg et al., 2021). Similarly, human disease-associated variants may present somewhat differently than null variants do in mice or *Xenopus*. This could reflect evolutionary

differences in genetic background or that the human variants are not null variants and, instead, could be productively modelled as germline mutant lines in *X. tropicalis*, if the locus is well conserved.

Finally, an important variable to consider when modelling human genetic disorders is the effect of biological sex. Many genetic disorders are sex biased, especially psychiatric disorders, with important implications and opportunities for modelling (Christiansen et al., 2022). *Xenopus* have a genetic mechanism of sex determination. Although the sex determining-gene is currently unidentified in *X. tropicalis*, it is well characterised in the closely related species *X. laevis* (Yoshimoto et al., 2008). Conveniently, frog gonads form later than most other organs, morphologically and molecularly diverging at the beginning of metamorphosis (Pipek et al., 2017). As the majority of disease modelling work in the field right now uses frogs at earlier developmental stages, the effect of gonadal biology on disease course is bypassed. Convenient experimental manipulations can, however, affect sex determination: treating animals with estradiol causes sex reversal in genetic males and aromatase inhibitor treatment reverses the sex of genetic females (Chang and Witschi, 1956; Olmstead et al., 2009; Villalpando and Merchant-Larios, 1990). These tractable approaches may be useful for understanding the influence of biological sex on specific diseases in the future. Overall, *Xenopus* is a powerful model for diagnosing and modelling rare diseases, but there are challenges that researchers will often need to address.

#### Outlook

As our knowledge of human disease genetics expands, *Xenopus* is available to meet the fast-growing need to model gene and variant function *in vivo*. With a wide variety of experimental techniques and opportunities for many organ systems at scale, *Xenopus* will be a key model for working quickly and affordably. The ability to study many genes and variants in parallel in a diploid vertebrate is an unparalleled advantage of the system, and provides the opportunity to combat genetic pleiotropy and meet the challenges of modelling complex genetic disorders. Further, *Xenopus* holds great promise for studying particular variants of interest since it is feasible to generate and study many germline mutant lines in parallel. This is especially attractive as many human disease-causing variants are often not simply LOF. This approach maintains the appropriate genomic context of the variant and allows for the dissection of the precise effect of the variant. Finally, given the ease of drug screening, *Xenopus* can also provide a simple *in vivo* therapeutics screening platform.

Currently, a main barrier to success is a disconnect between clinicians, geneticists and model organism experts (Khokha et al., 2020). There is a real need to streamline and simplify the way genetic data and associated patient phenotypes are collected and disseminated to the experimentalists who want to study these genes. This can be incredibly frustrating to the patient families who want to share their data and is complicated by the essential need for informed consent and rigorous ethical considerations. Currently, many model organism researchers find accessing these data challenging, slowing and even inhibiting progress.

Such access must be carefully balanced with the protection of patient data and there is no global consensus on what is considered 'identifiable' information. A model that facilitates the seamless translation and safe data sharing between the research and clinical space is, thus, a key requirement, whereby there is global equity of access for clinicians, researchers, patients and functional laboratories to collaborate and participate in research that will directly impact diagnostics and therapeutics. Efforts like the Center for Mendelian Genetics and the DECIPHER database provide a model for how this



can be achieved at scale, and as patient groups grow and self-organise with help of foundations like COMBINEDbrain, there is the likelihood that these problems can be overcome. The *Xenopus* community is ready to meet patient and clinician needs.

It is clear from our experiences that there are more variants and genes of interest that need modelling than current resources allow. Model organisms are a core part of the solution to this major human health challenge. Within this space, the discussions throughout this article show that *Xenopus* has a key role, balancing its similarity with humans and broad experimental toolkit with its low cost and time effectiveness.

#### Acknowledgements

We thank the vibrant *Xenopus* community and our clinician partners for helpful discussions and viewpoints.

#### Competing interests

The authors declare no competing or financial interests.

#### Funding

We acknowledge funding from the following sources: an investigator award from the Chan Zuckerberg Biohub - San Francisco (to H.R.W.); a National Institutes of Health grant (grant nos EY017400 and EY022954 to R.M.G.); research awards from the Sharon Stewart Aniridia Trust and Vision For Tomorrow (to R.M.G.); a Medical Research Council grant (grant no. MR/V012177/1 to S.E. and M.G.); and an EPSRC grant (grant no. EP/Y01720X/1 to S.E. and M.G.).

#### References

- Abu-Daya, A. and Godwin, A. (2023). CRISPR/Cas9 gene disruption studies in *F<sub>0</sub>* *Xenopus* tadpoles: understanding development and disease in the frog. *Methods Mol. Biol.* **2633**, 111-130. doi:10.1007/978-1-0716-3004-4\_10
- Allou, L. and Mundlos, S. (2023). Disruption of regulatory domains and novel transcripts as disease-causing mechanisms. *BioEssays* **45**, e2300010. doi:10.1002/bies.202300010
- Appendix G, Single-Gene Disorders. (2010). *Understanding Genetics: A District of Columbia Guide for Patients and Health Professionals*. Genetic Alliance; District of Columbia Department of Health. <https://www.ncbi.nlm.nih.gov/books/NBK132154/>
- Aslan, Y., Tadjuidje, E., Zorn, A. M. and Cha, S.-W. (2017). High-efficiency non-mosaic CRISPR-mediated knock-in and indel mutation in *F<sub>0</sub>* *Xenopus*. *Development* **144**, 2852-2858. doi:10.1242/dev.152967
- Aztekin, C., Hiscock, T. W., Marioni, J. C., Gurdon, J. B., Simons, B. D. and Jullien, J. (2019). Identification of a regeneration-organizing cell in the *Xenopus* tail. *Science* **364**, 653-658. doi:10.1126/science.aav9996
- Bamshad, M. J., Shendure, J. A., Valle, D., Hamosh, A., Lupski, J. R., Gibbs, R. A., Boerwinkle, E., Lifton, R. P., Gerstein, M., Gunel, M. et al. (2012). The Centers for Mendelian Genomics: a new large-scale initiative to identify the genes underlying rare Mendelian conditions. *Am. J. Med. Genet. A* **158A**, 1523-1525. doi:10.1002/ajmg.a.35470
- Bamshad, M. J., Nickerson, D. A. and Chong, J. X. (2019). Mendelian gene discovery: fast and furious with no end in sight. *Am. J. Hum. Genet.* **105**, 448-455. doi:10.1016/j.ajhg.2019.07.011
- Barbosa, S., Greville-Heygate, S., Bonnet, M., Godwin, A., Fagotto-Kaufmann, C., Kajava, A. V., Laouteouet, D., Mawby, R., Wai, H. A., Dingemans, A. J. M. et al. (2020). Opposite modulation of RAC1 by mutations in TRIO1 is associated with distinct, domain-specific neurodevelopmental disorders. *Am. J. Hum. Genet.* **106**, 338-355. doi:10.1016/j.ajhg.2020.01.018
- Bhattacharya, D., Marfo, C. A., Li, D., Lane, M. and Khokha, M. K. (2015). CRISPR/Cas9: an inexpensive, efficient loss of function tool to screen human disease genes in *Xenopus*. *Dev. Biol.* **408**, 196-204. doi:10.1016/j.ydbio.2015.11.003
- Blackburn, A. T. M. and Miller, R. K. (2019). Modeling congenital kidney diseases in *Xenopus laevis*. *Dis. Model. Mech.* **12**, dmm038604. doi:10.1242/dmm.038604
- Blackburn, A. T. M., Bekheirnia, N., Uma, V. C., Corkins, M. E., Xu, Y., Rosenfeld, J. A., Bainbridge, M. N., Yang, Y., Liu, P., Madan-Khetarpal, S. et al. (2019). DYRK1A-related intellectual disability: a syndrome associated with congenital anomalies of the kidney and urinary tract. *Genet. Med.* **21**, 2755-2764. doi:10.1038/s41436-019-0576-0
- Boskovski, M. T., Yuan, S., Pedersen, N. B., Goth, C. K., Makova, S., Clausen, H., Brueckner, M. and Khokha, M. K. (2013). The heterotaxy gene GALNT11 glycosylates notch to orchestrate cilia type and laterality. *Nature* **504**, 456-459. doi:10.1038/nature12723
- Bowes, J. B., Snyder, K. A., Segerdell, E., Jarabek, C. J., Azam, K., Zorn, A. M. and Vize, P. D. (2010). Xenbase: gene expression and improved integration. *Nucleic Acids Res.* **38**, D607-D612. doi:10.1093/nar/gkp953
- Boyling, A., Perez-Siles, G. and Kennerson, M. L. (2022). Structural variation at a disease mutation hotspot: strategies to investigate gene regulation and the 3D genome. *Front. Genet.* **13**, 842860. doi:10.3389/fgene.2022.842860
- Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W. and Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360**, eaar5780. doi:10.1126/science.aar5780
- Brinkman, E. K., Chen, T., Amendola, M. and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168. doi:10.1093/nar/gku936
- Cacheiro, P., Muñoz-Fuentes, V., Murray, S. A., Dickinson, M. E., Bucan, M., Nutter, L. M. J., Peterson, K. A., Haselimiashadi, H., Flenniken, A. M., Morgan, H. et al. (2020). Human and mouse essentiality screens as a resource for disease gene discovery. *Nat. Commun.* **11**, 655. doi:10.1038/s41467-020-14284-2
- Carotenuto, R., Pallotta, M. M., Tussellino, M. and Fogliano, C. (2023). *Xenopus laevis* (Daudin, 1802) as a model organism for bioscience: a historic review and perspective. *Biology* **12**, 890. doi:10.3390/biology12060890
- Chang, C. Y. and Witschi, E. (1956). Genic control and hormonal reversal of sex differentiation in *Xenopus*. *Proc. Soc. Exp. Biol. Med.* **93**, 140-144. doi:10.3181/00379727-93-22688
- Christiansen, D. M., McCarthy, M. M. and Seeman, M. V. (2022). Where sex meets gender: how sex and gender come together to cause sex differences in mental illness. *Front. Psychiatry* **13**, 856436. doi:10.3389/fpsy.2022.856436
- Cleaver, O. and Krieg, P. A. (1999). Expression from DNA injected into *Xenopus* Embryos. *Methods Mol. Biol.* **127**, 133-153. doi:10.1385/1-59259-678-9:133
- Colleluori, V. and Khokha, M. K. (2023). Mink1 regulates spemann organizer cell fate in the *Xenopus* gastrula via Hmga2. *Dev. Biol.* **495**, 42-53. doi:10.1016/j.ydbio.2022.11.010
- Conant, D., Hsiau, T., Rossi, N., Oki, J., Maures, T., Waite, K., Yang, J., Joshi, S., Kelso, R., Holden, K. et al. (2022). Inference of CRISPR edits from Sanger trace data. *CRISPR J* **5**, 123-130. doi:10.1089/crispr.2021.0113
- Cummings, B. B., Karczewski, K. J., Kosmicki, J. A., Seaby, E. G., Watts, N. A., Singer-Berk, M., Mudge, J. M., Karjalainen, J., Satterstrom, F. K., O'Donnell-Luria, A. H. et al. (2020). Transcript expression-aware annotation improves rare variant interpretation. *Nature* **581**, 452-458. doi:10.1038/s41586-020-2329-2
- Dale, L. and Slack, J. M. (1987). Fate map for the 32-cell stage of *Xenopus laevis*. *Development* **99**, 527-551. doi:10.1242/dev.99.4.527
- Date, P., Ackermann, P., Furey, C., Fink, I. B., Jonas, S., Khokha, M. K., Kahle, K. T. and Deniz, E. (2019). Visualizing flow in an intact CSF network using optical coherence tomography: implications for human congenital hydrocephalus. *Sci. Rep.* **9**, 6196. doi:10.1038/s41598-019-42549-4
- DeLay, B. D., Corkins, M. E., Hanania, H. L., Salanga, M., Deng, J. M., Sudou, N., Taira, M., Horb, M. E. and Miller, R. K. (2018). Tissue-specific gene inactivation in *Xenopus laevis*: knockout of *lhx1* in the kidney with CRISPR/Cas9. *Genetics* **208**, 673-686. doi:10.1534/genetics.117.300468
- den Dunnen, J. T., Dalglish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., Roux, A.-F., Smith, T., Antonarakis, S. E. and Taschner, P. E. M. (2016). HGVS recommendations for the description of sequence variants: 2016 update. *Hum. Mutat.* **37**, 564-569. doi:10.1002/humu.22981
- Deniz, E., Jonas, S., Hooper, M., N Griffin, J., Choma, M. A. and Khokha, M. K. (2017). Analysis of craniocardiac malformations in *Xenopus* using optical coherence tomography. *Sci. Rep.* **7**, 42506. doi:10.1038/srep42506
- Deniz, E., Mis, E. K., Lane, M. and Khokha, M. K. (2018). CRISPR/Cas9 *F<sub>0</sub>* screening of congenital heart disease genes in *Xenopus tropicalis*. *Methods Mol. Biol.* **1865**, 163-174. doi:10.1007/978-1-4939-8784-9\_12
- Deniz, E., Pasha, M., Guerra, M. E., Viviano, S., Ji, W., Konstantino, M., Jeffries, L., Lakhani, S. A., Medne, L., Skraban, C. et al. (2023). CFAP45, a heterotaxy and congenital heart disease gene, affects cilia stability. *Dev. Biol.* **499**, 75-88. doi:10.1016/j.ydbio.2023.04.006
- Devotta, A., Juraver-Geslin, H., Gonzalez, J. A., Hong, C.-S. and Saint-Jeanet, J.-P. (2016). Sf3b4-depleted *Xenopus* embryos: a model to study the pathogenesis of craniofacial defects in Nager syndrome. *Dev. Biol.* **415**, 371-382. doi:10.1016/j.ydbio.2016.02.010
- Dimitrakopoulou, D., Tulkens, D., Van Vlierberghe, P. and Vleminckx, K. (2019). *Xenopus tropicalis*: joining the armada in the fight against blood cancer. *Front. Physiol.* **10**, 48. doi:10.3389/fphys.2019.00048
- Edwards, N. A., Shacham-Silverberg, V., Weitz, L., Kingma, P. S., Shen, Y., Wells, J. M., Chung, W. K. and Zorn, A. M. (2021). Developmental basis of trachea-esophageal birth defects. *Dev. Biol.* **477**, 85-97. doi:10.1016/j.ydbio.2021.05.015
- Exner, C. R. T. and Willsey, H. R. (2021). *Xenopus* leads the way: frogs as a pioneering model to understand the human brain. *Genesis* **59**, e23405. doi:10.1002/dvg.23405
- Fainsod, A. and Moody, S. A. (2022). *Xenopus: From Basic Biology to Disease Models in the Genomic Era*. CRC Press.
- Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpes, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R. M. and Carter, N. P. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am. J. Hum. Genet.* **84**, 524-533. doi:10.1016/j.ajhg.2009.03.010

- Fish, M. B., Nakayama, T., Fisher, M., Hirsch, N., Cox, A., Reeder, R., Carruthers, S., Hall, A., Stemple, D. L. and Grainger, R. M. (2014). Xenopus mutant reveals necessity of rax for specifying the eye field which otherwise forms tissue with telencephalic and diencephalic character. *Dev. Biol.* **395**, 317-330. doi:10.1016/j.ydbio.2014.09.004
- Fisher, M., James-Zorn, C., Ponferrada, V., Bell, A. J., Sundararaj, N., Segerdell, E., Chaturvedi, P., Bayyari, N., Chu, S., Pells, T. et al. (2023). Xenbase: key features and resources of the Xenopus model organism knowledgebase. *Genetics* **224**, iyad018. doi:10.1093/genetics/iyad018
- Fortriede, J. D., Pells, T. J., Chu, S., Chaturvedi, P., Wang, D., Fisher, M. E., James-Zorn, C., Wang, Y., Nenni, M. J., Burns, K. A. et al. (2020). Xenbase: deep integration of GEO & SRA RNA-seq and ChIP-seq data in a model organism database. *Nucleic Acids Res.* **48**, D776-D782. doi:10.1093/nar/gkz933
- Getwan, M. and Lienkamp, S. S. (2017). Toolbox in a tadpole: Xenopus for kidney research. *Cell Tissue Res.* **369**, 143-157. doi:10.1007/s00441-017-2611-2
- Getwan, M., Hoppmann, A., Schlosser, P., Grand, K., Song, W., Diehl, R., Schroda, S., Heeg, F., Deutsch, K., Hildebrandt, F. et al. (2021). Ttc30a affects tubulin modifications in a model for ciliary chondrodysplasia with polycystic kidney disease. *Proc. Natl. Acad. Sci. USA* **118**, e2106770118. doi:10.1073/pnas.2106770118
- Gore, S. V., James, E. J., Huang, L.-C., Park, J. J., Berghella, A., Thompson, A. C., Cline, H. T. and Aizenman, C. D. (2021). Role of matrix metalloproteinase-9 in neurodevelopmental deficits and experience-dependent plasticity in Xenopus laevis. *Elife* **10**, e62147. doi:10.7554/eLife.62147
- Grainger, R. M. (2012). Xenopus tropicalis as a model organism for genetics and genomics: past, present, and future. *Methods Mol. Biol.* **917**, 3-15. doi:10.1007/978-1-61779-992-1\_1
- Grand, K., Stoltz, M., Rizzo, L., Röck, R., Kaminski, M. M., Salinas, G., Getwan, M., Naert, T., Pichler, R. and Lienkamp, S. S. (2023). HNF1B alters an evolutionarily conserved nephrogenic program of target genes. *J. Am. Soc. Nephrol.* **34**, 412-432. doi:10.1681/ASN.2022010076
- Grant, I. M., Balcha, D., Hao, T., Shen, Y., Trivedi, P., Patrushev, I., Fortriede, J. D., Karpinka, J. B., Liu, L., Zorn, A. M. et al. (2015). The Xenopus ORFeome: a resource that enables functional genomics. *Dev. Biol.* **408**, 345-357. doi:10.1016/j.ydbio.2015.09.004
- Griffin, J. N., Sondalle, S. B., Robson, A., Mis, E. K., Griffin, G., Kulkarni, S. S., Deniz, E., Baserga, S. J. and Khokha, M. K. (2018). RPSA, a candidate gene for isolated congenital Asplenia, is required for pre-rRNA processing and spleen formation in Xenopus. *Development* **145**, dev166181. doi:10.1242/dev.166181
- Griffin, J. N., Sempou, E. and Liu, K. (2020). *Xenopus Models of Organogenesis and Disease*. Frontiers Media SA.
- Gudmundsson, S., Singer-Berk, M., Watts, N. A., Phu, W., Goodrich, J. K., Solomonson, M., Genome Aggregation Database Consortium, Rehm, H. L., MacArthur, D. G. and O'Donnell-Luria, A. (2022). Variant interpretation using population databases: lessons from gnomAD. *Hum. Mutat.* **43**, 1012-1030. doi:10.1002/humu.24309
- Gull, M., Schmitt, S. M., Kälin, R. E. and Brändli, A. W. (2023). Screening of chemical libraries using xenopus embryos and tadpoles for phenotypic drug discovery. *Cold Spring Harb. Protoc.* **2023**, pdb.prot098269. doi:10.1101/pdb.prot098269
- Hamosh, A., Scott, A. F., Amberger, J. S., Bocchini, C. A. and McKusick, V. A. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* **33**, D514-D517. doi:10.1093/nar/gki033
- Harland, R. M. and Grainger, R. M. (2011). Xenopus research: metamorphosed by genetics and genomics. *Trends Genet.* **27**, 507-515. doi:10.1016/j.tig.2011.08.003
- Hellsten, U., Harland, R. M., Gilchrist, M. J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N. H., Shu, S., Taher, L. et al. (2010). The genome of the Western clawed frog *Xenopus tropicalis*. *Science* **328**, 633-636. doi:10.1126/science.1183670
- Hill, R. E., Favor, J., Hogan, B. L., Ton, C. C., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan, T., Hastie, N. D. and van Heyningen, V. (1991). Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354**, 522-525. doi:10.1038/354522a0
- Hiramoto, M. and Cline, H. T. (2021). Precisely controlled visual stimulation to study experience-dependent neural plasticity in Xenopus tadpoles. *STAR Protoc* **2**, 100252. doi:10.1016/j.xpro.2020.100252
- Hirose, G. and Jacobson, M. (1979). Clonal organization of the central nervous system of the frog: I. Clones stemming from individual blastomeres of the 16-cell and earlier stages. *Dev. Biol.* **71**, 191-202. doi:10.1016/0012-1606(79)90163-5
- Hoppler, S. and Conlon, F. L. (2020). Xenopus: experimental access to cardiovascular development, regeneration discovery, and cardiovascular heart-defect modeling. *Cold Spring Harb. Perspect. Biol.* **12**, a037200. doi:10.1101/cshperspect.a037200
- Hubrecht, R. C. and Carter, E. (2019). The 3Rs and humane experimental technique: implementing change. *Animals* **9**, 754. doi:10.3390/ani9100754
- Hwang, W. Y., Marquez, J. and Khokha, M. K. (2019). Xenopus: driving the discovery of novel genes in patient disease and their underlying pathological mechanisms relevant for organogenesis. *Front. Physiol.* **10**, 953. doi:10.3389/fphys.2019.00953
- Ismail, V., Zachariassen, L. G., Godwin, A., Sahakian, M., Ellard, S., Stals, K. L., Baple, E., Brown, K. T., Foulds, N., Wheway, G. et al. (2022). Identification and functional evaluation of GRIA1 missense and truncation variants in individuals with ID: an emerging neurodevelopmental syndrome. *Am. J. Hum. Genet.* **109**, 1217-1241. doi:10.1016/j.ajhg.2022.05.009
- Jacobson, M. and Hirose, G. (1981). Clonal organization of the central nervous system of the frog. II. Clones stemming from individual blastomeres of the 32- and 64-cell stages. *J. Neurosci.* **1**, 271-284. doi:10.1523/JNEUROSCI.01-03-00271.1981
- Takebe, A. D., Chitsazan, A. D., Williams, M. C., Saunders, L. M. and Wills, A. E. (2020). Chromatin accessibility dynamics and single cell RNA-Seq reveal new regulators of regeneration in neural progenitors. *Elife* **9**, e52648. doi:10.7554/eLife.52648
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P. et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434-443. doi:10.1038/s41586-020-2308-7
- Kariminejad, A., Szenker-Ravi, E., Lekszas, C., Tajsharghi, H., Moslemi, A.-R., Naert, T., Tran, H. T., Ahangari, F., Rajaei, M., Nasser, M. et al. (2019). Homozygous null TBX4 mutations lead to posterior amelia with pelvic and pulmonary hypoplasia. *Am. J. Hum. Genet.* **105**, 1294-1301. doi:10.1016/j.ajhg.2019.10.013
- Khokha, M. K., Liu, K. J. and Wallingford, J. B. (2020). Challenges and opportunities at the interface of birth defects, human genetics and developmental biology. *Development* **147**, dev197871. doi:10.1242/dev.197871
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424. doi:10.1038/nature17946
- Kostiuk, V. and Khokha, M. K. (2021). Xenopus as a platform for discovery of genes relevant to human disease. *Curr. Top. Dev. Biol.* **145**, 277-312. doi:10.1016/bs.ctdb.2021.03.005
- Kulkarni, S. S. and Khokha, M. K. (2018). WDR5 regulates left-right patterning via chromatin-dependent and independent functions. *Development* **145**, dev159889. doi:10.1242/dev.159889
- Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H. and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171-W174. doi:10.1093/nar/gkz365
- Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M. and Maglott, D. R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42**, D980-D985. doi:10.1093/nar/gkt1113
- Lane, M. and Khokha, M. K. (2022a). Obtaining Xenopus tropicalis embryos by natural mating. *Cold Spring Harb. Protoc.* **2022**, Pdb.prot106609. doi:10.1101/pdb.prot106609
- Lane, M. and Khokha, M. K. (2022b). Obtaining Xenopus tropicalis embryos by in vitro fertilization. *Cold Spring Harb. Protoc.* **2022**, Pdb.prot106351. doi:10.1101/pdb.prot106351
- Lennon, G., Auffray, C., Polymeropoulos, M. and Soares, M. B. (1996). The I. M. A. G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**, 151-152. doi:10.1006/geno.1996.0177
- Macken, W. L., Godwin, A., Wheway, G., Stals, K., Nazlamova, L., Ellard, S., Alfares, A., Aloraini, T., AlSubaie, L., AlFadhel, M. et al. (2021). Biallelic variants in COPB1 cause a novel, severe intellectual disability syndrome with cataracts and variable microcephaly. *Genome Med.* **13**, 34. doi:10.1186/s13073-021-00850-w
- Malarte, M., Short, S. and Sharpe, C. (2006). Xenopus embryos lacking specific isoforms of the corepressor SMRT develop abnormal heads. *Dev. Biol.* **292**, 333-343. doi:10.1016/j.ydbio.2006.01.007
- Marquez, J., Bhattacharya, D., Lusk, C. P. and Khokha, M. K. (2021). Nucleoporin NUP205 Plays a Critical Role in Cilia and Congenital Disease. *Dev. Biol.* **469**, 46-53. doi:10.1016/j.ydbio.2020.10.001
- Martin, S. A. (2023). A CRISPR/Cas-based method for precise DNA integration in xenopus laevis oocytes followed by intracytoplasmic sperm injection (ICSI) fertilization. *Methods Mol. Biol.* **2633**, 131-143. doi:10.1007/978-1-0716-3004-4\_11
- Mascibroda, L. G., Shboul, M., Elrod, N. D., Colleaux, L., Hamamy, H., Huang, K.-L., Peart, N., Singh, M. K., Lee, H., Merriman, B. et al. (2022). INTS13 variants causing a recessive developmental ciliopathy disrupt assembly of the integrator complex. *Nat. Commun.* **13**, 6054. doi:10.1038/s41467-022-33547-8
- Masho, R. (1988). Fates of animal-dorsal blastomeres of eight-cell stage Xenopus embryos vary according to the specific patterns of the third cleavage plane: (Xenopus Embryos/animal-Dorsal Blastomeres/third Cleavage/developmental Fates/fluorescein Dextran Amine). *Dev. Growth Differ.* **30**, 347-359. doi:10.1111/j.1440-169X.1988.00347.x
- Masho, R. and Kubota, H. Y. (1986). Developmental fates of blastomeres of eight-cell-stage Xenopus laevis Embryos: (intracellular Injection/horseradish Peroxidase/developmental fate/Xenopus Embryo). *Dev. Growth Differ.* **28**, 113-123. doi:10.1111/j.1440-169X.1986.00113.x



- Mecklenburg, N., Kowalczyk, I., Witte, F., Görne, J., Laier, A., Mamo, T. M., Gonschior, H., Lehmann, M., Richter, M., Sporb, A. et al. (2021). Identification of disease-relevant modulators of the SHH pathway in the developing brain. *Development* **148**, dev199307. doi:10.1242/dev.199307
- Moody, S. A. (1987a). Fates of the blastomeres of the 16-cell stage *Xenopus* Embryo. *Dev. Biol.* **119**, 560-578. doi:10.1016/0012-1606(87)90059-5
- Moody, S. A. (1987b). Fates of the blastomeres of the 32-cell-stage *Xenopus* Embryo. *Dev. Biol.* **122**, 300-319. doi:10.1016/0012-1606(87)90296-X
- Moody, S. A. and Kline, M. J. (1990). Segregation of Fate during Cleavage of Frog (*Xenopus Laevis*) Blastomeres. *Anat. Embryol.* **182**, 347-362. doi:10.1007/BF02433495
- Moreno-Mateos, M. A., Vejnár, C. E., Beaudoin, J.-D., Fernández, J. P., Mis, E. K., Khokha, M. K. and Giraldez, A. J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* **12**, 982-988. doi:10.1038/nmeth.3543
- Naert, T., Tulkens, D., Edwards, N. A., Carron, M., Shaidani, N.-I., Wlizia, M., Boel, A., Demuyne, S., Horb, M. E., Coucke, P. et al. (2020). Maximizing CRISPR/Cas9 phenotype penetrance applying predictive modeling of editing outcomes in *Xenopus* and zebrafish embryos. *Sci. Rep.* **10**, 14662. doi:10.1038/s41598-020-71412-0
- Naert, T., Çiçek, Ö., Ogar, P., Bürgi, M., Shaidani, N.-I., Kaminski, M. M., Xu, Y., Grand, K., Vujanovic, M., Prata, D. et al. (2021a). Deep learning is widely applicable to phenotyping embryonic development and disease. *Development* **148**, dev199664. doi:10.1242/dev.199664
- Naert, T., Tulkens, D., Van Nieuwenhuysen, T., Przybyl, J., Demuyne, S., van de Rijn, M., Al-Jazrawe, M., Aliman, B. A., Coucke, P. J., De Leeneer, K. et al. (2021b). CRISPR-SID: identifying EZH2 as a druggable target for desmoid tumors via in vivo dependency mapping. *Proc. Natl. Acad. Sci. USA* **118**, e2115116118. doi:10.1073/pnas.2115116118
- Nakamura, O. and Kishiyama, K. (1971). Prospective fates of blastomeres at the 32 cell stage of *Xenopus Laevis* embryos. *Proc. Jpn. Acad.* **47**, 407-412. doi:10.2183/pjab.1945.47.407
- Nakayama, T. and Grainger, R. M. (2023). Best practices for *Xenopus* tropicalis husbandry. *Cold Spring Harb. Protoc.* **2023**, pdb.top106252. doi:10.1101/pdb.top106252
- Nakayama, T., Fish, M. B., Fisher, M., Oomen-Hajagos, J., Thomsen, G. H. and Grainger, R. M. (2013). Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. *Genesis* **51**, 835-843. doi:10.1002/dvg.22720
- Nakayama, T., Fisher, M., Nakajima, K., Odeleye, A. O., Zimmerman, K. B., Fish, M. B., Yaoita, Y., Chojnowski, J. L., Lauderdale, J. D., Netland, P. A. et al. (2015). *Xenopus* pax6 mutants affect eye development and other organ systems, and have phenotypic similarities to human aniridia patients. *Dev. Biol.* **408**, 328-344. doi:10.1016/j.ydbio.2015.02.012
- Nakayama, T., Grainger, R. M. and Cha, S.-W. (2020). Simple embryo injection of long single-stranded donor templates with the CRISPR/Cas9 system leads to homology-directed repair in *Xenopus tropicalis* and *Xenopus laevis*. *Genesis* **58**, e23366. doi:10.1002/dvg.23366
- Nakayama, T., Roubroeks, J. A. Y., Veenstra, G. J. C. and Grainger, R. M. (2022a). Preparation of intact nuclei for single-nucleus omics using frozen cell suspensions from mutant embryos of *Xenopus tropicalis*. *Cold Spring Harb. Protoc.* **2022**, 641-652. doi:10.1101/pdb.prot107825
- Nakayama, T., Cox, A., Howell, M. and Grainger, R. M. (2022b). Gynogenetic production of embryos in *Xenopus tropicalis* using a cold shock procedure: rapid screening method for gene editing phenotypes. *Cold Spring Harb. Protoc.* **2022**, 616-623. doi:10.1101/pdb.prot107648
- Nakayama, T., Grainger, R. M. and Cha, S.-W. (2022c). Homology-directed repair by CRISPR-Cas9 mutagenesis in *Xenopus* using long single-stranded donor DNA templates via simple microinjection of embryos. *Cold Spring Harb. Protoc.* **2022**, 606-615. doi:10.1101/pdb.prot107599
- Nenni, M. J., Fisher, M. E., James-Zorn, C., Pells, T. J., Ponferrada, V., Chu, S., Fortriede, J. D., Burns, K. A., Wang, Y., Lotay, V. S. et al. (2019). Xenbase: facilitating the use of *Xenopus* to model human disease. *Front. Physiol.* **10**, 154. doi:10.3389/fphys.2019.00154
- Noble, A., Abu-Daya, A. and Guille, M. J. (2022). I-SceI-mediated transgenesis in *Xenopus*. *Cold Spring Harb. Protoc.* **2022**, Pdb.prot107011. doi:10.1101/pdb.prot107011
- Ogino, H., McConnell, W. B. and Grainger, R. M. (2006). High-throughput transgenesis in *Xenopus* using I-SceI meganuclease. *Nat. Protoc.* **1**, 1703-1710. doi:10.1038/nprot.2006.208
- Ogino, S., Gulley, M. L., den Dunnen, J. T. and Wilson, R. B. and Association for Molecular Pathology Training and Education Committee. (2007). Standard mutation nomenclature in molecular diagnostics: practical and educational challenges. *J. Mol. Diagn.* **9**, 1-6. doi:10.2353/jmoldx.2007.060081
- Olmstead, A. W., Kosian, P. A., Korte, J. J., Holcombe, G. W., Woodis, K. K. and Degitz, S. J. (2009). Sex reversal of the amphibian, *Xenopus tropicalis*, following larval exposure to an aromatase inhibitor. *Aquat. Toxicol.* **91**, 143-150. doi:10.1016/j.aquatox.2008.07.018
- Panagiotakos, G. and Pasca, S. P. (2022). A matter of space and time: emerging roles of disease-associated proteins in neural development. *Neuron* **110**, 195-208. doi:10.1016/j.neuron.2021.10.035
- Panthi, S., Chapman, P. A., Szyszka, P. and Beck, C. W. (2023). Characterisation and automated quantification of induced seizure-related behaviours in *Xenopus laevis* tadpoles. *J. Neurochem.* doi:10.1111/jnc.15836
- Park, D.-S., Yoon, M., Kweon, J., Jang, A.-H., Kim, Y. and Choi, S.-C. (2017). Targeted base editing via RNA-guided cytidine deaminases in *Xenopus laevis* embryos. *Mol. Cells* **40**, 823-827. doi:10.14348/molcells.2017.0262
- Piprek, R. P., Kloc, M., Tassan, J.-P. and Kubiak, J. Z. (2017). Development of *Xenopus laevis* bipotential gonads into testis or ovary is driven by sex-specific cell-cell interactions, proliferation rate, cell migration and deposition of extracellular matrix. *Dev. Biol.* **432**, 298-310. doi:10.1016/j.ydbio.2017.10.020
- Potter, P. K., Bowl, M. R., Jeyarajan, P., Wisby, L., Blease, A., Goldsworthy, M. E., Simon, M. M., Greenaway, S., Michel, V., Barnard, A. et al. (2016). Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nat. Commun.* **7**, 12444. doi:10.1038/ncomms12444
- Rana, A. A., Collart, C., Gilchrist, M. J. and Smith, J. C. (2006). Defining synphenotype groups in *Xenopus tropicalis* by use of antisense morpholino oligonucleotides. *PLoS Genet.* **2**, e193. doi:10.1371/journal.pgen.0020193
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E. et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405-424. doi:10.1038/gim.2015.30
- Roberts, A. E., Araki, T., Swanson, K. D., Montgomery, K. T., Schiripo, T. A., Joshi, V. A., Li, L., Yassin, Y., Tamburino, A. M., Neel, B. G. et al. (2007). Germline gain-of-function mutations in SOS1 cause Noonan syndrome. *Nat. Genet.* **39**, 70-74. doi:10.1038/ng1926
- Robinson, P. N., Köhler, S., Bauer, S., Seelow, D., Horn, D. and Mundlos, S. (2008). The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am. J. Hum. Genet.* **83**, 610-615. doi:10.1016/j.ajhg.2008.09.017
- Rosenthal, S. B., Willsey, H. R., Xu, Y., Mei, Y., Dea, J., Wang, S., Curtis, C., Sempou, E., Khokha, M. K., Chi, N. C. et al. (2021). A convergent molecular network underlying autism and congenital heart disease. *Cell Syst.* **12**, 1094-1107.e6. doi:10.1016/j.cels.2021.07.009
- Samocha, K. E., Kosmicki, J. A., Karczewski, K. J., O'Donnell-Luria, A. H., Pierce-Hoffman, E., MacArthur, D. G., Neale, B. M. and Daly, M. J. (2017). Regional missense constraint improves variant deleteriousness prediction. *bioRxiv* 148353.
- Schatz, M. C., Philippakis, A. A., Afgan, E., Banks, E., Carey, V. J., Carroll, R. J., Culotti, A., Ellrott, K., Goecks, J., Grossman, R. L. et al. (2022). Inverting the model of genomics data sharing with the NHGRI genomic data science analysis, visualization, and informatics lab-space. *Cell Genom.* **2**, 100085. doi:10.1016/j.xgen.2021.100085
- Schietroma, C., Parain, K., Estivalet, A., Aghaie, A., Boutet de Monvel, J., Picaud, S., Sahel, J.-A., Perron, M., El-Amraoui, A. and Petit, C. (2017). Usher syndrome type 1-associated cadherins shape the photoreceptor outer segment. *J. Cell Biol.* **216**, 1849-1864. doi:10.1083/jcb.201612030
- Seaby, E. G. and Ennis, S. (2020). Challenges in the diagnosis and discovery of rare genetic disorders using contemporary sequencing technologies. *Brief. Funct. Genomics* **19**, 243-258. doi:10.1093/bfpp/ela009
- Seaby, E. G., Smedley, D., Taylor Tavares, A. L., Brittain, H., van Jaarsveld, R. H., Baralle, D., Rehm, H. L., O'Donnell-Luria, A. and Ennis, S. and Genomics England Research Consortium. (2022). A gene-to-patient approach uplifts novel disease gene discovery and identifies 18 putative novel disease genes. *Genet. Med.* **24**, 1697-1707. doi:10.1016/j.gim.2022.04.019
- Sega, A. G., Mis, E. K., Lindstrom, K., Mercimek-Andrews, S., Ji, W., Cho, M. T., Juusola, J., Konstantino, M., Jeffries, L., Khokha, M. K. et al. (2019). De novo pathogenic variants in neuronal differentiation factor 2 (NEUROD2) cause a form of early infantile epileptic encephalopathy. *J. Med. Genet.* **56**, 113-122. doi:10.1136/jmedgenet-2018-105322
- Sempou, E., Kostiuik, V., Zhu, J., Guerra, M. C., Tyan, L., Hwang, W., Camacho-Aguilar, E., Caplan, M. J., Zenisek, D., Warmflash, A. et al. (2022). Membrane potential drives the exit from pluripotency and cell fate commitment via calcium and mTOR. *Nat. Commun.* **13**, 6681. doi:10.1038/s41467-022-34363-w
- Sestan, N. and State, M. W. (2018). Lost in translation: traversing the complex path from genomics to therapeutics in autism spectrum disorder. *Neuron* **100**, 406-423. doi:10.1016/j.neuron.2018.10.015
- Shaidani, N.-I., McNamara, S., Wlizia, M. and Horb, M. E. (2020). Animal maintenance systems: *Xenopus laevis*. *Cold Spring Harb. Protoc.* **2020**, pdb.prot106138. doi:10.1101/pdb.prot106138
- Shi, Z., Xin, H., Tian, D., Lian, J., Wang, J., Liu, G., Ran, R., Shi, S., Zhang, Z., Shi, Y. et al. (2019). Modeling human point mutation diseases in with a modified CRISPR/Cas9 system. *FASEB J.* **33**, 6962-6968. doi:10.1096/fj.201802661R
- Shibata, Y., Suzuki, M., Hirose, N., Takayama, A., Sanbo, C., Inoue, T., Umeson, Y., Agata, K., Ueno, N., Suzuki, K.-I. T. et al. (2022). CRISPR/Cas9-based simple transgenesis in *Xenopus laevis*. *Dev. Biol.* **489**, 76-83. doi:10.1016/j.ydbio.2022.06.001
- Smith, S. J., Towers, N., Demetriou, K. and Mohun, T. J. (2020). Defective heart chamber growth and myofibrillogenesis after knockout of *adprhl1* gene function by



- targeted disruption of the ancestral catalytic active site. *PLoS One* **15**, e0235433. doi:10.1371/journal.pone.0235433
- Solomon, B. D., Lacbawan, F., Jain, M., Domené, S., Roessler, E., Moore, C., Dobyms, W. B. and Muenke, M.** (2009). A novel SIX3 mutation segregates with holoprosencephaly in a large family. *Am. J. Med. Genet. A* **149A**, 919-925. doi:10.1002/ajmg.a.32813
- Sun, Y. A. and Poo, M. M.** (1985). Non-quantal release of acetylcholine at a developing neuromuscular synapse in culture. *J. Neurosci.* **5**, 634-642. doi:10.1523/JNEUROSCI.05-03-00634.1985
- Tryka, K. A., Hao, L., Sturcke, A., Jin, Y., Wang, Z. Y., Ziyabari, L., Lee, M., Popova, N., Sharopova, N., Kimura, M. et al.** (2014). NCBI's database of genotypes and phenotypes: dbGaP. *Nucleic Acids Res.* **42**, D975-D979. doi:10.1093/nar/gkt1211
- Villalpando, I. and Merchant-Larios, H.** (1990). Determination of the sensitive stages for gonadal sex-reversal in *Xenopus laevis* tadpoles. *Int. J. Dev. Biol.* **34**, 281-285.
- Wheeler, G. N. and Liu, K. J.** (2012). *Xenopus*: an ideal system for chemical genetics. *Genesis* **50**, 207-218. doi:10.1002/dvg.22009
- Willsey, H. R.** (2021). Whole-mount RNA in situ hybridization and immunofluorescence of embryos and tadpoles. *Cold Spring Harb. Protoc.* **2021**, pdb.prot105635. doi:10.1101/pdb.prot105635
- Willsey, A. J., Morris, M. T., Wang, S., Willsey, H. R., Sun, N., Teerikorpi, N., Baum, T. B., Cagney, G., Bender, K. J., Desai, T. A. et al.** (2018a). The psychiatric cell map initiative: a convergent systems biological approach to illuminating key molecular pathways in neuropsychiatric disorders. *Cell* **174**, 505-520. doi:10.1016/j.cell.2018.06.016
- Willsey, H. R., Walentek, P., Exner, C. R. T., Xu, Y., Lane, A. B., Harland, R. M., Heald, R. and Santama, N.** (2018b). Katanin-like protein *Katnal2* is required for ciliogenesis and brain development in *Xenopus* embryos. *Dev. Biol.* **442**, 276-287. doi:10.1016/j.ydbio.2018.08.002
- Willsey, H. R., Xu, Y., Everitt, A., Dea, J., Exner, C. R. T., Willsey, A. J., State, M. W. and Harland, R. M.** (2020). The neurodevelopmental disorder risk gene *DYRK1A* is required for ciliogenesis and control of brain size in *Xenopus* embryos. *Development* **147**, dev189290. doi:10.1242/dev.189290
- Willsey, H. R., Exner, C. R. T., Xu, Y., Everitt, A., Sun, N., Wang, B., Dea, J., Schmunk, G., Zaltsman, Y., Teerikorpi, N. et al.** (2021). Parallel in vivo analysis of large-effect autism genes implicates cortical neurogenesis and estrogen in risk and resilience. *Neuron* **109**, 1409. doi:10.1016/j.neuron.2021.03.030
- Willsey, H. R., Guille, M. and Grainger, R. M.** (2022a). Modeling human genetic disorders with CRISPR technologies in *Xenopus*. *Cold Spring Harb. Protoc.* **2022**, pdb.prot106997. doi:10.1101/pdb.prot106997
- Willsey, H. R., Willsey, A. J., Wang, B. and State, M. W.** (2022b). Genomics, convergent neuroscience and progress in understanding autism spectrum disorder. *Nat. Rev. Neurosci.* **23**, 323-341. doi:10.1038/s41583-022-00576-7
- Womble, M., Pickett, M. and Nascone-Yoder, N.** (2016). Frogs as integrative models for understanding digestive organ development and evolution. *Semin. Cell Dev. Biol.* **51**, 92-105. doi:10.1016/j.semcdb.2016.02.001
- Wong, S., Tan, Y. X., Loh, A. Y. T., Tan, K. Y., Lee, H., Aziz, Z., Nelson, S. F., Özkan, E., Kayserili, H., Escande-Beillard, N. et al.** (2023). *RAF1* deficiency causes a lethal syndrome that underscores RTK signaling during embryogenesis. *EMBO Mol. Med.* **15**, e17078. doi:10.15252/emmm.202217078
- Yoshimoto, S., Okada, E., Umamoto, H., Tamura, K., Uno, Y., Nishida-Umehara, C., Matsuda, Y., Takamatsu, N., Shiba, T. and Ito, M.** (2008). A W-linked DM-domain gene, *DM-W*, participates in primary ovary development in *Xenopus laevis*. *Proc. Natl. Acad. Sci. USA* **105**, 2469-2474. doi:10.1073/pnas.0712244105
- Zahn, N., James-Zorn, C., Ponferrada, V. G., Adams, D. S., Grzymkowski, J., Buchholz, D. R., Nascone-Yoder, N. M., Horb, M., Moody, S. A., Vize, P. D. et al.** (2022). Normal table of *Xenopus* development: a new graphical resource. *Development* **149**, dev200356. doi:10.1242/dev.200356