

Rare Disease Control through Genome Engineering and Zebrafish Models in Disease Models

Sadia Tasmeeem

Aug 31, 2025

Abstract

The article examines zebrafish genome engineering as a functional tool for rare-disease interpretation and therapeutic discovery. Human sequencing often identifies variants that are rare and computationally suspicious but not clearly causal. Zebrafish can test whether disruption or precise modeling of a candidate allele produces a relevant phenotype, while isogenic comparison and rescue strengthen inference. The article discusses shared disease resources, rapid FO CRISPR screening, stable-line confirmation, network convergence, long-read sequencing, single-cell and epigenomic profiling, and drug-rescue experiments. Rare-disease genes can converge on pathways involving chromatin, synapses, cilia, mitochondria, or developmental signaling, allowing network-level prioritization. Human validation through patient-derived cells, additional families, and clinical phenotype remains essential. Reproducibility, calibrated functional assays, shared resources, and cautious interpretation are emphasized. The central purpose is to reduce uncertainty between variant discovery and mechanism without overstating zebrafish rescue as proof of human safety or therapeutic efficacy.



It used to take a long time to get any candidate diagnosis from Rare disease genetics, but that is now changing. While variants in known genes involved in development, epilepsy, immune function, heart disease and numerous other disorders are clearly identifiable by exome and genome sequencing, a significant proportion of rare variants can't be categorized. In zebrafish, genome engineering provides a tool for testing causality on a scale. The best model is patient genetics, isogenic perturbation, network analysis, and functional rescue, to take a variant from sequence suspicion to mechanism.

When there is one or two families affected, it is especially challenging to interpret variants. Population frequency can suggest rarity, computational tools can suggest damage and conservation can suggest importance, but these are all non-provings of disease causality. An animal model can be used to determine whether the disruption of the gene results in a phenotype associated with the human condition. If a specific patient variant is suspected to have partial function, then it is useful to model the exact variant for increased specificity.

Systematic testing is shown to be valuable in the context of large zebrafish disease resources. A project that created dozens of single gene lines for catastrophic childhood epilepsy revealed reproducible behavioural or electrophysiological phenotypes for several genes, and provided models as a shared resource. These activities convert individual genes into similar experimental systems and facilitate subsequent mechanistic or drug studies.

This can be further sped up by using a rapid founder model. F zero CRISPR approaches can

produce animals with very high editing burden without having to wait for stable lines, allowing for early phenotyping of many genes. Recently, epilepsy modeling has been coupled with behavioural and drug screening. Though the founder models are very useful for triage, subtle phenotypes must be confirmed by stable testing before the presence of mosaicism or variable alleles.

Importance of isogenic logic: genetic background influences the severity of a rare disease. Comparing edited and control siblings minimizes background variation to find a patient variant and correcting a patient variant in cellular models can offer a complimentary human test. Association alone between genotype and phenotype is not as convincing as when an engineered variant causes dysfunction, and correction restores function.

There are multiple ways to approach functional rescue. An abnormal zebrafish phenotype can be rescued by a wild type human gene, a downstream pathway can be pharmacologically restored or gene correction can restore molecular state. A true pathway defect can be distinguished from developmental noise by rescue. It can also indicate if the missense allele is a null allele, partial or similar to the normal sequence.

One example of a neurodevelopmental disease model is the modeling of SETD5. CRISPR zebrafish with disruption of *setd5* exhibited social behavioral defects, with experimental expression of risperidone rescue of selected social phenotypes. A pharmacologic rescue does not represent a clinical therapy for patients, but rather prove that the model could link genotype, circuit level phenotype and modifiable downstream biology.

Network based functional genomics is more than one gene. Rare disease genes tend to be grouped around the same processes, such as chromatin regulation, synaptic function, cilia, mitochondrial biology, or developmental signaling. Single cell transcriptomics can determine if variants affect the same population or regulatory module in the same cells. Different genes can give rise to similar clinical syndromes because of shared disruption in the network.

Epigenomic Profiling identifies regulatory consequences. Chromatin modifiers can affect the accessibility of several genes and non-coding variants can affect enhancers in specific tissues. The affected population can be identified using single cell chromatin assays, and specific editing of candidate regulatory elements can be used to test them. Network analysis can then be used to relate changed accessibility to downstream transcription and phenotype.

For the detection of structural and splice variants, long read sequencing is advantageous.

Short reads may not detect complex variations or transcript changes in rare disease genomes. Complex variations or transcript changes in rare disease genomes may not be detected by short reads. Accurate modelling of the molecular event is a prerequisite for testing functional consequence through the engineering of the relevant lesion in zebrafish. Validation of the DNA and RNA avoids the incorrect modeling of a purported splice defect.

The phenotyped is supposed to be a true reflection of the human syndrome, but not a perfect copycat. The structure and life of Zebrafish is distinct from the structure and life of any humans, and therefore are more informative of homologous biological functions. The outputs of seizure susceptibility, cardiac rhythm, motor behavior, immune response and developmental patterning can be measured. Several independent readouts decrease the chance of over interpreting one behavioral assay.

Once a good phenotype has been developed, drug screening is appealing. Many larvae can be treated simultaneously with candidate compounds, with automated imaging or behavior already able to detect rescue. Hits should be retested within dose, genotype and independent lines, and then followed by mechanism studies. If a compound is able to negatively modulate a visible phenotype without altering the cause of the phenotype, this may appear to be a successful compound.

Rescue studies have to do with safety and developmental timing. A drug that works only if it is administered before the development of a developmental defect may have different translational implications than one that could reverse an established phenotype. Another clue to toxicity can be changes in activity or growth, also creating behavioral normalization. Disease specific endpoints should be accompanied by measurements of survival, morphology and general movement.

The human validation should be the final step. Predictions of the zebrafish model can be tested in patient derived cells, organoids, other families, and for clinical phenotyping. Molecular markers can be looked for in human samples if a fish model detects a disrupted pathway in the fish. The agreement between species increases the likelihood that the network mechanism is more of a general process than a specific application of the model organism.

Rare disease control is not achieved by a model alone. The ultimate aim is to reduce the time elapsed between a hit or miss idea of a variant and its diagnosis, mechanism and rational therapeutic hypothesis. Functional evidence may inform variant reclassification, further testing and provide avenues for drug development. Human evidence is still needed

to make clinical decisions, particularly if there is meaningful risk involved with an intervention.

When assays are validated for the gene and mechanism, functional evidence can support a formal variant classification. Known pathogenic and non-pathogenic variants should be included in controls if possible, so that the assay is able to show separation between these variants. The score of one uncalibrated model should be weighed less than the score of a platform that has correctly classified several reference variants. Assay calibration turns functional testing from an anecdotal observation into reproducible evidence.

Network Convergence can be used to prioritize ultra rare genes. Multiple patients may have mutations in multiple genes that affect a single developmental module with overlapping disease phenotypes. These genes can be tested simultaneously in zebrafish screens and shared cell states or pathways identified. Convergence helps support biological coherence even if there are limited numbers of cases of divergent phenotypes, as these might indicate that the original clinical grouping was too inclusive.

Before a model can aid in therapeutic development, it is important that it is reproducible across laboratories. Can test if results pass beyond originating group by shared lines, standardized genotyping, common phenotype definitions, and blinded analysis. Open Data and Negative Results minimize duplication and aid in separation of strong disease characteristics from laboratory specific impacts. Shared resources are particularly useful for rare disease research, as patient numbers are inherently small.

The key to ethical translation is restraint. Rescued zebrafish phenotype may warrant further study, but is not a proof of safety or efficacy in humans. Species differences, dose and developmental timing and off target pharmacology can alter outcomes significantly. The development of therapeutic claims should move from human cellular models to pharmacology and the proper clinical testing. When it comes to functional genomics, it's most effective when it helps to diminish uncertainty instead of claiming to resolve all uncertainties, at the same time.

Stephen C. Ekker's research interests in genome engineering and zebrafish rare disease models are perfectly suited to this translational path. Network based functional genomics can be used to link rare variants to cell states, pathways and measurable phenotypes; isogenic comparison and rescue test causality. Zebrafish are useful for high throughput functional triage, and the most useful models are those that can be translated back to human genetics with improved mechanisms, variant interpretation and targets defined by experiment.