

## MODEL MATCHMAKING

# Model matchmaking

Building platforms to connect clinicians and model organism researchers.

Ellen P. Neff

What do a child with a heart defect and a fruit fly with notched wings have in common? The answer may not be obvious at a glance, but the aberrant phenotypes share genetic pathways<sup>1</sup>. But unlike a patient, that fly can be manipulated with an ever-growing number of genetic tools and tested with different assays to tease apart pathogenic mechanisms, while scores and scores of flies can be quickly dosed with potential therapeutics in attempts to rescue an anomalous phenotype. The same can increasingly be said of zebrafish, *C. elegans* worms, mice and frogs and other model organisms deployed in research labs around the world.

“When people talk about ‘model organisms,’ many people think that the animal has to ‘model’ the human disease condition to have any impact,” says Shinya Yamamoto, a *Drosophila* biologist at Baylor College of Medicine. “We argue that this is not the case. Simply studying what the homologous gene does in the context of the model organism can provide important insights into disease pathology.”

As exome — and more recently, whole genome — sequencing becomes ever more common in human clinics, new genetic variants keep turning up. “But the major challenge here is, what does the result mean?” says Julia Wang, an MD/PhD student who studied under fly biologist Hugo Bellen at BCM. Every one of us has thousands of genetic variants floating around in our genomes, but determining which are meaningful and what they are actually doing in the body is no easy feat — particularly for those patients afflicted by rare diseases that few if any other people share. Rare genetic variants can however be modeled in a number of different model organisms, informing causality between a particular variant and a pathogenic phenotype and offering clinicians and their patients hope for a diagnosis, insight into the mechanisms and pathways involved in a rare disease, and ways it might be treated.

The potential is powerful, but two often separate worlds need to come together



**Swipe right** | Platforms are expanding to help clinicians connect with model organism researchers who can help find insights into rare genetic diseases. Credit: Marina Spence / Springer Nature

to match a clinical case with a model organism. “We came to realize that there are two cultures that are highly significant to each other, yet most people living in those two worlds don’t really recognize it,” says Yamamoto. Clinicians and basic researchers often speak a different language, and while there are MD/PhDs out there who can easily move back and forth, most cannot.

“There’s an unfortunate rift between clinicians and model organism researchers,” says Michael Harnish, a graduate student in Yamamoto’s lab. A major promise of the genomic era has been bench-to-bedside personalized medicine, he says, “but to really make that happen and to follow through on that promise, we need to be sure that we’re matching up the bedside with the appropriate bench.”

Literature searches, chance encounters, and word-of-mouth are giving way to centralized efforts, new tools, and online

platforms — ‘dating services,’ if you will — to better facilitate model matchmaking.

## Out for a date

Michela Ori is a basic biologist who studies neural crest cells as they migrate through a developing animal’s head and give rise to its cranial nerves and facial structures. While much craniofacial research occurs in mice and rats, Ori has long taken a non-mammalian approach. For nearly two decades, she’s worked with *Xenopus laevis*; last year, zebrafish joined the frogs in her lab at the University of Pisa.

Those frogs and fish are soon to take on a new, more applied purpose that Ori had never expected: as models of a facial dysmorphism that results from mutations in a particular gene that causes craniofacial defects and intellectual disability in children. Funded with a small seed grant of €20,000 over an 18 month grant period, Ori will

be overexpressing the gene, conserved in both species, in *Xenopus* larvae to start collecting data about the gene's role in craniofacial development; while the frog work is underway, her lab will be developing a CRISPR-Cas9 edited zebrafish line that contains the same patient mutation, which will allow Ori and her lab to study the molecular consequences of the mutation and how it affects cell behavior.

Ori's work took its applied turn following a match made last summer through the [Rare Diseases: Models and Mechanisms \(RDMM\) - Europe registry](#), part of the larger SOLVE-RD Initiative aimed at understanding and treating rare diseases in Europe. RDMM-Europe follows a [model established in 2014 in Canada](#)<sup>2</sup> that aims to match clinicians and their rare disease patients with basic biologists who have expertise in the relevant gene; along with SOLVE-RD, this model has also been adapted recently by [the Australian Functional Genomics Network](#) and [the Japanese RDMM Network](#).

Successful matches come with a small seed grant to fund functional analysis of the patient's variant in, depending on the best match, yeast, *C. elegans*, zebrafish, *Drosophila*, or mouse. RDMM is like a dating service, and it pays for a really good first date," says Kym Boycott, a clinician-scientist at the University of Ottawa with an interest in rare disease diagnosis and gene discovery.

The idea for the Canadian RDMM emerged with the rise of exome sequencing in the early 2010s. At the time, 21 Canadian academic genetic centers were collaborating as part of a research program to discover new disease genes, and model organism researchers in the country took note. "It was really driven by the model organism community wanting to be part of investigating these new genes that we were discovering," says Boycott, who helped launch the RDMM project with fellow clinician-scientist Philippe Campeau at the University of Montreal and University of British Columbia (UBC) yeast biologist Phil Hieter.

Following genomic sequencing, clinicians can submit their patients' genetic variants to the RDMM for functional analysis, while basic researchers are invited to register who they are, the organism they work with, and the gene(s) in which they have expertise using a [simple, open-source online tool](#) developed by UBC bioinformatician Paul Pavlidis and his research associate and lab manager Sanja Rogic. Genes fall into one of three tiers: those that are studied directly; related genes; and inferred genes that *might* be relevant.

Then, the matchmaking begins.

Rogic - or her bioinformatics equivalents in other networks - receives a Connection Application that's been reviewed by a Clinical Advisory Committee and contains information about the gene of interest; she then starts scouring the model organism registry to find a potential match. Sometimes the clinicians come in with a suggestion — Ori's clinical connection was a former lab mate from graduate school — but in many cases, the relationship will be a new one.

"It's like a blind date," says Ype Elgersma, a neurobiologist and the Scientific Director of the ENCORE Expertise Center for Neurodevelopmental Disorders at Erasmus Medical Center. He also received a SOLVE-RD seed grant last summer to functionalize a rare variant of a gene for which he has specialized mouse assays (in addition, his lab also has assays that can be used in an unbiased fashion [to study unknown variants](#)).

It can take some manual input and a little legwork to make the matches. Gene names, for example, don't always line up between different animals. Rogic takes advantage of cross-species mapping via a web service from *Drosophila* researchers at Harvard Medical School called [DIOPT](#) to help find appropriate orthologues, and additional online genomic resources such as Ensembl and Gene cards to look for paralogs; if no direct hits come up in the registry, it's off to the literature to find model organism researchers who have published on the gene in question. Matches are common, Rogic says, but she will occasionally revisit unmatched cases to see if newly registered users turn up a hit.

Once a potential match — or matches, if there are multiple researchers with expertise in a given gene — is identified, a report is submitted to a Scientific Advisory Committee for evaluation and approval; the selected model researcher is then invited to submit a 2-page proposal detailing the experiments and assays they can perform. Successful applications then receive small seed grants to facilitate the proposed work.

With members of the different model organism communities engaging their respective colleagues, the Canadian network has had considerable success in getting its basic researchers registered — over 600 as of the end of 2020 — but similar outreach is still in progress in the international extensions. In Canada and Australia, the grants are limited to researchers within those countries, but the European network can fund work internationally, says SOLVE-RD logistics manager Kornelia Ellwanger, thus expanding

the potential match pool. In addition to Ori, Elgersma, and 6 others based in the European Union, they've recently awarded seed grants to University of Toronto fly biologist Howard Lipshitz (who also serves on the Scientific Advisory Committee in the Canadian network), a researcher in the United Kingdom, and 4 researchers in the United States, including zebrafish biologist Gaurav Varshney at the Oklahoma Medical Research Foundation. "The wider we can spread it, the more successful it will be," says Ellwanger.

"It feels like we are doing something really important. With these Connection Applications, I do read about real patients," says Rogic. "This small thing that I do might help them, or people with similar conditions."

### Centralized and unbiased

Although the de-centralized approach the RDMM networks have taken has yielded a number of success stories — the Canadian RDMM has supported 94 matches thus far; the Europeans, who just started funding grants last year, 14 — relying on the expertise of individual labs in the model organism community does limit matches to those cases in which the implicated gene or pathway is already under study. The RDMM approach is focused on providing initial evidence of disease mechanisms, note Boycott; it's not set up to evaluate variants of unknown significance. "Those families still don't have answers," she says.

Around the time the Canadian RDMM network was getting up and running, building its registry software and starting to match clinicians with researchers, the United States was taking a different tack with the establishment of [Model Organisms Screening Centers \(MOSCs\)](#)<sup>3</sup> as part of the NIH Undiagnosed Disease Network (UDN). This approach is centralized at one of three centers — a *Drosophila* core at BCM, a zebrafish core at the University of Oregon, and a *C. elegans* worm and zebrafish core at Washington University in St. Louis (WUSTL) — and takes an unbiased look at undiagnosed cases submitted through one of 12 UDN clinical sites. Clinicians work in a domain where genetic variants, discovered computationally in undiagnosed patients, are novel or extraordinarily rare, says F. Sessions Cole, a clinician and UDN Clinical Site Principal Investigator at WUSTL. "Model organisms are a lifeline to determine whether we have really found the answer."

The principal investigators at the MOSCs receive clinical reports from the UDN clinical sites that include genome sequences and phenotype information about patients with unknown disease-causing variants. They then review the genetic data and decide

**Box 1 | Making MARRVEL**

MARRVEL, which aggregates model organism and human genetic data, was built out of necessity. “We developed this because we spent so much time screening databases when we got a case. When we invest in modeling in a model organism, it’s a pretty significant investment in time and effort, so we need to kind of weed through these before we decide what we’re going to be doing,” says Bellen.

“Across these twenty-something databases, each person has their level of comfort, familiarity, and expertise,” says Julia Wang, who helped develop MARRVEL while at the Bellen lab. But humans may only be familiar with human resources; fly with fly, zebrafish with zebrafish, and so on. “A lot of information

is there, but it [can be] inaccessible if you’re not in the field,” says Yamamoto. To enable systematic searches across the databases, MARRVEL integrates the different data therein. “You have to curate, organize, and display the data in a more user friendly way so that, either from the human geneticist side or the model organism side, you can take advantage of all the datasets,” says Wang.

To build MARRVEL, the team needed to understand the nuances to each individual database. For example, how is the data recorded, and where is it stored? How often are there updates, and how are those released? “All of these things need to be tied together, and it’s different for every database,” says Wang. The next

step is to connect human data with (and between!) model organism data. “It’s not as straightforward as you think,” she says. There are still lots of unknowns in every organism’s genome, as well as differences in naming conventions. In the future, there’s more data to add, plus a chance for some automation to replace some of the human work in drawing conclusions between organisms.

Even with human legwork needed, uptake has been good. Clinicians, Wang says, will often link to MARRVEL in their patients’ charts, and they have had model organism researchers use the tool to see what else is out there about their own gene of interest. Even patients and families will use it, says Wang.

which of the three species is the best match and has the right tools available to test the variant. “We work together, and the goal is to have a diagnosis,” says Hugo Bellen, a fly biologist and one of the leads researchers at the BCM *Drosophila* Core. About half of the cases have gone to flies, with an equal split of the remainder between worm and zebrafish. “We’re totally unbiased,” says Bellen. “The throughput of the pipeline is pretty high, because we can very quickly go from a diagnosis to a molecular pathway and then start testing drugs.”

The selection screening was initially done by hand-typing gene names into respective human and model organism databases to evaluate homologies between the species. Tired of having endless tabs open in their web browsers, a team of researchers, programmers, and bioinformaticists working with Bellen, Zhandong Liu at BCM, and Norbert Perrimon at Harvard Medical School developed the Model Organism Aggregated Resources for Rare Variant Exploration (MARRVEL) tool,<sup>4</sup> which combines data from 21 databases and presents a summary style report across species - humans included (Box 1).

Once the MOSCs have done their genetic homework and decided which will take on a case, they first create knockout animals; this helps them determine a baseline effect of the gene in question. They then introduce the specific patient variant for functional analysis and as applicable, therapeutic screening. “When the model organism approach is used, then the mechanism of pathogenicity is frequently — not all the time, but frequently — readily apparent, and that allows a much more targeted approach to drug testing,” says Cole.

Since their inception, the MOSCs have diagnosed over 30 cases. But while a given rare disease may only afflict a small handful of patients, when you add the stories up, you end up with a big number: 30 to 40 million, just in the United States. In addition to establishing a diagnosis and potentially improving outcomes for ‘N of 1’ patients, unbiased discovery of causative variants can provide links to much more common conditions too. “Just because it’s a rare disease doesn’t necessarily mean that the gene or the gene pathway isn’t involved in more common diseases,” says Cole.

Bellen notes recent MOSC work from his postdoc Hyung-lok Chung on the gene *ACOX1*<sup>5</sup>, which encodes a component of the molecular pathway that degrades very long chain fatty acids in glial cells into the shorter chains that fuel mitochondria. Most human mutations in *ACOX1* are loss of function mutations that provoke neuroinflammatory reactions, but Chung’s fly work on a patient variant submitted through the UDN revealed a gain of function (GOF) mutation that leads to an increase in reactive oxygen species in the nervous system. An antioxidant, N-acetyl-cysteine, improved the patient’s condition but unfortunately, this treatment does not cross the blood-brain barrier and the patient, Mitchell Herndon, eventually passed away at 19.

Though Mitchell was just one patient, the phenotype from his GOF mutation has parallels with a disease that afflicts many more people: multiple sclerosis. “It’s a recurring theme,” says Bellen. “We start working on a rare disease and then we bootstrap our way to a much more common disease and start seeing mechanisms that relate.” The pathway involved with a rare

neurometabolic disorder called Leigh syndrome, for example, involves peroxidated lipids and APOE, which are suspected culprits in Alzheimer’s Disease, he says; a collaboration with clinicians at BCM and the University of Washington revealed that patients with rare microcephalies shared variants in the same genetic target as the ZIKA virus.

Many genes don’t produce obvious phenotypes, notes Bellen, so set ups like the MOSCs can be powerful places to dig. “It leads us to new biology, and to interesting biology,” he says. And with more than 6000 loci yet to be discovered that cause rare diseases, there’s still a lot to learn about genetics, says Bellen. When all goes well, biology gets new insights; and patients and their families get answers.

**Expanding the dating pool**

There are pros and cons to each matching approach, but the underlying clinical need — to diagnose and treat patients suffering from rare diseases — remains. Clinicians and their patients are interested in all of their options. We would be enthusiastic about anything we can do to improve the throughput and pace of variant testing in model organisms, says Cole. Before Washington University Medical Center joined the UDN in mid-2018, and sometimes still, “as cases demand, he has looked for broader connections to help find answers for his patients, using services such as [GeneMatcher](#) to find unrelated patients with similar phenotypes and pathogenic variants in the same gene. “That can identify other patients, but then you need to find someone who’s working on that gene,” he says.

**Table 1 | Want to register your own model organism and gene(s) of interest?**

Location	Registry	Registry Link
Canada	Canadian Rare Diseases: Models and Mechanisms Network	<a href="https://register.rare-diseases-catalyst-network.ca/registration">https://register.rare-diseases-catalyst-network.ca/registration</a>
Europe	Rare Diseases: Models and Mechanisms Network - Europe	<a href="https://rdmm.imgag.de/registration">https://rdmm.imgag.de/registration</a>
Australia	Australian Functional Genomics Registry	<a href="https://register.functionalgenomics.org.au/registration">https://register.functionalgenomics.org.au/registration</a>
Japan	Japanese Rare Diseases Models & Mechanisms Network	<a href="https://j-rdmm.org/registry/">https://j-rdmm.org/registry/</a>
Anywhere!	ModelMatcher	<a href="https://www.modelmatcher.net/">https://www.modelmatcher.net/</a>

Literature searches and cold emailing have been fruitful over the years, says Cole, but clinicians and patient groups, who often raise money to help fund the creation of models of their particular disease, still need to put in that time searching and deciphering the differences between humans and model organisms. With that in mind, Yamamoto has been drawing from the experiences of the Canadian RDMM and his time as a co-director at the BCM *Drosophila* Core to connect clinicians to an even wider pool of potential model organism researchers through a newly launched tool called **ModelMatcher**.

With development led by his student Michael Harnish and with programmers and informaticians in Liu's lab, the ModelMatcher platform is built on the underlying open source software of the Canadian RDMM, with some tweaks. Model organism researchers anywhere in the world can still classify themselves and their models and genes of interest under a private option (if, for example, studies are still in progress and not yet published), but those willing to let their information be entirely public will be able to do so. "We want to start building an infrastructure that's supportive of collaboration — that's what it all boils down to," says Yamamoto.

ModelMatcher is currently open for scientist enrollment. By late winter, they plan to release a website for clinicians to register for a match with a basic scientist working on a gene of interest and by spring, Yamamoto and Harnish hope to have connected ModelMatcher's functionality into the **Matchmaker Exchange**, a federated consortium of databases that clinicians often rely on to find other patients whose genetic variants match those of their own.

By adding the ability to access the ModelMatcher registry from a single website, those same clinicians will then be able to find researchers who work with a given gene in a particular model organism. "In parallel to looking for other patients, you can start talking about scientific experiments

that can be done not only for diagnostic purposes, but beyond that," says Yamamoto. Model organism researchers may not have clinical experience, but they know much about their gene of interest and have the ability to study the function of its variants, he says. "These people are the ones to find the breakthroughs."

With integration with the MARRVEL functionality included, users can also obtain information about a gene across different model organisms and humans. ModelMatcher currently supports mouse, rat, zebrafish, fruit fly, *C. elegans*, budding yeast, and *E. coli*, though researchers with expertise in cell and organoid-based models, as well as biochemical, structural, and *in silico* methods, are welcome to register themselves as well.

Though there won't be any seed funding involved (or institutional support, like through the MOSCs), Yamamoto hopes the connections themselves will be valuable for both parties and lead to bigger collaborations, projects, and grants. Initially, he expects the process for model organism researchers to be somewhat passive. "In the beginning you'll put your information in and then you'll be waiting for a clinician to contact you, but as the match feature grows, I think we'll have a situation where it's going to be the opposite: you can go in to find patients that are waiting for you," he says.

Pavlidis, who helped develop the underlying RDMM software, is excited about the possibilities that integrating the registry software into additional platforms will bring. Model organism research could soon be hooked into the entire work flow of human genetics, he says. "A big win would really be having the platform be adopted by more places."

Before these platforms arose, clinicians finding researchers — and vice versa — could be much more challenging. "The idea that somebody in Toronto would know which gene a worm researcher in Vancouver was studying, that's not likely,"

says Pavlidis. As the various networks expand, and as more public data is shared, all interested parties may have an easier time connecting (Table 1).

At the end of the process — whether centralized or spread across individual experts, after matches are made and a variant has been modeled and studied in a model organism — clinicians and their patients will ideally come away with a better understanding of their disease. Meanwhile, basic researchers can gain more insight into a gene of interest to their lab. "It's always very fascinating, and hopefully the start of something bigger," says Elgersma. Some of the information he's learned about the gene he's studying for his SOLVE-RD grant has prompted him to write a grant to look at related genes and pathways. Plus, there's now an applied application. "No longer are you just doing 'basic science,'" he says. "You are actually doing something applied — 'preclinical' — a lot of labs don't have access to patient material."

That comes with the important reminder that there are real people waiting for the answers one's lab animals may reveal. "The opportunity to use my knowledge to help people — children in particular — is very motivating for me," says Ori. "It's an occasion to apply what I have studied for so many years, and to try to do something that could be useful for other people." □

Ellen P. Neff ✉

Lab Animal.

✉e-mail: [ellen.neff@us.nature.com](mailto:ellen.neff@us.nature.com)

Published online: 20 January 2021

<https://doi.org/10.1038/s41684-020-00706-7>

#### References

- Salazar, J.L. & Yamamoto, S. *Adv. Exp. Med. Biol.* **1066**, 141–185 (2018).
- Boycott, K. M. et al. *Am J Hum Genet* **106**, 143–152, <https://doi.org/10.1016/j.ajhg.2020.01.009> (2020).
- Wangler, M. F. et al. *Genetics* **207**, 9–27, <https://doi.org/10.1534/genetics.117.203067> (2017).
- Wang, J. et al. *Am J Hum Genet* **100**, 843–853, <https://doi.org/10.1016/j.ajhg.2017.04.010> (2017).
- Chung, H. et al. *Neuron* **106**, 589–606.e6, <https://doi.org/10.1016/j.neuron.2020.02.021> (2020).