Discovering Genetic Disease Mechanisms

Lindley Darden and John Moult
University of Maryland

Email addresses: darden@umd.edu, jmoult@umd.edu

Key Words: discovery, disease mechanism, mechanism schema, human-computer interaction

Abstract

Knowing what is to be discovered guides the process to discover it. Biologists seek to discover mechanisms. Philosophers of science have analyzed the concept of a biological mechanism and proposed reasoning strategies to guide mechanism discovery. These analyses are now guiding the development of a graphical, web-based representation system to aid humans in discovering genetic disease mechanisms.

Here is an analysis of biological mechanisms: mechanisms are entities and activities organized such that they engage in regular changes producing a phenomenon of interest. A mechanism schema represents a hypothesis about how a mechanism works. Reasoning strategies guide the construction of mechanism schemas, including abstraction/instantiation, modular subassembly, and forward/backward chaining.

This prior work in philosophy of science provides a framework for building a graphical representation system that aids humans in discovering genetic disease mechanisms. What is to be discovered is the mechanism by which a genetic variant produces a disease phenotype (or raises its risk). The system is called MecCog (" Mechanismus Cognito," Latin for mechanism known or discovered). (See the system and examples at www.meccog.org). It provides a web-based, abstract, schematic, representation system for drawing genetic disease mechanisms and linking to the medical literature. The implemented ontology proceeds in stages of biological organization, from the molecular (e.g. DNA, RNA, protein, ligands, and receptors), cellular, tissue, organ, to the organismal phenotype. Each step is a substate perturbation (SSP), deviation from the normal. Each step is driven by a specific causal activity (or a group of productive entities and activities, such as protein synthesis) designated a mechanism module (MM). Icons for each type are available to the schema builder to depict triplets of SSP-MM-SSP. Construction can proceed from the beginning stage, i.e., the DNA variant and proceed forward, examining the activity enabling properties that give clues as to the next step (e.g., the charges in the active site of a protein that enable subsequent binding to a ligand). Alternatively, the schema builder can work backwards from the end, the disease phenotype, to examine activity signatures that indicate what happened in a previous step. Black boxes indicate missing entities and activities, not yet known, where additional research is needed. Drop down boxes provide links to papers via PubMed IDs to indicate the evidence for or against each step. Additional icons indicate sites in the mechanism to which to direct drug discovery.

In sum, the MecCog system implements a representation of a genetic disease mechanism and is available to aid researchers in mechanism discovery. Currently, schemas are being developed for a variety of diseases, including rare disease (e.g., cystic fibrosis), cancer (e.g., Lynch syndrome), and genetic loci associated with complex trait diseases, such as Crohn's and Alzheimer's Disease. Current and future work: develop AI tools to find relevant papers in the scientific literature, to find components of mechanisms, to locate empirical findings that provide evidence for and against the components, and to propose new therapeutic avenues.

References

Craver, Carl F. and Lindley Darden (2013), *In Search of Mechanisms: Discoveries across the Life Sciences.* Chicago, IL: University of Chicago Press.

Darden, Lindley (2002), "Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/Backward Chaining," *Philosophy of Science* 69: S354-S365.

Darden, Lindley (2006), *Reasoning in Biological Discoveries: Mechanisms, Interfield Relations, and Anomaly Resolution.* New York: Cambridge University Press.

Darden, Lindley (2007), "Mechanisms and Models," in David L. Hull and Michael Ruse (eds.), *Cambridge Companion to Philosophy of Biology.* New York: Cambridge University Press, pp. 139-159.

Darden, Lindley (2008), "Thinking Again About Mechanisms," *Philosophy of Science* 75 (5): 958-969.

Darden, Lindley (2018), "Strategies for Mechanism Discovery," in Stuart Glennan and Phyllis Illari (eds.), *Routledge Handbook of Mechanisms and Mechanical Philosophy.* New York: Routledge, Ch. 19, pp. 255-266.

Darden, Lindley, Lipika R. Pal, Kunal Kundu, and John Moult (2018), "The Product Guides the Process: Discovering Disease Mechanisms," in *Building Theories: Heuristics and Hypotheses in the Sciences.* Ed. D. Danks and E. Ippoliti. Springer, pp. 101-117.
Preprint: http://philsci-archive.pitt.edu/13176/

Darden, Lindley, Kunal Kundu, Lipika R. Pal, and John Moult (2018), "Harnessing Formal Concepts of Biological Mechanism to Analyze Human Disease," *PLoS Computational Biology* 14 (12): e1006540.
https://doi.org/10.1371/journal.pcbi.1006540

Kundu, Kunal, Lindley Darden, and John Moult (2021), "MecCog: A Knowledge Representation Framework for Genetic Disease Mechanism," *Bioinformatics 37, (22): 4180–4186.*
https://doi.org/10.1093/bioinformatics/btab432

Machamer, Peter, Lindley Darden, and Carl F. Craver (2000), "Thinking About Mechanisms," *Philosophy of Science* 67: 1-25.