

3rd Conference of the European Philosophy of Science Association

Athens, Greece, 5-8 October 2011

Contributed paper

Section 3. Philosophy of the life sciences

Mendelian population as a model, intended as a “stable target of explanation”

Emanuele Serrelli

University of Milano Bicocca

Piazza dell'Ateneo Nuovo 1, 20126 Milano, ITALY

emanuele.serrelli@unimib.it

Models constitute an increasingly important object of study for philosophy of biology. There is, however, no univocal and sufficiently comprehensive definition of modeling and model (cf. e.g. Downes 1992, Godfrey-Smith 2006, Thomson-Jones in press). Here, I adopt a specific notion of a model as a “stable target of explanation” (Keller 2002), and use it to explore population genetics in a uncommon way.

Mathematical population genetics is unanimously considered as a major pillar of neo-Darwinian evolutionary theory, and commonly referred to as a great set (or “family”) of powerful and elaborate models - whose historical merits include the reconciliation of the Mendelian theory of inheritance with the Darwinian theory of natural selection in 1930s. Arguably, by “mathematical models” one commonly means *equations*, calculating, e.g., gene frequencies or frequency spectrums in a population, or phenotypic change within a

variance-covariance matrix. This notion of a model appears to imply the definition of modeling as an activity of equations specification, tuning, and calculation.

On the contrary, the notion of a model as a “stable target of explanation” I adopt here does not apply to population genetics equations. It fits best what experimental biologists call “model organisms”, like e.g. *Drosophila melanogaster*, or *Caenorhabditis elegans*. Yet, I argue, the notion can be brought to bear onto population genetics. More specifically, it is suitable for capturing Mendelian population, i.e. a formal system whose nature is not equations: Mendelian population is a combination space population genetics equations are *about*. Therefore, one interesting result of my approach is to liken (at least for some epistemological characteristics) a formal system to organic systems, both seen as models in the same sense. A single notion of a model, one that emphasizes model autonomy - with interesting epistemological problems about representation, explanation, and prediction - seems to capture effectively both Mendelian population and model organisms. Models as stable targets of explanation are systems selected for intensive research, yielding their stability and a cost-effective apparatus of experimental resources; they feature some degree of artificiality, and are never exhaustively known, even in case of complete artificiality. My aim is to contribute to the discussion on models and modeling, towards a better clarification of the notions and perhaps, eventually, some clarifying constraints in our use of the terms.

Many authors (e.g. Creager et al. 2007, Morgan & Morrison 1999) insist today on the *autonomy* of model as a direct object of research. Such “stubborn autonomy” is two-

fold: from theory, and from outside data input. A model as a «stable target of explanation» is, by definition, autonomous in both respects. Models are targets of explanation, not immediately tools for explaining (least of all embedded in other kinds of explanation, e.g. theories). For Keller (2002:115) «...the primary meaning of the term model in experimental biology is an organism [...]. A model in this sense is not expected to serve an explanatory function in itself, nor is it a simplified representation of a more complex phenomenon for which we already have explanatory handles. Rather, its primary function is to provide simply a stable target of explanation». As for Mendelian population, its autonomy was effectively expressed by Lewontin when he wrote that (1974:189) «For many years population genetics was an immensely rich and powerful theory with virtually no suitable facts on which to operate. It was like a complex and exquisite machine, designed to process a raw material that no one had succeeded in mining [...] for the most part the machine was left to the engineers, forever tinkering, forever making improvements, in anticipation of the day when it would be called upon to carry out full production». Today, as everybody knows, equations developed for understanding Mendelian population are put to work on molecular and phenotypic data, e.g. for reconstructing phylogeny or population history, or to predict gene frequency.

Autonomy of model from data creates particular epistemological questions (dilemmas?) about representation, explanation, and prediction.

If, as several authors point out (e.g., Creager et al. 2007), models are not chosen because they are typical of a certain set of systems, nor they are built to represent some

other system by reduction, deduction, encoding (Casti & Karlqvist 1989, Rosen 1989) or the like, *how can they ever represent?* Creager and colleagues (2007:2) talk about models *for* (as opposed to models *of*), acting as “representatives”, not “representations”: their similarities to other systems stand to be always better explored, and instances of “unpredictable relevance” might well happen (Hubbard 2007). For Keller «To be sure, a model is expected to bear some resemblance to that which is being modeled, but in science as in art, the degree of resemblance is generally understood to be a matter of perspective. The more critical question is whether it is a “good” model, and in both science and art the measure of how good a model is varies notoriously» (2002:46). Context-dependent similarity has been proposed by several authors (e.g. Giere 1988, Plutynski 2006).

The possible decoupling of model and *explanation* is well expressed in the following quote by John von Neumann: «the sciences do not explain... they mainly make models» (cit. in Creager et al. 2007:11). Models can explain, for von Neumann, only with the addition of verbal interpretation and narratives containing assumptions. Like resemblance, explanation is, for Keller (2002), a context-dependent use of models: what counts as an explanation (and as knowledge) depends on values, so that for example many biologists could reject physical or mathematical models of life as valid explanations (cf. e.g. E.B. Wilson 1934).

Models can also be *predictive* (Oreskes 2007, Dahan Dal Medico 2007), but prediction is, again, outside the notion of a model I am using here.

The fantastic degree of idealization of Mendelian population has been topic of much philosophical literature focusing on representation and explanation relationships Mendelian population should, may, or may not have with other systems (real populations). Such relationships are obviously important, but the point is that discussing them *is not essential within a notion of a model as a stable target of explanation*. That is, if we choose this notion of model we can provisionally remain silent on how and what the model represents and explains. We can concentrate on what the model is, its structure, assuming that it will somehow be an adequate description of some object of study.

According to Creager et al. «A model system in biology refers to an organism, object, or process selected *for intensive research* as an exemplar of a widely observed feature of life» (2007:5, my emphasis). Model-systems constitute answers to the challenge of producing lawlike knowledge in fields such as experimental biology, and their stability is due, amongst other things, to self-reinforcing: «...the more the model system is studied, and the greater the number of perspectives from which it is understood, the more it becomes established as a model system. Even for the many biologists who do not study one of the canonical model organisms, these systems tend to serve as benchmarks and methodological guides when they turn to other organisms and objects as researchers» (Creager et al. 2007:6). Another strength of a model is constituted by the “experimental resources” associated with it, i.e., those elements that ease its study by making it, e.g., cost effective. Simulations and calculations made upon Mendelian population yield theoretical results saving the costs for collection and analysis of data from

organic systems - moreover, the latter would probably never allow the same results. Knowledge of the well-studied model is cumulative, and can be expanded, modulated differently, and re-combined.

A difference between Mendelian population and model organisms seems to lay in its complete *artificiality*. Mendelian population is a formal system built up from the simple rules of inheritance discovered and coded by Mendel, scaled on a large number of factors and alleles, and provided with fitness (a value coming from the Darwinian theory of natural selection). But such architectural artificiality does not rise a clear-cut discontinuity: a degree of (or even complete) artificiality is inherent to models in experimental biology, too. In the history of developmental biology, for example, Keller found that «To many authors writing in the early part of the twentieth century [...] the question of what life is was to be answered not by induction but by production, not by analysis but by synthesis» (Keller 2002:18). So, artificial production does not subtract Mendelian population from the notion of a model suitable for experimental biology. Moreover, artificiality of Mendelian population does not imply that the model is completely known (another feature that would separate it from model organisms). On the contrary, the behavior of Mendelian population is unknown: various, non-exhaustive mathematical methods are devised and tried in order to approach it. Such *inexhaustibility* feature once again relates Mendelian population to model organisms: «...although model organisms are standardized in order to facilitate highly controlled biological experimentation, their inherent complexity means that the systems are

never fully understood and can continue to generate surprising results» (Creager et al. 2007:7).

[1499 words]