Unifying diseases through common genetic mechanisms: the example of the genetic theory of infectious diseases

Marie Darrason

Unifying diseases from a genetic point of view: the example of the genetic theory of infectious diseases

Marie Darrason

Institut d'Histoire et de Philosophie des Sciences et des Techniques (IHPST)-13 rue du Four 75006 Paris Université Paris 1 Panthéon Sorbonne

Email: marie.darrason@gmail.com

Abstract: In the contemporary biomedical literature, every disease is considered genetic. This extension of the concept of genetic disease is usually interpreted either in a trivial or genocentrist sense, but it is never taken seriously as the expression of a genetic theory of disease. However, a group of French researchers defend the idea of a genetic theory of infectious diseases. By identifying four common genetic mechanisms (Mendelian predisposition to multiple infections, Mendelian predisposition to one infection, and major gene and polygenic predispositions), they attempt to unify infectious diseases from a genetic point of view. In this article, I analyze this explicit example of a genetic theory relying on mechanisms and applied only to a specific category of diseases, what we call "a regional genetic theory." I have three aims: to prove that a genetic theory of disease can be devoid of genocentrism, to consider the possibility of a genetic theory applied to every disease, and to introduce two hypotheses about the form that such a genetic theory could take by distinguishing between a genetic theory of *diseases* and a genetic theory of *Disease*. Finally, I suggest that network medicine could be an interesting framework for a genetic theory of Disease.

Keywords: Geneticization. Genetic disease. Genocentrism. Causal Selection. Disease mechanisms. Disease explanation. Disease theory.

Introduction

The concept of genetic disease has gone through several shifts [1-2]. In the 1960s, a paradigmatic example of genetic disease was phenylketonuria, a rare, monogenic Mendelian, hereditary disorder, for which the equation one mutation = one gene = one phenotype was explicitly assumed. But the following years witnessed three major changes resulting in the collapse of this model. First, the core of the concept of monogenic Mendelian disease [3] was called into question by different scientific discoveries, such as allelic heterogeneity (several allelic mutations of the same locus can cause the disease), locus heterogeneity (several genes can cause the same disease), and modifier genes (most monogenic disorders are also influenced by the intervention of other genes) [4]. Secondly, there has been an increasing interest in the genetics of common, non-hereditary and polygenic disorders, such as cancer or diabetes. Finally, the development of bioinformatics, rapid DNA sequencing techniques (such as recombinant DNA technology, sequencing by hybridization and whole-genome sequencing), and "big science" projects such as the Human Genome Project have led to an extraordinary upsurge of genetic data and of gene-disease correlations. So, while the concept of genetic disease originally designated a very restricted class of rare, Mendelian, hereditary, monogenic disorders, nowadays it encompasses common, non-Mendelian, non-hereditary, polygenic disorders to the point where every disease seems to be genetic. Abby Lippman has coined the word "geneticization" to describe this phenomenon of understanding all diseases as the result of genes [5].

Two related but distinct issues arise here: what is a genetic disease and is the geneticization of diseases legitimate? A common strategy for addressing these questions, shared by several philosophers [6-9], is to begin by approaching the project

of defining the concept of genetic disease as an instance of the causal selection problem [10], which consists in picking out the main cause of an event occurring in a multicausal context. Applied to the problem of genetic disease, this means that labeling a disease "genetic" implies that genes are the most important cause in disease explanation. If we understand the concept of genetic disease in the context of the causal selection problem, geneticization can then be understood as an expansion of the concept of genetic disease to all diseases. In that case, geneticization amounts to an acceptance of genocentrism—the belief that genes are the most important causal factor in explaining any biological phenomenon. Genocentrism, however, has already been heavily criticized [6-9]. I will not review here the numerous arguments against genocentrism. It is enough to say that genocentrism seems to be both scientifically unjustified and ethically questionable. Since geneticization, on this causal selection understanding of the term, is identifiable with genocentrism, it follows that geneticization cannot be an acceptable approach to disease. These philosophers, therefore, attempt to give a more restricted account of genetic disease, whereby it addresses the causal selection problem without leading to the pervasive geneticization of disease. They do this by defining the concept of genetic disease strictly in an attempt to distinguish between diseases that are "true" genetic diseases, where genes are necessary and sufficient to cause the disease (usually the Mendelian monogenic diseases), and diseases where gene-environment interactions are more difficult to assess (usually the polygenic disorders).

At this point, some scientific issues with the concept of genetic disease [1-4], to which I alluded above arise again: there are few, if any, "true" Mendelian monogenic diseases and the frontier between monogenic and polygenic diseases gets increasingly blurred. So, while the problems of geneticization and genocentrism are avoided, the

result is an overly restrictive and unclear meaning of the concept of genetic disease. One response at this point might be to abandon the concept of genetic disease. Is it truly useful? Pragmatism is sometimes invoked to explain this lingering attachment to the concept of genetic disease [11], but which pragmatic reasons and pragmatic for whom? I noted above the scientific difficulties encountered with the concept of genetic disease. From a clinical point of view too, it is unclear what purpose the concept serves. It does not guide genetic testing or genetic counseling (where the notion of "inherited disease" is more useful), nor does it define diseases that are targets for genetic therapy (as Caplan has pointed out [12]), nor does it pick out diseases that need special funding because of their rarity ("orphan" or "rare" diseases would be more useful concepts for this purpose).

I noted above how geneticization was abandoned in an attempt to salvage the concept of genetic disease and that the key move in this analysis was to approach the matter through the causal selection problem. However, since there seems to be little point in saving the concept of genetic disease, perhaps one can salvage the concept of geneticization. Of course, I do not mean "geneticization" understood as essentially equivalent to genocentrism, which, as I noted, is subject to significant objections. Rather, I suggest a meaningful interpretation of geneticization that bypasses the issue of causal selection. Rather than interpreting geneticization as an expansion of the concept of genetic disease, I propose that geneticization be understood as the development of a common mechanistic explanation for the genetic side of diseases, what we call "a genetic theory of disease." This account is definitely not genocentrist—by no means would I wish to suggest that genes are the most important factor in causal explanations of all diseases. Indeed, the account I will propose embraces interactionism and

acknowledges a multicausal model of disease causation for every disease. No disease can be understood without appealing to both genes and environment.

But, if the theory stops at this point, it has moved only from genocentrism to weak interactionism, and that is not a very interesting achievement. Indeed, weak interactionism is probably true but certainly trivial: it does not tell us anything truly meaningful about causal explanations of diseases. Furthermore, there is still the lingering temptation to come back to the causal selection problem and to view diseases on a causal continuum where both genes and environment would play a part in causing every disease but where some diseases would still be "more genetic" or "more environmental" than others. That is why, in this article, I propose a first step towards a strong and meaningful interactionism. This strong interactionism asserts that diseases share some common genetic mechanisms in their development and tries to assess which types of genetic mechanisms are at play in disease explanation. In this kind of account, it does not make sense to consider some mechanisms "more genetic" than others. Rather, one can identify various mechanisms that could provide an interesting basis to reclassify diseases according to the type of mechanisms that they exhibit, thus providing a new way of understanding disease causation.

To make sense of the evolution of contemporary biomedical science, the best method may be to take the recent biomedical literature as a starting point. I will focus on the genetic theory of infectious diseases, one of the rare examples of an explicit genetic theory [13]. This theory is defended by a small but renowned group of scientists and aims to unify infectious diseases around the identification of four common genetic mechanisms. Therefore, I will first describe the structure of this genetic theory of infectious diseases before discussing the benefits and limits of this approach. From this genetic theory, which is restricted to a specific class of diseases (a "regional" genetic

theory), I will try to introduce two hypotheses about the form that a general genetic theory could take by distinguishing between a genetic theory of *diseases* and a genetic theory of *Disease*. Finally, I will suggest that network medicine could provide an interesting framework for developing a genetic theory of *Disease*.

The example of the genetic theory of infectious diseases

From germ theory to the genetic theory of infectious diseases

Infectious diseases were born as an independent entity at the end of the nineteenth century with the development of germ theory. This is best captured by the four Henle-Koch postulates [14], which state that for an agent to be considered the infectious cause of a disease, it must fulfill the following conditions:

(1) the agent must be present in all cases of the disease;

(2) the agent must be isolated from someone with the disease and grown in pure culture;

(3) inoculation into a susceptible organism of the agent—from a pure culture—must produce the disease;

(4) the agent must be recovered from the infected–inoculated organism and grown again in culture.

In the years following the establishment of these postulates, several issues raised by germ theory have been pointed out [15]. I will concentrate here on two specific difficulties. First, the third postulate cannot account for the problem of asymptomatic carriers. For example, it cannot explain the fact that of the over one hundred people infected by the influenza virus, only ten will develop the flu. The

"agent" of a given disease can be inoculated in an organism and yet fail to produce the disease in the infected organism. Secondly, these postulates do not address the question of the inter-individual variability of the symptoms. The example of leprosy is particularly telling on this point [16]. Leprosy has two main clinical subtypes: the paucibacillary form and the multibacillary form. Whereas in the paucibacillary form, there is a limited number of hypopigmented and anesthesic lesions without any microscopically discernable bacteria, the multibacillary form exhibits numerous sensitive or anesthesic lesions with high bacillary loads. In the nineteenth century, G.A. Hansen identified the agent responsible for these two forms of leprosy as *Mycobacterium leprae*, thus giving leprosy its other name, "Hansen's disease." How can the same pathogen be responsible for two different clinical subtypes of diseases that receive two different types of treatment and do not have the same prognosis? To some extent, the first problem can be understood as a limiting case of the second: the problem is explaining how the same pathogen inoculated in different organisms can produce so many different subtypes of the same disease, from completely asymptomatic to severe forms.

It is precisely in order to fill this explanatory gap that the genetic theory of infectious diseases has been designed. "The field of human genetics of infectious disease aims to define the genetic variations accounting for inter-individual variability in the course of human infections" [13, p. 915]. Infectious diseases, then, are no longer understood as purely environmental diseases, but also as determined in part by genetic factors, thus stepping out of the monocausal model to an explanation of infectious diseases in general that can fit both the individual and population levels. The genetic theory of disease is not incompatible with and does not try to refute the germ theory. Nor does the genetic theory claim to be a complete picture of inter-individual clinical

variability. In fact, Jean-Laurent Casanova, Laurent Abel, and Alexandre Alcaïs acknowledged at least three other theories that contribute to a global explanation of inter-individual clinical variability. [17, p.404] For instance, the genetics of the microbe itself (some bacteria may carry antibiotic resistance genes or specific virulence genes) should be taken into account for understanding that the "same" infectious disease can be more severe in some individuals than others. Non-microbial environmental factors may also be involved, with air temperature or humidity and the availability of an animal vector being particularly crucial (the ecological theory of infection). Finally, non-genetic host factors, such as age or, in the last century, personal vaccination history may have a key role (the immunological theory of infection) [17, p. 404].

Indeed, the genetic theory of infectious diseases does not even aim to provide a complete picture of the causal factors involved in the pathogenesis of infectious diseases. While acknowledging other possible factors at play, it only focuses on the genetic mechanisms of infectious diseases. To put it differently, the genetic theory of infectious diseases only aims at providing an explanation of "the genetic host-side" of infectious diseases.

The proponents of a genetic theory of infectious diseases

Evidence supporting the genetic theory of infectious diseases first came from observations of familial aggregations of both rare and common infections, and also from follow-up studies of adoptive children and twins. Nevertheless, with the exception of a few diseases, genetic susceptibility to infections was poorly understood until the completion of the Human Genome Project [18].

This may explain why, even if the involvement of genes in the host reaction to infectious diseases was implicitly recognized by every infectious disease specialist,

only a small group of researchers explicitly theorized this involvement. These researchers are mainly Casanova, Alcaïs, and Abel. They worked in the laboratory of human genetics of infectious diseases at Necker Hospital Medicine School in Paris, and they have written approximately thirty articles over the last thirty years to defend this theory. My account of the genetic theory of infectious diseases will rest on two of their most recent and explicit articles [13,19] and on a chapter titled "Human Genetics of Infectious Diseases," which they published in a reference book on human genetics in 2010 [17]. In these three papers, they attempt to unify infectious diseases from a genetic point of view by identifying four genetic mechanisms.

Before describing these mechanisms, however, one should note that the scientists who wrote these papers explicitly use the term "mechanisms," a term that will be discussed in more detail later in this article.

Description of four mechanisms

1. Mendelian predisposition to multiple infections: mutations in one gene cause susceptibility to multiple infections. For example, the X-linked agammaglobulinemia is caused by mutations in the Bruton's tyrosine kinase gene. This gene plays an essential role in the maturation of B cells in the bone marrow. When mutated, immature B-lymphocytes cannot develop into functional B cells, thus causing susceptibility to multiple bacterial infections at early stages of the infected male's life. This is also referred to as "conventional primary immunodeficiency" (conventional PID).

2. Mendelian predisposition to one infection: mutations in one gene cause susceptibility to one infection. Take, for example, Herpes Simplex Encephalitis (HSE): *Herpes Simplex Virus-1* (HSV1) infects around 80% of the population, but only a small fraction will develop HSE, which still remains the most common form of sporadic

encephalitis in Western countries. The diseased people have an autosomal recessive UNC93B deficiency. This deficiency impairs the recognition of RNA intermediates of HSV1 in the central nervous system, resulting in impaired interferon production and causing enhanced viral replication and cell death. The diseases in this category are also called “novel primary immunodeficiencies” (new PIDs), as they were discovered after the conventional PIDs, which were Mendelian predispositions to multiple infections.

3. Major gene / Resistance to one infection: the “major gene” or “major locus” concept was developed in the context of complex segregation analysis in order to understand the phenomenon of incomplete penetrance. Penetrance is the frequency of individuals carrying a particular allele that also express an associated trait. For a given disease causing mutation, penetrance can be incomplete, meaning, only a portion of the people having the given allele will exhibit the corresponding disease. A major gene creates the immunodeficiency, but its penetrance may be lowered due to the combined effect of other genes and the environment. The main assumption is that only one mutated gene causes the corresponding disease but other genes or environmental factors may influence the expression of this gene, thus, explaining its variable penetrance. The concept of “resistance” mirrors the major gene concept: some specific mutations on a given allele confer resistance to a given pathogen because they result in the lack of expression of receptors needed by the invading microbes. For example, consider the case of malaria caused by *Plasmodium vivax*. *P. vivax* is one of the pathogens that cause malaria. To penetrate into the blood cells, *P.vivax* needs to fix onto a Duffy blood group chemokine coreceptor, called DARC and located at the surface of the erythrocytes. A single nucleotide mutation on the promoter of the DARC gene prevents the expression of the DARC receptor at the cell surface, conferring a resistance to malaria caused by *P. vivax*.

4. Polygenic predisposition to one or multiple infections: Polygenic inheritance differs from the major gene concept. It implies that the global phenotype results not from one single gene influenced by other genes or the environment, but from the combined effects of a large number of loci. Depending on the number and relative impact of the genes influencing disease, we may distinguish between oligogenic predisposition and "true" polygenic predisposition. Oligogenicity implies that the phenotype is dependant on two or a few major genes, while other genetic loci make a relatively lower contribution. In "true" polygenic inheritance, no major gene is involved and the occurrence of disease depends on a large number of genetic loci, each making a small contribution. These four mechanisms are summarized in Table 1.

Table 1: Four mechanisms in the genetic theory of infectious diseases

MECHANISM	**DESCRIPTION**	**EXAMPLE**
Mendelian predisposition to multiple infections	One gene, complete penetrance, multiple infections	***X-linked agammaglobulinemia:*** Mutations in Bruton's tyrokinase gene ⇒ immature B lymphocytes ⇒ multiple bacterial infections
Mendelian predisposition to one infection	One gene, complete penetrance, one infection	***Herpes Simplex Encephalitis:*** Autosomal recessive UNC93B deficiency ⇒ impaired recognition of HSV1 by the CNS ⇒ impaired production of interferon ⇒ viral replication in the CNS
Major gene / Resistance to one infection	One major gene, high penetrance, one infection	***Malaria caused by* P. vivax*:*** Mutations in the promoter of DARC gene ⇒ lack of DARC coreceptor of *P.vivax* ⇒ *P. vivax* cannot enter erythrocytes (resistance)

Polygenic predisposition to one or multiple infection(s)	Multiple genes, low penetrance, one or multiple infection(s)	***HLA associated infections***

From common genetic mechanisms to the concept of genetic continuum

Common genetic mechanisms

With four genetic mechanisms at play in the inter-individual clinical variability of infectious diseases, one might expect infectious diseases to be split into four mutually exclusive categories, each defined by its own genetic mechanism. In fact, however, the categories overlap to create a continuum. Indeed, the previously described mechanisms are said to be common, meaning that the same disease can combine two or

three mechanisms. As an example, genetic predisposition to tuberculosis, which was previously considered to be purely polygenic, was recently shown also to reflect a Mendelian predisposition to one infection in some patients and a major gene effect in others [20, 21]. It is precisely because these mechanisms are not the property of a specific category of diseases that there are non-mutually exclusive classes of diseases and that one can talk about a mechanistic continuum and not of a simple typology. The term "continuum" must be understood here in its usual mathematical sense. It indicates that the genetic differences between infectious diseases are not discrete, but just a matter of degree. This concept of continuum is well represented in various figures in the articles of Alcaïs, Abel, and Casanova [13, 17, 19], a version of one of which appears below (Fig. 1).

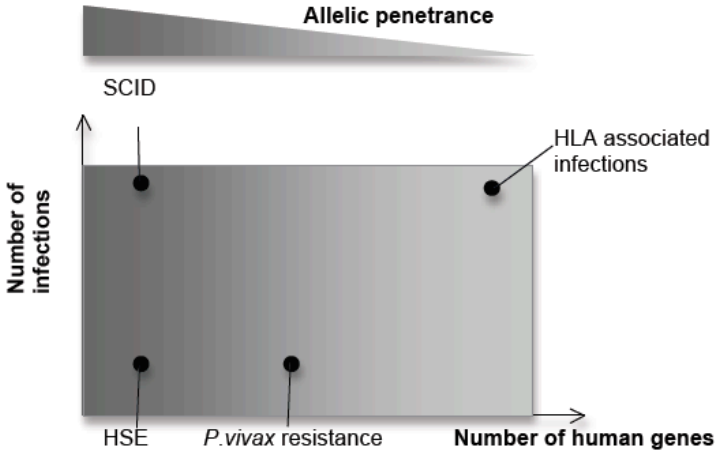

Fig. 1: Schematic representation of the genetic continuum of infectious diseases (adapted from [19]). The ordinate is the number of infections at risk of development. The abscissa is the number of genes at play. Allelic penetrance is represented by the triangle above the graph: when just one or a major gene is involved, penetrance is high. Conversely, when multiple genes are involved, each gene has a limited effect on the global phenotype and penetrance is lower. Finally, the shades of gray (from dark gray to light gray) represent the genetic continuum between infectious diseases and are also highly correlated with allelic penetrance as suggested by the common color code. Four diseases are represented on this graph exemplifying the four previously described genetic mechanisms in infectious diseases.

A new concept of genetic continuum

This is a rather new way of representing the concept of genetic continuum. Genetic diseases are represented in a very different manner, for example, in a textbook published by the National College of French Teachers and Practitioners of Medical Genetics (CNEPGM) in 2004 [22]. The graph, entitled "joint action of genetic and environmental factors in diseases," consists of a single line, made of three segments (black, gray, and white), each corresponding to a specific disease category. The black segment represents "the diseases that are mostly genetic," whereas the white segment represents "the diseases that are mostly environmental." Between these two extremes is a gradation of gray, indicating diseases in which both genetic and environmental factors are at play but in different proportions. In other words, this is a typical representation of a genetic continuum as framed by the "causal selection problem." Some diseases are more genetic than others, and the main issue at stake is to determine the proportions in which genetic and environmental factors interact. In this kind of genetic continuum, the previous examples of infectious diseases will probably be on the "mostly environmental" side of the graph.

The genetic continuum represented in Figure 1 differs from the CNEPGM version in two aspects. First, the causal selection problem is not an issue—there is no distinction between diseases based on how much genes and environment influence their phenotypes. Indeed, the continuous gradation of gray represents allelic penetrance, that is, the fraction of people having the gene(s) and the corresponding disease. A disease with lower penetrance is not a disease with less genetic influence. Indeed, polygenic inheritance does not suppose less genetic influence than Mendelian predisposition. It is only a difference in the way genes cause the disease. In Mendelian predisposition, one

gene is responsible for the disease whereas in polygenic inheritance several genes are responsible together.

Second, whereas it is impossible to distinguish between different causal mechanisms for the same disease on the traditional representation of the genetic continuum, it is absolutely possible in the kind of representations used by Abel, Alcaïs, and Casanova. For example, tuberculosis should appear in at least three different points of Figure 1 since it can be caused by at least three different genetic mechanisms, as previously mentioned.

Consequences of the genetic continuum

What are the epistemological consequences of this mechanistic continuum? First, it provides a unifying explanation of inter-individual clinical variability from a genetic point of view. It is assumed that for each infectious disease, one or more of these mechanisms can be instantiated to explain why a fraction of the infected individuals exhibit symptoms while others stay asymptomatic. So, the mechanistic continuum exhibited by the genetic theory of infectious diseases represents an important gain in our understanding of the pathogenesis of infectious diseases, compared to the germ theory, which did not provide an explanation for the phenomenon of inter-individual clinical variability and could not account for the problem of asymptomatic carriers or variations in the symptoms exhibited by individuals.

Not only does the genetic continuum of infectious diseases provide a unifying background that accounts for inter-individual clinical variability, it also provides a satisfactory explanation of infectious diseases at both the individual and population levels. The germ theory could only provide an explanation for individuals who have tuberculosis because they have been infected by *Mycobacterium tuberculosis*. On the

other hand, the genetic theory of infectious diseases allows two kinds of explanations. At the population level, it allows a general account of every genetic mechanism implied in the predisposition to a given disease. In this case, some individuals develop tuberculosis because they have either a Mendelian predisposition to *Mycobacterium tuberculosis* or a major gene effect. But at the individual level, it would be theoretically possible to distinguish between these different mechanisms to explain why an individual, in a particular case, developed tuberculosis.

Is this genetic theory a real mechanistic explanation?

Preliminary comments on the concept of "mechanism"

Now that I have presented the main contents of the genetic theory of infectious diseases, I will comment on the term "mechanisms," which is explicitly used by Abel, Alcaïs, and Casanova but may raise some justified concerns for those who are familiar with the recent debates about the definition of mechanisms and their relevance to biological explanations [23]. These debates refer to Machamer, Darden, and Craver's definitive characterization of mechanisms, which comes from their classic paper published in 2000, "Thinking about Mechanisms" [24]. The so-called "MDC account" of mechanisms has become the received philosophical view of mechanisms in recent years, superseding other attempted definitions. The account has even crossed over into the scientific community, making the original MDC paper the most-cited paper ever published in *Philosophy of Science*. The MDC account characterizes biological mechanisms as "entities and activities organized such that they are productive of regular changes from start or set up conditions to termination or finish conditions" [24, p. 3].

The genetic theory of infectious diseases does, to an extent, have the entities called for in the MDC account in the form of genes or diseases. It also seems that a regular organization between entities and activities is assumed in each mechanism between genes and the development of diseases. Still, these are not the entities and forms of organization expected when one speaks of "genetic mechanisms." When talking about "genetic mechanisms", one expects to be confronted with molecular activities, such as DNA replication and transcription, regulatory networks of gene expression, and so on. Do the mechanisms described above really deserve to be called "genetic" mechanisms? Indeed, are they even specific enough to be considered mechanisms at all?

Imprecise activities, missing entities and problematic concepts

There are three specific critiques that can be seen to expand on the questions raised in the last paragraph. First, as I pointed out, the described mechanisms are imprecise. For a genetic mechanism, one may expect a detailed molecular description. For example, the description of the fourth mechanism, that is, the polygenic predisposition to one or multiple infections, clearly remains vague. Indeed, the identification of a truly polygenic predisposition requires a large number of individuals, both because of the small expected effect attributable to each gene and because of the additive nature of these genetic effects. This may explain why evidence of such genetic mechanisms at both the population and individual levels has only been provided by studies of susceptibility to infectious diseases in animal models of experimental infectious diseases, and has yet to be provided by human studies. The description of the third mechanism, "major gene / resistance to one infection," suffers similar shortcomings. Very little is said about how other genes and the environment may affect the expression of the major gene.

Secondly, one could argue that the genetic theory of infectious diseases does not take into account some entities involved in the inter-individual clinical variability of infectious diseases, especially the genetics of the microbiome and the genetics of the pathogen [25]. On the one hand, the microbiome is the complex community of bacteria, archaea, eukaryotes, and viruses that infect humans and live permanently in our bodies. It is firmly believed that the genetics of this microbiome interacts with our immune system, thus modulating its response to infections. On the other hand, the genetics of the pathogen itself are probably of great importance to understanding inter-individual clinical variability; different individuals of the same pathogen species do not necessarily carry the same type of resistance to antibiotics, the same genes of virulence, etc. It is not that the genetic theory of infectious diseases developed by Casanova, Abel, and Alcaïs is not incompatible with these theories; it just does not mention them.

Thirdly, one may question the concepts chosen for describing these mechanisms. Indeed, concepts such as "Mendelian predisposition" or "monogenic" are borrowed from classical human genetics. But, as has already been suggested above, these concepts are not as straightforward as they may seem, since several of them have been challenged recently. Indeed, non-Mendelian modes of inheritance have been discovered [26], and monogenic disease is no longer considered a simple category [27].

Mechanism sketches?

These objections are not so much obstacles to a mechanistic description of the genetic theory of infectious diseases as a problem of the degree of explanation—molecular mechanisms are not so much absent here as implicit. What the proponents of the genetic theory of infectious diseases propose is neither the explanation of a specific case of genetic susceptibility for a given infectious disease (in which case the described

entities and activities would be more specific), nor is it a complete general description of the molecular level of each mechanism (in which case one could expect some general schema to describe each mechanism). As the theory itself is a work in progress (some of these mechanisms, such as "Mendelian resistance," have only been recently described), the description is necessarily incomplete. It still constitutes, however, what Craver would describe as a "mechanism sketch."

A sketch is an abstraction for which bottom out entities and activities cannot (yet) be supplied or which contains gaps in its stages. The productive continuity from one stage to the next has missing pieces, black boxes, which we do not know yet how to fill [24, p.18]. Thus, the mechanisms of the genetic theory of infectious diseases seem closer to mechanism sketches than to a complete mechanistic description of genetic susceptibility in infectious diseases. However, even mechanism sketches have a purpose: they constitute heuristic tools designed to indicate what further work needs to be done to get a better mechanistic explanation. More importantly, the incompleteness of this theory does not weaken my main argument, as I am not so much interested here in the genetic theory of infectious diseases itself as in the conclusions that one can draw from such an example of a regional genetic theory.

What about the genetic theory of diseases in general?

What about a genetic theory of disease in general?

The genetic theory of infectious diseases is an example of what a regional genetic theory, that is, a genetic theory that applies only to a specific group of diseases, could look like. With the genetic theory of infectious diseases, I considered the example of a genetic theory devoid of genocentrism that relies on four common genetic

mechanisms (or mechanism sketches) to unify infectious diseases from a genetic point of view. However, the aim was not the extension of the concept of genetic disease to infectious diseases, but its extension to any disease. What is of interest is a genetic theory of disease in general, which could apply to any category of disease. One can consider two ways to progress from this example of a regional genetic theory to a more general theory. One is to progress to a genetic theory of *diseases*, and the other is to move on to a genetic theory of *Disease*.

My distinction between a genetic theory of *diseases* and a genetic theory of *Disease* derives from Paul Thagard's history of medical theories [28]. In this history, Thagard makes a clear distinction between "ancient" medical theories and modern ones. Every ancient theory, such as humoral medicine, traditional Chinese medicine, or traditional Indian medicine, relies on a general definition of Disease as an imbalance (even if the nature of this imbalance differs from one ancient theory to another). Conversely modern medicine emerged with the development of the microbial theory that identifies a specific cause for a specific class of diseases. Later, other specific theories for other classes of diseases arose, giving birth to our current medical theory, which is a collection of different theories for different classes of diseases. So there is a clear opposition between the ancient theories, which are general theories of Disease in this respect, and our modern medical theory, which is a collection of distinct theories for different classes of diseases. The distinction between "diseases" and "Disease" does not bear any ontological commitment. It only aims to distinguish between two different kinds of disease explanations: explaining diseases as distinct, individual, and separate entities, or trying to find common biological features to the concept of disease.

Representing a genetic theory *of diseases* and a genetic theory *of Disease*

By applying the distinction between a general theory of Disease and theories of diseases made by Thagard to the search for a genetic theory, one ends up with two different possibilities—a genetic theory *of diseases* and a genetic of *Disease*, as represented in Figures 2 and 3.

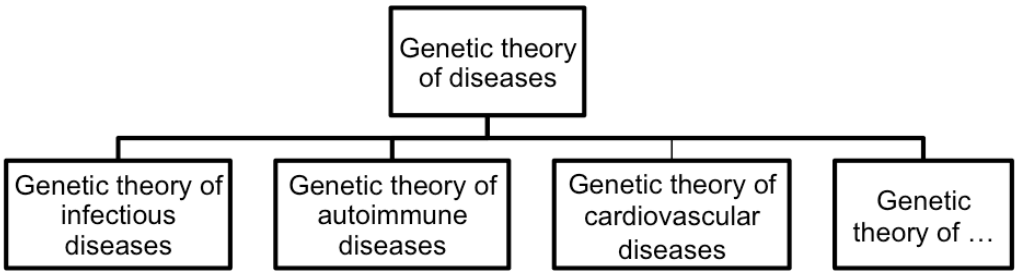

Fig. 2: Typical representation of a genetic theory *of diseases*. The genetic theory of *diseases* is a set of regional genetic theories. For each category of diseases, there is a specific genetic theory with specific mechanisms. Genetic mechanisms may differ for each class of diseases. This kind of theory does not change the way we classify diseases.

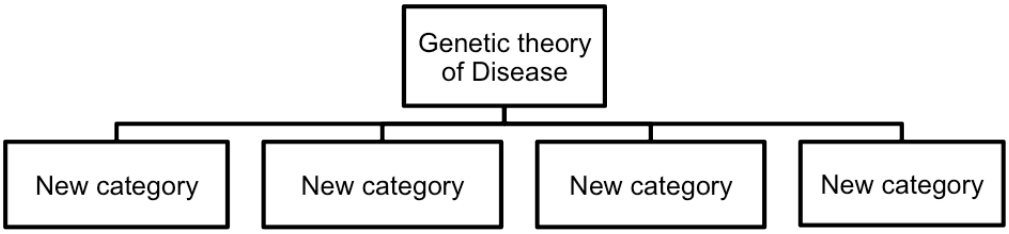

Fig. 3: Typical representation of a genetic theory *of Disease.* In a genetic theory of Disease, we may expect a genetic definition of Disease in general. Depending on this definition, some classificatory principles would appear and these principles would likely renew the way we classify diseases.

The first diagram (Fig. 2) is a representation of what we may call a genetic theory of *diseases*. It is a set of regional genetic theories, an extension of the example of the genetic theory of infectious diseases. Each regional theory would be defined either by distinct and specific genetic mechanisms (each regional theory would have its

own genetic mechanisms) or by applying the same kinds of genetic mechanisms for each regional theory. These mechanisms could use similar concepts as those in the example of the genetic theory of infectious diseases, provided these concepts (e.g., Mendelian inheritance, monogenic disease) have been clarified in the meantime. In this approach, each category of diseases as we know it now (autoimmune diseases, infectious diseases, cardiovascular diseases, and so on) would remain the same, except for the elucidation of the genetic components of their physiopathology.

The second diagram (Fig. 3) represents what would be a genetic theory of *Disease* in general. In such an approach, it is the very definition of Disease that is likely to change and to receive a genetic interpretation. If there were a genetic definition of disease, one would expect radical changes in the way diseases are classified. For example, one might expect the re-classification of the disease categories, as currently known, into new subclasses of diseases that are yet to be defined. The idea of a genetic definition of Disease is still theoretical, but there are some hints towards such an idea, e.g., in network medicine [29].

Network medicine: a genetic theory of Disease?

Network theory is a set of solid mathematical and computational methods developed to decipher the underlying architecture behind apparently anarchic networks, such as the World Wide Web, social networks, and biological networks. Network medicine was born of the synthesis between network theory, genomic medicine, and systems biology. Network medicine aims to develop network-based approaches to disease by analyzing the interactions between different kinds of genomic networks in a given disease and between apparently distinct pathophenotypes.

The core of network medicine relies mainly on two biological properties of the cell: interconnectivity and functional redundancy [30]. The interconnectivity of the cell components implies that disease can never been understood as the result of a single mutation in a single gene. On the contrary, disease is defined as a perturbation in a functional module, that is, in a complex network of intra- and extra- cellular components (genes, transcription factors, proteins, etc.) that interact to achieve a specific function. But a single perturbation in a functional module does not necessarily imply the occurrence of the disease. Indeed, the disorganization of a functional module does not necessarily lead to its inactivation. It can also lead to a rerouting of the function or to a less efficient achievement of the function. Moreover, most cellular functions depend not on a single functional module but on several—a property called "functional redundancy"—which contributes to the robustness of the function. If a single mutation or a single environmental perturbation could breakdown a functional module, humans would be permanently ill. But there is some robustness in the way our bodies are able to adjust to a certain level of stress and genetic or lifestyle-induced perturbations. Based on this functional redundancy, disease can be defined in a more specific way. A disease is a dynamic and complex phenomenon that occurs with the progressive inactivation of several functional modules initially used to achieve the function.

In what sense can network medicine be considered a theory of Disease? First, this definition of disease is supposed to apply to most (if not all[1]) diseases. There is a common definitional framework for every disease, which is the first requirement of a theory of Disease, unlike a theory of diseases, which is merely a collection of disease

[1] The application of this framework is easy to imagine for most monogenic and complex diseases, including the infectious diseases I discussed above. It may prove difficult for some specific cases, such as environmental poisoning or brutal accidents. That said, defining these cases as diseases may itself be problematic.

classes whose mechanisms or explanatory frameworks may differ from one class to another.

Second, from such a perspective, the *explanandum* of disease explanation changes. Our current classification of disease delineates diseases based on similar phenotypes and symptoms, neglecting the different ways in which the same disease can occur. Network medicine, by contrast, aims to identify disease in a more specific and sensitive way. For each disease a functional subnetwork (the entire set of redundant functional modules) is identified and based on this modular identification, a disease can be defined in its preclinical state and in an unequivocal way. Moreover, the aim is not to explain separately the occurrence of every disease but to understand how diseases are functionally related to each other. Diseases themselves have intertwined relationships and are understood as functionally related entities, since different diseases may share some components in the composition of their module and the failure of one functional module in a disease A can have an influence on the disorganization of one of the functional modules of disease B. It is based on these hypotheses that the proponents of network medicine hope to explain not only comorbidity (for example, the relationship between obesity, metabolic syndrome, and cardiovascular diseases) but also syndrome families and the extraordinary importance of some genes in common diseases [31]. Searching for a common origin to different individual diseases can thus be considered a step toward a theory of Disease.

Third, network medicine may completely change the way we classify diseases [32]. What matters here is not the main organ disturbed by diseases, as in most of the anatomo-clinical classifications, nor is it the identification of a main cause (infectious, genetic, autoimmune). What matters is the identification of a given module composed of genetic and non-genetic components at the cellular level. It is still not clear on which

classificatory principles network medicine would rely since a functional subnetwork is supposed to be specific for a given disease. Therefore, to some extent, classifying diseases into classes or categories does not make sense and each disease is a class of its own (identified by a unique functional subnetwork). Still, in this respect, network medicine seems closer to a theory of Disease, which is supposed to renew our disease categories, than to a theory of diseases, which keeps our disease classification and our disease categories intact.

This explains why network medicine might be considered a theory of Disease and not a theory of diseases. But in what sense is this theory of Disease genetic? And in what sense is it a general framework, or at least a first step toward a strong interactionism in disease explanation?

Both genocentrism and weak interactionism approach the multicausal model of disease explanation as a binary choice between genes and environment, with options being defined in a rather loose way. In network medicine, this multicausal model is refined, and genes and environment are defined in a stricter way, thus offering a more fine-grained causal background for an interactionist disease-explanation. Human disease genes are all those genes known to be involved in diseases, but not every gene is a human disease gene. For example, essential genes that are involved in key cellular functions or key developmental features cannot be human disease genes, since mutations in these essential genes are usually lethal *in utero*. There are other biological properties of human disease genes [33, 34], but I cannot review them here in detail. I only want to observe that network medicine accounts for the fact that not all genes have the same functional role in a cell [35]. Different types of environment are acknowledged as well. For example, some proponents of network medicine distinguish between the environment E, which designates external environmental modifiers

commonly shared between individuals close to each other (such as nutriments, bioclimatic conditions, or pollutants), and the environment E', which designates a more internal environment, depending on the individual history, epigenetics, intrinsic stochasticity, and which is strictly independent of the genotype [30]. This distinction between E and E' is necessary to understand how two monozygous twins raised in a similar environment may have a different set of functional modules at some point in their life. Not only does network medicine provide a framework for us to redefine the initially unsatisfying dichotomy between genes and environment in disease explanation, it also allows a redefinition of the distinction between Mendelian monogenic diseases and polygenic disorders. Indeed, the causal selection problem was deeply entangled with an unsatisfying account of Mendelian monogenic diseases, in which genes are necessary and sufficient for the occurrence of disease, and of polygenic diseases, in which genes and environment interact in a more complex way. In network medicine, Mendelian monogenic diseases are understood to have low redundancy and weak robustness, and polygenic disorders are understood to have high redundancy and strong robustness. This explains why just a few genetic mutations can lead to the occurrence of Mendelian monogenic diseases while many mutations and environmental perturbations are necessary to trigger polygenic disorders; and it does so without compelling us to consider monogenic diseases to be "more genetic" than others.

While one cannot assume that network medicine, a field in its infancy, has all the characteristics of a genetic theory of Disease, it seems promising. One issue that remains to be addressed is whether such a theory can be given a full and explicit mechanistic account, given the dynamic and complex relationships that exist between the different components of the functional modules and the interactions that exist between these functional modules and the different types of environmental background.

Conclusion

As long as it is embedded in a misdirected quest to deem genes the most important causal factor in disease causation, the current geneticization of diseases cannot be interpreted as anything other than an unsatisfactory expression of genocentrism. This essay presents an alternative interpretation of geneticization, wherein a strong interactionist model underlies a unified mechanistic explanation for the genetic side of diseases.

Despite being closer to a mechanism sketch than a full mechanistic model, the genetic theory of infectious diseases supports this reinterpretation. It unifies infectious diseases from a genetic point of view through the identification of common genetic mechanisms and achieves a better explanation of the pathogenesis of infectious diseases than the germ theory previously did, while acknowledging a multicausal model in disease explanation.

Eventually, this genetic theory of infectious diseases may be considered a heuristic tool to imagine two different types of genetic theory. On the one hand, one would have a "genetic theory of diseases" as a set of regional genetic theories, where each category of disease could exhibit some specific mechanisms or where the same genetic mechanisms could apply to every category of disease. In this case, medicine would keep the same subclasses of diseases that are already in use (infectious diseases, autoimmune diseases, cardiovascular diseases, etc.). One would have a unified explanation for the genetics of each disease category but that would not change the conceptualization of disease. On the other hand, one could have a "genetic theory of

Disease" with a new genetic definition of disease and a reclassification of every disease into new disease categories. Network medicine might offer the conceptual framework for developing such a genetic theory of Disease.

Acknowledgements: This article is an expanded version of a paper presented at the International Philosophy of Medicine Roundtable in San Sebastian, Spain. I have greatly benefited from many discussions with the participants of the roundtable, especially with Fred Gifford. This paper also owes much to the many discussions with and very helpful comments from members of the Institut d'Histoire et de Philosophie des Sciences et des Techniques (IHPST), namely, Jean Gayon, Thomas Pradeu, Maël Lemoine, and Hélène Richard. Finally, I would like to thank Julia Bursten for kindly correcting my English mistakes.

References

1. Melendro-Olivier, Sarah. 2004. Shifting concepts of genetic disease. *Science Studies* 1: 20-33.

2. Lindee, Susan M. 2000. Genetic Disease since 1945. *Nature Reviews Genetics* 1: 236-241.

3. Scriver, Charles R., and Paula J. Waters. 1999. Monogenic traits are not simple: lessons from phenylketonuria. *Trends in Genetics* 15: 267-272.

4. Dipple, Kathrina M., and Edward R.B. McCabe. 2000. Modifier genes convert 'simple' Mendelian disorders to complex traits. *Molecular Genetics and Metabolism* 71: 43–50.

5. Lippman, Abby, 1994. The genetic construction of prenatal testing: choice, consent or conformity for women? In *Women and prenatal testing: facing the challenges of genetic testing*, ed. Karen H. Rothenberg and Elizabeth J. Thompson, 8-34. Columbus: Ohio State University Press.

6. Magnus, David. 2004. The concept of genetic disease. In *Health, disease and illness*, ed. Arthur L. Caplan, James J. McCartney, and Dominic A. Sisti, 233-242. Washington D.C.: Georgetown University Press.

7. Gifford, Fred. 1990. Genetic traits. *Biology and Philosophy* 3: 327-347.

8. Smith, Kelly C. 2007. Towards an adequate account of genetic disease. In *Establishing medical realities*, ed. Harold Kincaid and Jenifer M. McKitrick, 83-100. Dordrecht: Springer.

9. Hull, Richard T. 1979. Why genetic disease? In *Genetic counseling: facts, values and norms*, ed. Alexander M. Capron, Marc Lappé, Robert F. Murray, Jr., Tabitha M. Powledge, and Sumner B. Twiss, vol. 15, no. 2 of March of Dimes Birth Defects Original Article Series, 57-69. New York: Alan R. Liss, Inc.

10. Mackie, John L. 1965. Causes and conditions. *American Philosophical Quarterly* 2: 245-264.

11. Gannett, Lisa. 1999. What's in a cause? The pragmatic dimensions of genetic explanations. *Biology and Philosophy* 14: 349-374.

12. Caplan, Arthur L. 1992. If gene therapy is the cure, what is the disease? In *Gene mapping: using laws and ethics as guides*, ed. George J. Annas and Sherman Elias, 128-141. New York: Oxford University Press.

13. Casanova, Jean-Laurent, and Laurent Abel. 2007. Human genetics of infectious diseases: a unified theory. *EMBO Journal* 4: 915-922.

14. Evans, Alfred S. 1976. Causation and disease: effects of technology on postulates of causation. *Yale Journal of Biology and Medicine* 64: 513-528.

15. Stewart, Georges T. 1968. Limitations of germ theory. *Lancet* 291: 1077-1081.

16. Alter, Andrea, Audrey Grant, Laurent Abel, Alexandre Alcaïs, and Erwin Schurr. 2011. Leprosy as a genetic disease. *Mammalian Genome* 22: 19-31.

17. Alcaïs, Alexandre, Laurent Abel, and Jean-Laurent Casanova. 2010. Human genetics of infectious diseases. In *Vogel and Motulsky's human genetics (problems and approaches)*, ed. Michael R. Speicher, Stylianos E. Antonorakis, and Arno G. Motulsky, 403-416. Heidelberg: Springer-Verlag Berlin/Heidelberg.

18. Kwiatkowski, Dominic. 2000. Science, medicine and the future: susceptibility to infection. *British Medical Journal* 321: 1061-1065.

19. Abel, Laurent, Alexandre Alcaïs, and Jean-Laurent Casanova. 2009. Human genetics of infectious diseases: between proof of principle and paradigm. *Journal of Clinical Investigation* 9: 2506-2514.

20. Abel, Laurent, and Jean-Laurent Casanova. 2000. Genetic predisposition to clinical tuberculosis: bridging the gap between simple and complex inheritance. *American Journal of Human Genetics* 67: 274-277.

21. El Baghdadi, Jamila, Orlova Marianna, Andrea Alter, et al.. 2006. An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. *Journal of Experimental Medicine* 203: 1679-1684.

22. Collège National des Enseignants et Praticiens de Génétique Médicale. 2004. *Génétique médicale, formelle, chromosomique, moléculaire, clinique*. Paris: Masson.

23. Bechtel, William, and Adele Abrahamsen. 2005. Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 421-441.

24. Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science* 67: 1-25.

25. Virgin, Herbert W., and John A. Todd. 2011. Metagenomics and personalized medicine. *Cell* 147: 44-56.

26. Badano, José L., and Nicholas Katsanis. 2002. Beyond Mendel: an evolving view of human genetic disease transmission. *Nature Genetics* 3: 779-789.

27. Dipple, Katrina M., and Edward R.B. McCabe. 2000. Phenotypes of patients with 'simple' Mendelian disorders are complex traits: thresholds, modifiers and systems dynamics. *American Journal of Human Genetics* 66: 1729-1735.

28. Thagard, Paul. 2006. What is a medical theory? In *Multidisciplinary approaches to theory in medicine*, ed. Ray Paton and Lauren A. McNamara. 47-62. Amsterdam: Elsevier.

29. Barabasi, Albert-Làszlò, Natali Gulbahce, and Joseph Loscalzo. 2011. Network medicine: a network-based approach to human disease. *Nature Genetics Reviews* 12: 56-68.

30. Debret, Gaëlle, Camille Jung, Jean-Pierre Hugot, Leigh Pascoe, and Annick Lesne. 2011. Genetic susceptibility to a complex disease: the key role of functional redundancy. *History and Philosophy of the Life Sciences* 33: 497-514.

31. Oti, Martin, and Han G. Brunner. 2007. The modular nature of genetic diseases. *Clinical genetics* 71: 1-11.

32. Loscalzo, Joseph, Isaac Kohane, and Albert-Làszlò Barabasi. 2007. Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Molecular Systems Biology* 3:124.

33. Tu, Zhidong, Li Wang, Min Xu, Xianghong Zhou, Ting Chen, and Fengzhu Sun. 2006. Further understanding human disease genes by comparing with housekeeping genes and other genes. *BMC Genomics* 7: 31.

34. Chavali, Sreevenas, Fredrik Barrenas, Kartiek Kanduri, and Michael Benson. 2010. Network properties of human disease genes with pleiotropic effects. *BMC Systems Biology* 4(78): 1-11.

35. Jimenez-Sanchez, Gerardo, Barton Childs, and David Valle. 2001. Human disease genes. *Nature* 409: 853-855.