Scientific exploration of causation in biomedical research: The case of gene targeting on mouse embryonic stem cells

[Draft]

[This paper will be presented at ISHPSSB 2023 conference in Toronto on 14 July.]

Ruey-Lin Chen
Department of Philosophy
National Chung Cheng University

Abstract

This paper aims to explore scientific exploration of causation in biomedical science. "Scientific exploration" means searching for new information, data or knowledge beyond what is known about some objective by plural means. I use this concept in contrast to "scientific explanation," because a scientific exploration does not answer a why-question. A scientific exploration is a scientific action or practice that involves not only exploratory experiments but also other means such as background knowledge and ideas, the production of research tools, standardized operational procedures, database searches, animal modeling, the realization of assumptions, and analogical reasoning. A scientific exploration of causation aims to discover new information, data or knowledge about a particular causal relationship. It can also determine whether a supposed causal relationship between a factor and a phenomenon exists or not. Hence, a scientific exploration of causation involves the establishment of a criterion of causation, which biomedical scientists frequently appeal to Robert Koch's postulates. To flesh out a scientific exploration of causation, I offer a case study of gene targeting on mouse embryonic stem cells to illustrate, in particular, employing this technique to produce gene knockout mice that express the symptoms of cystic fibrosis and to establish a mouse model for human disease of cystic fibrosis.

Keywords: Scientific Exploration, Experimentation, Causation, Koch's postulates, Gene-Targeting, Gene Knockout, Animal Models

1. Introduction

The topic of causation is one of the hottest issues in both metaphysics and the philosophy of science. To date, however, most work in the philosophy of science aims to provide a general theory of causation, focusing on connecting the precise understanding of the causation concept with scientific explanation and causal reasoning. Philosophers of science seemingly pay little attention to the establishment of causation criteria in scientific practice, especially in biomedical research.

The main goal of biomedical scientists concerned with causation is to search for causal relationships between some causes or factors and some effects of interest so that they can solve problems or eliminate undesirable effects (say,

diseases) rather than explain those phenomena. How do they find a causal relationship between a specific cause and a specific effect? According to what criterion of causation do they derive a reliable claim that a specific factor is the real cause of an effect of interest? I call the general action that biomedical scientists conduct for their goal and practice "scientific exploration of causation." This paper aims to explore the details of the scientific exploration of causation by supplying a case study of gene targeting on mouse embryonic stem (ES) cells to illustrate those details.

A scientific exploration of causation is a compound action that integrates background knowledge and ideas, the production of research tools, operational techniques, standardized procedures, exploratory experiments, database searches, animal modeling, the realization of assumptions, and analogical reasoning. This kind of action aims to explore unknowable causal relationships between some specific factors and their possible effects. It not only includes exploratory experiments that have been discussed by philosophers in recent years (Brian 1998, Stern 1998), but also involves other non-experimental procedures that have been listed above. The kind of action is best illustrated by using gene targeting on mouse ES cells to produce gene knockout mice to investigate functions or dysfunctions of specific genes (Capecchi 2005).

Investigating a function or a dysfunction of a specific gene is exploring the causal relationship between the targeted gene and some unknowable expression, for example, a proper function or a dysfunction in mice. To explore such a kind of causal relationship, biomedical scientists have first to select a doable gene in the mouse genome from searching public databases or by means of gene-sequencing and then produce gene knockout mice by means of the homologous recombination technique (Hall, Limaye, and Kulkarni 2009; Doyle, et. al. 2012).[1] In producing gene knockout mice and observing the expressions (whether including some normal function or some dysfunction) of the produced mice, biomedical scientists may identify a causal relationship between a functional feature and the targeted gene. An actual case is the establishment of the mouse model for cystic fibrosis by gene targeting in 1992 (Snouwaert et al. 1992). This kind of investigation is exploratory since what functional features will be expressed is unknown before investigation. In other words, this is that biomedical scientists establish a reliable causal relationship between some specific expression (e.g. cystic fibrosis in mice) and a specific gene (the murine cystic fibrosis transmembrane conductance regulator gene) through a scientific exploration of causation. Why can such a causal relationship be established reliably? The key is the method of gene targeting for knockouts.

A gene knockout refers to the technique according to which a targeted gene in an organism is deleted from or made inactive in the genome of the organism. The technique offers sufficient evidence for judgments of causation, which biomedical scientists frequently appeal to Robert Koch's postulates of causation in order to

[1] Scientists sometimes use the method of gene knock-in. A gene knock-in refers to the technique that involves inserting an exogenous gene (in general, a DNA sequence) into a targeted locus of a genome in an organism or substituting an original gene in the genetic locus with another new one. To focus on the gene knockout, however, this paper puts the gene knock-in aside.

justify (Falkow 1988; Fredrick and Relman 1996). To fit with the requirements of their particular cases or the new development, however, biomedical scientists also revised Koch's original postulates once again. A revised version of Koch's postulates used as an exploratory procedure of causation in molecular genetics may be formulated as: (1) Loss of a function/feature associated with a mutated gene should be discovered. (2) Specific inactivation or deletion of the normal gene should lead to loss of the function/feature. (3) The specific functional difference between the individuals with the modified gene and the individuals with the normal gene should be observed. I argue that this and other revised versions of Koch's postulates can serve as valid criteria of causation well for scientific explorations of causation in biomedical science. More details will be discussed in section 5.

This paper will go as follows. Section 2 develops a general characterization of scientific exploration, comparing it with the conception of exploratory experimentation. In section 3, I introduce the case of the gene targeting on mouse embryonic stem cells, investigating how biomedical scientists to identify a causal relationship between a specific gene and a specific expression of a mouse by means of the gene knockout technique. The actual example is the construction of a mouse model for the human cystic fibrosis disease. Section 4 argues that such an investigation is the very scientific exploration of causation, because it not only fully satisfies the general characterization of scientific exploration but also implies a criterion of causation. In section 5, I show that biomedical scientists revise Koch's postulates of causation to fit with their requirements in exploring causal relationships via the gene-targeting technique on cystic fibrosis.

2. Scientific Exploration

"Scientific exploration" is a concept extended from that of exploratory experimentation. In the philosophy of science, the conception and practice of exploratory experimentation has been discussed since Burian (1997) and Steinle (1997). Burian (1997) provided an analysis of a historical case (Jean Brachet's localization of nucleic acids) in detail to exemplify exploratory experimentation, but he did not provide a general characterization. Steinle (1997) gave a preliminary characterization of exploratory experimentation by comparing it with "theory-driven experimentation" which is "done with a well-formed theory in mind, from the very first idea, via the specific design and the execution, to the evaluation." In contrast, exploratory experimentation "is driven by the elementary desire to obtain empirical regularities and to find out proper concepts and classification by means of which those regularities can be formulated." (Steinle 1997: 70) Steinle's (1997, 2002) distinction between theory-driven experimentation and exploratory experimentation has become a standard frame in categorizing experimentation (Franklin, 2005: 888-889; O'Malley, 2007: 339; Burian, 2007: 286-288).

The distinction between theory-driven and exploratory experimentation seems to hint that the latter might be done without theories. Are exploratory experiments always done without theories? There are a few disputes over the role of theories play in exploratory experiments (Franklin 2005, O'Malley 2007, Colaço 2018). Some philosophers such Waters emphasize that such distinction does not mark a sharp division, as the latter is not free of theory (Waters, 2007: 277-279). Waters

notes that the difference between being theory-informed and being theory-directed and that the distinction between exploratory and theory-driven experimentations is made in the way under which an experiment depends on theories (p. 277). Nevertheless, the philosophers who develop exploratory experimentation agree that the two distinctive categories of experimentation, although not sharp, largely work for methodological analyses.

Elliot (2007: 324) develops a taxonomy of exploratory experiments, discerning different kinds along the three relatively independent dimensions: aims of experimental activity, role of theory in the activity, and methods or strategies for varying parameters. According to his taxonomy, "testing a hypothesis" is neither an aim of nor plays a role in an exploratory experiment. The aims of exploratory experiments include (1) "identifying regularities and developing new concepts," (2) "isolating or manipulating particular entities or phenomena," (3) "developing experimental techniques, instrumentation, or simulations," and (4) "resolving anomalies."

One still wonder whether or not there could be an experiment or a series of organized experiments that can be used to test hypotheses and to explore novel things. If there is one, then how should we should characterize it? That imagined experiment might test some hypothesis derived from a theory, falsify the hypothesis, find anomalous phenomena, and then enter into an unknown field and become exploratory. Thus, we should say that the imagined experiment is both theory-driven and exploratory. Furthermore, resolving anomalies may bring new information/data or knowledge which can be used to revise the tested theory and to guide other new explorations. The entire procedure goes beyond experimentation, interweaving experiments with theorizing, observing, and other scientific activities. Practicing the procedure iteratively can bring many new findings. To jump out the dichotomy of theory-driven and exploratory experimentation and to accommodate with the iterative procedure, I suggest a new conception of scientific exploration that is extended from that of exploratory experimentation.

Gelfert (2018) extends the concept of "exploration" to modelling and uses the case of the reaction diffusion models in the study of biological pattern formation to explore the exploratory function of modelling. He argues that "exploration should stand alongside explanation, prediction, and representation as a core function of scientific models." (Gelfert 2018: 246) Mättig (2022) suggests a general concept of exploration, as he says that "[e]xploration is a key scientific practice and percolates all scientific fields and methods." He lists many examples in modelling, theory development, experimentation, instrument development, observation, and natural history (Mättig 2022: 1-2). However, neither Gelfert nor Mättig give "exploration" a general characterization. To understand scientific exploration deeply, I will provide a general characterization and make sense of this characterization by a case study.

I use "scientific exploration" to refer to the action of searching for new information, data or knowledge beyond what is known about some objectives (including phenomena, regularities, properties, relations, structures, mechanisms, concepts, instruments, methods, and maybe others) by plural scientific means (involving experimentation, observation, measurement, modeling, hypothesizing,

theorizing, and so on). Hence, a scientific exploration aims at discovering or findings. Scientific discoveries are the end of scientific explorations, and scientific explorations in turn are the means, actions, and/or processes of scientific discoveries.

A scientific exploration does not involve an explanation, because it is not required to answer a why-question. A scientific exploration is an action to *answer the question of whether or not there are new information, data or knowledge beyond what is known about some objective.* Confirmation of a novel prediction can be part of a scientific exploration, because the confirmation produces new data or knowledge about the predicted phenomenon or event. Furthermore, we may use a theory to lead an exploration of new knowledge if we apply the theory to some unknown domain – this is *theory-driven exploration*. Differing from the distinction between theory-driven and exploratory experimentation, I rather suggest a new contrast between scientific exploration and scientific explanation, given many scientific activities aim at exploration rather than explanation. The aim of scientific explorations does not involve understanding the world. Nevertheless, a scientific exploration may be followed by a scientific explanation.

Scientific explorations differ from non-scientific explorations such as an ocean exploration, a geographical exploration, mining, an oil exploration, a space exploration, and so on, because those non-scientific ones aim at obtaining new things and/or interests rather than new knowledge. Scientific explorations are essentially epistemic explorations, motivating to search for new information or knowledge. Even those explorations with the help of scientific equipment are non-scientific if the explorers' goal and motive is not epistemic. Of course, some non-scientific explorations, for example, Darwin's voyage round the world by taking the Beagle, may accompany with a scientific exploration. The goal of the Beagle went around the world was not epistemic, but Darwin utilized the voyage of the Beagle to collect samples of organisms for his epistemic goal. Hence, Darwin's exploration was scientific. In addition, non-scientific explorations may bring new information/data about new things or objects. For example, ocean explorations in the 15th century brought much new information/data about new lands or territories and motivated scientific and epistemic explorations. Nonetheless, the obtaining of knowledge is not the goal that motivates those explorations. That is, the difference between the motivating goals distinguish between scientific explorations and non-scientific explorations.

In this paper, I focus on the scientific exploration of causation, which is a subkind of scientific exploration. I will flesh it out by a case study of gene targeting on mouse embryonic stem cells for the human cystic fibrosis. Next section starts to introduce the gene targeting technique and its application to human diseases.

3. Gene targeting on mouse embryonic stem cells for knockouts

Using mice as model organisms to investigate processes, causes or mechanisms of human diseases and other topics in the biomedical science has a long history over one hundred years (Rader 2004, Ericsson 2013, Gurumurthy and Lloyd 2019). However, it was not until the 1990s that scientists began using mice to investigate human diseases related to genes. Due to a revolutionary technique of

gene targeting which could remove a gene of interests from the genome of a mouse, scientists have the ability to modify the mouse genome. Scientists cultured gene-modified mice whose specific genes was "knocked out" from the stage of embryonic stem cells and observed the process of physiological defects expressed on the experimented mice. In doing such a kind of experiments, they built mouse models for genetic diseases such as cystic fibrosis (Snouwaert et. al. 1992) or diseases related to genes such as essential hypertension (Smithies and Maeda 1995) and cancer (Donehower, et al. 1992). They could not do such a kind of experiments without the new technique of gene targeting on mouse embryonic stem cells (ES cells in brief) via homologous recombination.

The technique of gene targeting was developed by Mario R. Capecchi, Martin J. Evans and Oliver Smithies, who were awarded the Nobel prize in 2007, in the late 1980s. It applies the mechanism of homologous recombination in the process of cell division to alter a targeted gene by inserting a marking sequence to a target locus of the gene. The result is either deleting or inactivating the targeted gene in the cases of the so-called gene knockout or introducing an exogenous gene to the genome at a specific locus in the cases of the so-called gene knock-in.

Homologous recombination is a genetic phenomenon first discovered by Thomas Hunt Morgan in the early 1910s. Morgan called it chromosomal crossover, because, in order to explain the data from the experiments with fruit flies, he hypothesized that two corresponding fragments in two homologous chromosomes exchange with each other (Morgan et al. 1915). Chromosomal crossover results in a recombination of genetic information. Later molecular geneticists recognized that the recombination of genetic information is a general phenomenon occurring in almost all organisms. They called the phenomenon "homologous recombination". There are two major functions in homologous recombination. One function is responsible for the generation of variation: organisms produce new combinations of DNA sequences in the process of reproduction, resulting in variation in offspring. The other function helps repair defects of DNA, because a sister strand of two homologous DNA strands can offer correct DNA sequences for the other strand that contains defective sequences.

The technique applying the mechanism of homologous recombination can efficiently produce gene-modified mice for a variety of biomedical research, which is characterized by the following two stages (Capecchi 1994, Cepecchi 2005). The first stage consists of the four steps:

1. Engineer copies of a gene in the test tube to produce targeting vectors by inserting a neomycin resistance gene (neo^r) into a targeted gene. Engineer the vector by attaching the herpes virus thymidine kinase (HSV-tk) gene at its one end. The neo^r gene can interrupt the targeted gene in the homologous chromosome and inactivate the gene on the one hand; and it also serves as a marker to indicate the cells that successfully absorb the vector by way of homologous recombination on the other hand. The HSV-tk gene serves as a marker to indicate the cells that randomly incorporate the vector.

2. Use the DNA calcium phosphate co-precipitate to introduce a vector into cells extracted from a mouse embryo.

3. Introducing the targeting vector into mouse ES cells results in three kinds of outcomes. (1) The vector is inserted into the target site at the homologous chromosome. The neo^r gene carried by the vector is successfully inserted into the targeted gene and inactivates it, which amounts to "knocking it out." (2) The vector is randomly incorporated into some chromosome so that the chromosome contains the neo^r gene and the HSV-tk gene. (3) The vector is not inserted into any chromosome.

4. Use two kinds of drug to make the positive-negative selection. The first kind of drug, a neomycin analogue (G418), is lethal to the cells that do not carry the neo^r gene. The second kind of drug, ganciclovir, can kill the cells that contain the HSV-tk gene. After conducting the procedure, only the cells containing the neo^r gene inserted by homologous recombination survive and proliferate.

The foregoing steps demonstrate that scientists can modify a targeted gene or make a targeted mutation in the genome of a mouse cell via homologous recombination. The next stage is to develop a method that can alter the genome of living mice. Given the result in the first stage, a targeted gene in the genome of an ES cells obtained from an early mouse embryo has been modified. On the basis of the result, scientists utilize the pluripotency of ES cells to proliferate mice. The pluripotency makes ES cells have the capability to generate all types of cells and to develop into complete mice.[2] In order to do this, scientists have to culture ES cells with the modified gene in dishes. After producing newborn mice, scientists explore what phenotypic features would occur on the modified mice and infer the causal relationship between the modified gene and some feature. The second stage contains the following steps.

5. Inject the ES cells obtained from a strain of mice with a brown coat (the brown strain of mice) into the embryos in the blastocyst stage obtained from a black strain of mice. The ES cells contain a targeted gene which has been modified according to the steps in the first stage.

6. Inject the blastocyst stage embryos into the black strain of mice that serve as surrogate mothers. Some of the newborn mice are brown with black stripes. They are chimeras that contain cells derived from two different strains of mice.

7. Mate chimeric males with the black strain of females according to the rules of Mendelian heritance. In the offspring, the proportion of the black mice to the brown mice is 1 to 3. Examine the genes in the brown mice and select those animals that inherit the modified gene.

8. Select the males and the females carrying the modified gene to mate with each other. All of the offspring produced by them carry the modified gene

[2] Regarding the philosophical research on embryonic stem cells, see Fagan (2013).

and thus lack the normal function. Carefully observe what physiological and behavioral abnormalities are expressed on the mice.

If some human disease is caused by lacking some gene, then scientists can build a mouse model for such a human disease by employing the technique of gene targeting to knock the gene out from the mouse genome. In the first step, scientists have to investigate whether or not mouse has a gene that is functionally like the human gene related to the disease. They then have to produce the gene knockout mice and to investigate what functions or dysfunctions would be expressed on those mice. They finally infer unknown details of the aimed human disease from the research of the mouse model. They can also develop therapeutic methods or drugs by using the gene-knockout mice as experimental animals. Below I will discuss the case of human cystic fibrosis.

Cystic fibrosis is a genetic disorder that mostly occur on children who inherit the lineage from Northern European ancestry. This disease affects the respiratory, digestive and reproductive systems and causes the symptoms like difficult breathing, coughing up, intestinal obstruction, putty-like meconium, liver cirrhosis, and infertility in males (Snouwaert et. al. 1992; Davis 2006; for a brief description, see Craver and Darden 2013: 189-193). Physiologists discovered a denser level of salt in the sweat of the patients with this disease than in that of normal humans. They hypothesize that the patients have a dysfunction because of a kind of protein that can control the transport of chloride ions into and out of cells. This kind of protein was called cystic fibrosis transmembrane conductance regulator (CFTR) that is encoded by the CFTR gene, which had been discovered in 1989. The scientists discovered that a mutation in the CFTR gene result in malformed protein and cause the disease of cystic fibrosis. At the same time, they located the CFTR gene on the long arm of the 7^{th} chromosome in human cells and identified the mutation with a three-base deletion in the DNA. This deletion results in the lacking of the amino acid at position 508 (Delta F508) and produce the abnormal protein (Rommens et. al. 1989: 1059-1065).

Despite the discoveries of the CFTR gene and its mutation, the complete mechanism of this disease had not yet been revealed in the 1990s. Many puzzles and blanks required to be solved and filled up. Using model organisms (say, mouse) to do experiments offers a great help in grasping the details in the pathogenic mechanism of cystic fibrosis. Moreover, if physicians want to find therapeutic drugs or methods, they have first to do animal experiments before human experiments. This indicates the importance of creating gene-modified mice by targeting the murine CFTR gene. How did scientists this?

First of all, a murine gene that is equivalent to the human CFTR gene had been discovered by means of gene-sequencing in 1991 (Tara et al. 1991). The codes of the equivalent DNA are transcribed by the correspondent RNA and then translated to synthesize the protein sequence that shares 78% identity with the human cystic fibrosis transmembrane regulator. Nevertheless, the scientists concluded that "[t]he functions of CFTR and its mouse homolog have yet to be determined." (Tara, et al 1991: 306) However, the scientists further assumed that deleting a phenylalanine residue of the mouse protein would produce a murine mutation corresponding to the human mutation Delta F508a. They pointed out

that these findings allow using the technique of gene-modification to make a "cystic fibrosis mouse" (Tara et al. 1991). The scientist Oliver Smithies and his team soon produced the gene-modified mice that have the correspondent mutation and constructed a mouse model for cystic fibrosis by the technique of gene targeting via homologous recombination in 1992 (Snouwaert, et al., 1992). In the research, they recognized that many pathological features of young human cystic fibrosis patients, including difficulty to grow, meconium ileus, alteration of mucous and serous glands, etc. also occurred in the produced mice that usually died before 40 days of age. The team also investigated whether or not the chloride transport was abnormal if the murine CFTR gene were interrupted and provided as affirmative answer (Clarke, et al. 1992).

4. Why is the research of mouse models by gene targeting a scientific exploration of causation?

The aim of a scientific exploration of causation directs toward finding new information, data or knowledge about a particular causal relationship. It is usually performed to answer the question of *whether or not there is a causal relationship between a factor and a phenomenon*. Sometimes the factor is known and the phenomenon is unknown or undetermined; sometimes the factor unknown or undetermined and the phenomenon known; and sometimes both are unknown or undetermined. In all situations, the supposed causal relationships are unknown or undetermined. A scientific exploration of causation is thus used to determine *whether or not a supposed causal relationship really exists*. Hence, a scientific exploration of causation always involves the establishment of a criterion of causation. In many cases, a criterion of causation is embedded in the exploratory methods or procedures used to discover the supposed causal relationship. This paper aims to explore an exploratory action or procedure of causation, focusing on causal relationships that possibly exist between genes and features/functions, which are illustrated by the case of gene targeting on mouse ES cells for knockouts.

The creation of the gene-knockout mice for cystic fibrosis by the technique of gene targeting *per se* is a research of animal experiment. This mouse experiment presupposes the following background assumptions: (A1) There is a gene equivalent to the human CFTR gene in mice, which is responsible for the production of the murine CFTR protein. (A2) The murine CFTR regulates the transport of chloride ions into and out of cells in mice. (A3) Altering or inactivating the murine CFTR gene would result in the abnormal CFTR and the abnormal transport of chloride. (A4) The abnormalities further result in the disease of cystic fibrosis in mice, which would be displayed on the symptoms similar to that of the human CF. The four background assumptions imply two groups of causal relationships. A1 and A2 imply the first group of causal relationships between the CFTR gene and its multiple effects, including the product of the normal CFTR protein and the normal chloride transport. A3 and A4 imply that the second group of causal relationships between the altered CFTR gene and its multiple effects, including the abnormal CFTR protein product and the abnormal chloride transport.

All these assumptions in turn stemmed from the background knowledge of the human cystic fibrosis disease: the symptoms, the disease as an autosomal recessive disease, the partially pathological mechanism, the identification of the genetic

factor (i.e., the CFTR gene), the location of the CFTR gene, the mutation that results in the abnormal CFTR protein, the identification of the mutation, etc.

Under those background knowledge and assumptions, the scientist Oliver Smithies and his team's conducted the experiment of the CF mice by the technique of gene targeting. This experiment derived the following three outcomes: (C1) The first outcome is *the realization of the four background assumptions* A1-A4 and thus the determination of the function and dysfunction of the murine CFTR and their genetic factor (i.e., the CFTR gene). (C2) The second is *the production of experimental animals as tools* for the investigation of the pathological mechanism of CF in mice. (C3) The third is *the construction of a mouse model* for the human CF and *an analogical inference* from the mouse model of CF to the human CF. All the outcomes are consequences of scientific explorations. Below let me explain why this is so.

One may want to say that Smithies' experiment of the CF mice confirmed the hypothesis about the pathological mechanism of the human CF, because the hypothesis guides the experiment and derives the assumptions. Saying such may be too hasty. In fact, the assumptions were not derived from the hypothesis about the disease, because the experimented objects are mice rather than humans. They are only assumptions about what may happen in mice, given the background knowledge about the physiological similarity of mice to humans and about the human CF. People did not know whether or not the equivalent CFTR gene in mice would function as the human CFTR gene before Smithies' experiment. People did not even know whether or not there is a murine gene equivalent to the human CFTR gene till the discovery by means of gene sequencing in 1991. The nature of Smithies' experiment is to create the CF disease in mice and investigate the causal pathway of the disease from the murine CFTR gene to the symptoms on the experimented mice (i.e., the chloride transport) rather than to confirm a give hypothesis of the human CF. Although the scientists conducted their experiment under the background assumptions A1-A4, the aim of the experiment is to see *whether or not the assumptions might be realized rather than to test some hypothesis about the human CF*. In consequence, the assumptions were realized by the experiment. According to the analysis above, we may well say:

(1) The experiment provided new data and knowledge about the murine CF disease by realizing the background assumptions. The data and knowledge are part of all knowledge about CF, helping puzzle out the complete mechanism of the CF disease in human and mouse. In this sense, the experiment is exploratory.

Smithies' experiment should be seen only as part of the research of the CF mice from the investigations of the genetic factor in the human CF disease to the experimental creation of the CF mice. Of course, the research of the CF mice is exploratory in the following more senses:

(2) In the experiment, the scientists intervene a specific factor/cause that has been targeted (i.e., the murine CFTR gene) and observes what effects would be resulted by this intervention. If the assumed effects (the symptoms of CF) occur, then the assumption about the causal relationship between the factor and the effects is realized and a new causal relationship in mice is discovered. Hence, the aim of

experiment is not providing a causal explanation for a given phenomenon. (In fact, the effect was unknown before the experiment). In this sense, it is exploratory.

(3) The research of CF offers a causal inference on a causal pathway from the abnormal regulation of the chloride transport to the symptoms of the human CF disease. However, the inference is analogical because it infers from the mouse model to the human conditions. In other words, the hypothesis that the CFTR regulates the chloride transport in human bodies is not confirmed because this research did not experiment on human subjects. On the contrary, the hypothesis established by the mouse modeling was waiting to be tested by some experiment on human cells or subjects. In this sense, it is exploratory but not only an exploratory experiment. It also makes modeling, (analogical) reasoning, and hypothesizing beyond mere experimenting.

(4) The research of CF mice establishes a standard method for other similar investigations and to help develop new therapeutic drugs or methods in the future, as the scientists who discovered the murine CFTR gene say that "[a]n animal model for cystic fibrosis would be useful in elucidating the role of CFTR in the normal and disease states and for testing new treatment modalities." (Tara et al. 1991: 307). All the aims are exploratory rather than explanatory.

The basis of the research of CF mice is the gene targeting technique that was developed from a series of experiments. The aim and nature of the technique is also exploratory. Applying the background idea and principle of homologous recombination, it explores the applicable potentiality and range of an available tool (mouse as an experimental organism) and the gene targeting technique. The mouse tool in turn is used to explore unknown causal relationships between genes and phenotypic features in mice and humans. In fact, the scientists who develop the gene targeting technique directly use the verb "explore" to describe their intention. For example, Capecchi (2005: 507) says that "I was exploring whether I could introduce DNA into nuclei of mammalian cells using extremely small glass needles." This description is completely consistent to my characterization of exploration: A scientific exploration is used to answer the question of whether or not there are new information, data or knowledge beyond what is known about some objective.

5. Molecular Koch's postulates as an exploratory procedure of causation

Why scientists judge that there is a causal relationship between a targeted /modified gene and a specific feature in experimented mice? According to what criterion or condition can scientists do such a judgment? Biomedical scientists frequently revisit the pathologist Robert Koch's postulates of causation. Discussing and revising Koch's original postulates to fit with new developments in biomedical science also forms a tradition of explorations on causation (Evans 1976, Falkow 1988, Fredricks and Relman 1996, Walker, LeVine, and Jucker 2006, Byrd and Segre, 2016).

In 1890, Robert Koch proposed his well-known postulates as guidelines or standards to demonstrate there is really a causal relationship between a bacterium

and a disease. According to Fredrick and Relman (1996: 19), Koch's original postulates may be formulated as follows.

> (K1) The parasite occurs in every case of the disease in question and under circumstances which can account for the pathological changes and clinical course of the disease.
> (K2) The parasite occurs in no other disease as a fortuitous and nonpathogenic parasite.
> (K3) After being fully isolated from the body and repeatedly grown in pure culture, the parasite can induce the disease anew.

Koch's follower Friedrich Loeffler reformulated Koch's original postulates as a series of operable steps that can be performed to demonstrate a causal relationship. According to modern scientists Walker, LeVine, and Jucker (2006: 2), the operational procedure is formulated as follows.

> (L1) The organism must be shown to be invariably present in characteristic form and arrangement in the diseased tissue.
> (L2) The organism, which from its relationship to the diseased tissue appears to be responsible for the disease, must be isolated and grown in pure culture.
> (L3) The pure culture must be shown to induce the disease experimentally.
> (L4) The organism should be re-isolated from the experimentally infected subject [this postulate was added after Loeffler.]

Biomedical scientists apply Koch's postulates in a very concrete way. If they found some assumed pathogens could not satisfy any of the postulates, they would revise Koch's postulates to fit with their cases. Fredricks and Relman (1996) offer an excellent discussion on a number of actual cases. For example, Thomas Rivers proposed revised Koch's postulates for viruses because viruses could not be cultured to induce diseases. The postulates about viruses are (R1) "a specific virus must be found associated with a disease with a degree of regularity;" and (R2) "the virus must be shown to occur in the sick individual not as an incidental or accidental finding but as the cause of the disease under investigation." (Fredricks and Relman 1996: 20-21). However, Rivers' postulates are not very operable. Fredricks and Relman (1996: 30) develop a set of exquisite guidelines of DNA sequence-based microbial identification methods for establishing causal relationships between microbes and diseases. In addition to those cases, Walker, LeVine and Jucker (2006: 3) investigate diseases from infectious proteins and adjust L1-L4 to fit with proteinaceous pathogens.

Can Koch's postulates be revised to fit with the cases of gene manipulation in genetic engineering? Falkow (1988: 274) proposed a set of molecular Koch's postulates for the demonstration of causal relationships between genetic variations and diseases. He formulates the following five guidelines.

> (F1) The phenotype or property under investigation should be associated with pathogenic members of a genus or pathogenic strains of a species.

(F2) Specific inactivation of the gene(s) associated with the suspected virulence trait should lead to a measurable loss in pathogenicity or virulence.
(F3) Reversion or allelic replacement of the mutated gene should lead to restoration of pathogenicity.

The three guidelines revised from Koch's original formulations don't involve the method of gene manipulation, so Falkow further suggests the following two alternatives (Falkow 1988: 274).

(F2a) The gene(s) associated with the supposed virulence trait should be isolated by molecular methods. Specific inactivation or deletion of the gene(s) should lead to loss of the function in the clone.
(F3a) The replacement of the modified gene(s) for its allelic counterpart in the strain of origin should lead to loss of function and loss of pathogenicity or virulence. Restoration of pathogenicity should accompany the reintroduction of the wild-type gene(s).

The molecular Koch's postulates F1, F2a, F3a suggested by Falkow are an operational procedure for exploring the causal relationship between a gene (usually a mutated gene) and a disease. The procedure indicates to inactivate the (mutated) gene to eliminate or reduce the supposed disease. In other words, Falkow's postulates indicates a procedure of gene therapy. All revisions of Koch's postulation show that the establishment of *a standardized procedure of operations* in a scientific exploration of causation.

The molecular Koch's postulates function well for many cases of genetic diseases in pathology. In the case of cystic fibrosis, however, scientists inactivate the normal gene (the murine CFTR gene) to make the disease of CF in the experimented mice. Hence, I reformulate the guidelines for such cases as CF and other cases in which expressed functions or features are not diseases by reference to Falkow's molecular postulates of causation:

(M1) Loss of a function/feature (in particular, a disease) associated with a mutation of a gene should be discovered in some individuals of a kind of organisms.
(M2) Specific inactivation or deletion of the normal gene should lead to loss of the function/feature (in particular, expressing a disease) in the experimented organisms (e.g., the knockout mice).
(M3) The replacement of the modified gene for its allelic counterpart in the strain of origin should lead to disappearance of the functional loss. Restoration of the normal function/feature should accompany the reintroduction of the normal gene.

Does the case of employing the gene knockout method to produce CF mice satisfy the guidelines M1-M3? The answer is affirmative. First, there is *really* a gene equivalent to the human CFTR gene in mice, which is responsible for the production of the murine CFTR protein. Hence, a murine CFTR gene associated with and the production of the murine CFTR protein has been discovered. M1 is satisfied. Second, the experiment demonstrates that the inactivation of the murine CFTR gene by the knockout method indeed results in the abnormal CFTR and the

disease of cystic fibrosis in mice. M2 is satisfied. The condition of M3 can be satisfied only if a gene therapy for the CF disease will be invented. For most cases, M3 is too strong to be a necessary condition for the demonstration of causal relationships. Here I suggest a weaken condition M3* based on the concept of causation as difference-making to replace M3.

(M3*) The specific functional difference between the individuals with the modified gene and the individuals with the normal gene should be observed.

I call the combination of M1, M2, and M3* the difference-making version of Koch's postulates. It offers a strong criterion of causation: if one finds that a specific gene and a specific features satisfy the three conditions, then one warrant belief in there is a causal relationship between the gene and the feature. The criterion is embedded in the standardized procedure of the gene knockout method on the basis of the gene targeting technique.

In the final part of this section, let me provide a preliminary analysis of the difference-making version from a philosophical perspective of causation. M1 implies *the condition of correlation*, because it states that there is a statistical correlation between loss of a function and a mutated gene. M2 implies *the condition of a counterfactual intervention*, because it states that an intervention is made to realize a counterfactual situation. M3* implies *the condition of difference-making-by-intervention*, because it states the difference between the intervened objects and the not-intervened objects is resulted by the intervention. Whether the intervention version of Koch's postulates for the case of the gene knockout can be generalized into a general criterion of causation or not? I leave the question and the deeper philosophical implication for another occasion.

6. Conclusion

The gene knockout method based on the technique of gene targeting is a strong exploratory method of causation. It used to explore unknown causal relationships between genes and phenotypes/features/functions in molecular biology and that between genes and diseases in biomedical science. It can demonstrate that a supposed causal relationship between a gene and a phenotype really exist on the basis of the intervention version of Koch's postulates. The nature of the gene knockout method is exploratory, because it is used to discover unknown causal relationships rather than to explain known phenomena. However, it is not only an exploratory experiment but also includes other non-experimental means. In particular, the gene-knockout mouse experiments involve the realization of background assumptions, the production of experimental animals as tools, and an analogical inference from animal models to human diseases. All of the actions aim at determining causal relationships rather than offering scientific explanations or causal explanations. In a summary, a scientific exploration of causation has been preliminary explored in this paper.

Reference

Burian, Richard M. (1997). "Exploratory Experimentation and the Role of Histochemical Techniques in the Work of Jean Brachet, 1938-1952." *History and Philosophy of the Life Science* 19 (1): 27-45.
Burian, Richard M. (2007). "On MicroRNA and the Need for Exploratory Experimentation in Post-Genomic Molecular Biology." *History and Philosophy of the Life Sciences*, 29(3): 285-311.
Byrd, Allyson L. and Julia A. Segre (2016). "Adapting Koch's postulates: Criteria for disease causation must take microbial interactions into account." Science, 351(6270): 224-226. https://doi.org/10.1126/science.aad6753
Capecchi, Mario R. (1994). Targeted Gene Replacement. *Scientific American*, 270(3): 52-59.
Capecchi, Mario R. (2005). "Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century." *Genetics*, 6: 507-512.
Clarke, L. L., B. R. Grubb, S. E. Gabriel, O. Smithies, B. H. Koller and R. C. Boucher (1992). "Defective Epithelial Chloride Transport in a Gene-Targeted Mouse Model of Cystic Fibrosis." *Science*, 257(5073): 1125-1128.
Colaço, David (2018). "Rethinking the role of theory in exploratory experimentation." *Biology & Philosophy*, 33: 38. https://doi.org/10.1007/s10539-018-9648-9
Craver, Karl and Lindley Darden (2013). *In Search of Mechanisms*. The University of Chicago.
Davis, Pamela (2006). "Cystic fibrosis since 1938." *American Journal of Respiratory and Critical Care Medicine*, 173: 475-481
Donehower L. A., M.Harvey, B. L. Slagle, M. J. McArthur, C. A. Jr. Montgomery, , J. S. Butel, and A. Bradley (1992). "Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours." *Nature*, 356: 215-221.
Doyle, Alfred, Michael P. McGarry, Nancy A. Lee, and James J. Lee (2012), "The Construction of Transgenic and Gene Knockout/Knockin Mouse Models of Human Disease." *Transgenic Research*, 21(2): 327–349. http://doi.org/10.1007/s11248-011-9537-3
Ericsson, Aaron (2013). "A Brief History of Animals." *Missouri Medicine*, 110(3): 201-205.
Evans, Alfred S. (1976). "Causation and Disease: The Henle-Koch Postulates Revisited." *The Yale Journal of Biology and Medicine*, 49: 175-195.
Fagan, Melinda B. (2013). *Philosophy of Stem Cell Biology: Knowledge in Flesh and Blood*. Palgrave Macmillan.
Fredrick, David N. and David A. Relman (1996). "Sequence-Based Identification of Microbial Pathogens: a Reconsideration of Koch's Postulates," *Clinical Microbiology Reviews*, 9(1): 18–33.
Franklin, L. R. (2005). "Exploratory Experiments." *Philosophy of Science*, 72(5): 888-899.
Falkow, Stanley (1988). Molecular Koch's Postulates Applied to Microbial Pathogenicity. *Reviews of Infectious Diseases*, Vol. 10, Supplement 2: S274-S276.
Gelfert, Axel (2018). "Models in Search of Targets: Exploratory Modelling and the Case of Turing Patterns." *European Studies in Philosophy of Science*, 9: 245-269. https://doi.org/10.1007/978-3-319-72577-2

Gurumurthy, C. B. and K. C. Lloyd (2019). "Generating mouse models for biomedical research: technological advances." *Disease Models & Mechanisms*, 12: 1-10. https://doi.org/10.1242/dmm.029462

Hall, Bradford, Advait Limaye, and Ashok B Kulkarni (2010). "Overview: Generation of Gene Knockout Mice." *Current protocols in cell biology*, Chapter 19: Unit 19.1217. doi:10.1002/0471143030.cb1912s44.

Mättig, Peter (2022). "Classifying exploratory experimentation – three case studies of exploratory experimentation at the LHC." *European Journal for Philosophy of Science*, 12: 66. https://doi.org/10.1007/s13194-022-00496-4

Morgan, T. H., A. H. Sturtevant, H. J. Muller, and C. B. Bridges (1915). *The Mechanism of Mendelian Heredity*. Henry Holt and Company.

O'Malley, Maureen A. (2007). "Exploratory Experimentation and Scientific Practice: Metagenomics and the Proteorhodopsin Case." *History and Philosophy of the Life Sciences*, 29(3): 337-360.

Rader, Karen (2004). *Making Mice: Standardizing Animals for American Biomedical Research, 1900-1955*. Princeton University Press.

Rommens, J. M., M. C. Iannuzzi, B. Kerem, M. L. Drumm, G. Melmer, M. Dean, R. Rozmahel, J. L. Cole, D. Kennedy, N. Hidaka, M. Zsiga, M. Buchwald, J. R. Riordan, L.-C. Tsui and F. S. Collins (1989). "Identification of the Cystic Fibrosis Gene: Chromosome Walking and Jumping." *Science*, 245(4922): 1059-1065.

Smithies Oliver and Nobuyo Maeda (1995). "Gene targeting approaches to complex genetic diseases: Atherosclerosis and essential hypertension." *Proceedings of the National Academy of Sciences*, 92: 5266-5272.

Snouwaert, John N., Kristen K. Brigman, Anne M. Latour, Nadia N. Malouf, Richard C. Boucher, Oliver Smithies and Beverly H. Koller (1992). "An Animal Model for Cystic Fibrosis Made by Gene Targeting." *Science*, 257(5073): 1083-1088.

Steinle, Friedrich (1997). "Entering New Fields: Exploratory Uses of Experimentation." *Philosophy of Science* (Proceedings) 64: S65-S74.

Steinle, Friedrich (2002). "Experiments in History and Philosophy of Science." *Perspectives on Science*, 10(4): 408-432.

Tara, F., P. Stainer, C. Wicking, S. Halford, H. Kruyer, N. J. Lench, P. J. Scambler, C. Hansen, J. C. Braman, R. Willamson, And B. J. Wainwright (1991). "Cloning the Mouse Homolog of the Human Cystic Fibrosis Transmembrane Conductance Regulator Gene." *Genomes*, 10: 301-307.

Waters, C. Kenneth (2007). "The Nature and Context of Exploratory Experimentation: An Introduction to Three Case Studies of Exploratory Research." *History and Philosophy of the Life Sciences*, 29(3): 275-284.

Walker, Lary, Harry LeVine and Mathias Jucker (2006). "Koch's Postulates and Infectious Proteins." *Acta Neuropathol*, 112(1): 1–4. https://doi.org/10.1007/s00401-006-0072-x