# Funding of societally relevant breakthroughs in biology

**Jelmer de Ronde**
Theoretical Biology
Department of Biology
Utrecht University

June, 2014

**Supervisors**
Rob de Boer
Albert Meijer

**Abstract**

Most biological research is supported by public institutions such as research councils, which are funded by taxpayers. This begs the question whether such publicly funded research should contribute to society and how such contributions can be stimulated. To provide insight into this issue, I studied how societally beneficial scientific breakthroughs have occurred. I traced the origins of several scientific breakthroughs and categorized the steps of their scientific backstory. I conclude with an analysis of the different developmental routes I identified and make recommendations for future research funding policies.

# Contents

**1 Introduction** **3**

1.1 Contents of this thesis . . . . . 3

**2 Methods** **5**

2.1 Science Breakthrough of the Year . . . . . 5

2.2 Tracing the backstory of discovery . . . . . 5

2.3 Categorizing each step of the backstory . . . . . 6

2.4 Further analysis . . . . . 7

**3 Results** **8**

3.1 1996: New hope in HIV . . . . . 8

3.2 1997: Dolly the sheep . . . . . 10

3.3 1999: Stem cells . . . . . 11

3.4 2000: Whole genome sequencing . . . . . 13

3.5 2002: RNA interference . . . . . 14

3.6 2007: Human genetic variation . . . . . 15

3.7 2008: Reprogramming cells . . . . . 17

3.8 2011: HIV treatment as prevention . . . . . 18

3.9 2013: Cancer immunotherapy . . . . . 19

3.10 Summary and analysis of results . . . . . 20

**4 Conclusion & discussion** **22**

4.1 Insights & policy recommendations . . . . . 22

4.2 Discussion & future improvement and additions . . . . . 23

4.3 Outlook . . . . . 23

# 1. Introduction

In recent years, the best way to fund research has been an often debated subject, both in The Netherlands and abroad[1;2]. The debate distinguishes between two modes of funding: directed or undirected. Much biological research is financed by public institutions such as research councils, which are funded by taxpayers. Because of this it has been argued that the public should direct research to be beneficial to society by changing the allocation of funds[3]. This then, would be directed funding. The main argument against such a scheme is the fact that it is impossible to predict the effects of research in advance. Fundamental research undertaken without any direct societal benefit in mind, has often led to advances that proved to be beneficial in the long run[4]. Those who hold this view are more inclined towards undirected funding: granting funds to scientists without setting the research agenda. Instead of focusing on either of these two funding schemes I sought an answer to the following research questions:

*How have recent research breakthroughs that were beneficial to society been reached?*

In my research, I have limited myself to the biological sciences because this is my field of expertise. In this thesis, I also concern myself with the implications my answer has on the best way to fund research beneficial to society.

## 1.1 Contents of this thesis

To answer the research question I posited above, I conducted a case study of recent biological research breakthroughs. Such a procedure to study research has been used before, see for example Hollingsworth (2002)[5]. The results of the case study have provided insight into how societally beneficial research breakthroughs were reached before. They also allowed me to make recommendations on the future funding of research. Doing such a study enabled me to gather relatively much detail on a limited set of research breakthroughs, allowing a nuanced view of the subject. A potential disadvantage of a case study is the introduction of a bias in the cases selected. Steps to avoid such a bias are described in the next paragraph. Using this methodology I hope to provide a starting point for a nuanced discussion on societally beneficial research and funding thereof.

As my source for societally relevant studies, I used the yearly awarded Science Breakthrough of the Year. Science's Breakthrough of the Year is widely recognized as selecting research which has

a positive and significant impact on society. Every year since 1996, Science selects a pivotal scientific breakthrough[6]. Of the 18 breakthroughs awarded since then, 10 have been in the biological sciences, providing me with an adequate sample to study.

Since few breakthroughs are reached in a vacuum, I then traced back the work on which each breakthrough was built. However, the development history of modern research is complex and often draws on diverse previous achievements and publications. I therefore resorted to reducing the developmental history of each breakthrough to a simple linear path. After defining the development route, I assigned the different steps leading to each breakthrough into three different categories of research: fundamental, applied and technological. Finally, I analyze these categorizations and draw conclusions on how research breakthroughs are reached.

A more precise description of the methodology I used for this case study is described in the next chapter. The actual breakthroughs, their history, their benefit to society, and how I categorized the steps in their developmental route is presented in the results chapter. Finally, the last chapter will list my recommendations and discuss avenues for future research. There, I will also outline the limitations of the research described in this thesis.

# 2. Methods

## 2.1 Science Breakthrough of the Year

As explained in the introduction, my first step to studying the societally relevant biological breakthroughs was to list all of Science's Breakthroughs of the Year since the first was awarded in 1996. Since then, Science has awarded the most significant development in research each year with this title. The Science Breakthroughs of the Year are widely recognized as one of the highest distinctions in science. It is awarded based on the impact the breakthrough had or will have on both research itself and society at large. A table listing all the breakthroughs is shown below (Table 2.1). The Table also shows whether I determined a breakthrough to be biological in nature, as determined by standard biology textbooks such as Biology by Reece[7]. If so, a reference to the original Science article describing the breakthrough is added.

The Science article about "Evolution in action" (the 2005 breakthrough) lists a number of research discoveries related to evolutionary processes. Since it mentions Darwin's 1859 publication of the Origin of Species as the real breakthrough, it falls outside of the scope of this thesis and will not be discussed any further.

## 2.2 Tracing the backstory of discovery

The original articles from Science formed the starting point of my analysis. Each article describes what research breakthrough was selected that year and what the reasoning behind this decision was. Using this as a lens, I identified the original scientific publication(s) that constituted the Breakthrough of the Year. The next step in my analysis was determining what earlier scientific work was needed to enable the breakthroughs to occur. I did this by reading the scientific publications detailing how the research was done, and using the references of each article to trace back earlier work. In addition I used material from the Internet and the Science articles to create a more complete backstory for each case. I repeated these steps as often as necessary for each Breakthrough to trace the backstory as far as necessary to get a complete picture.

As explained in the introduction, I then reduced each developmental history to a simple linear case by selecting the most important steps. I based this selection on the original articles describing why Science selected these publications as their Breakthrough of the Year. While being a simplification, the linear histories are easier to grasp and analyze. All information pertaining to each specific case and my reasoning is described in the results chapter. Having defined linear backstories for each

***Table 2.1:*** *Table of all of Science's Breakthroughs of the Year. Each row lists the year the breakthrough was published in Science, a short description of the breakthrough, whether I consider the breakthrough to be biological and a reference to the original article in Science for the selected breakthroughs. The original list was retrieved from Wikipedia*[6].

| Year | Breakthrough | Biological? | Reference |
|---|---|---|---|
| 1996 | Understanding HIV | yes | Balter (1996) |
| 1997 | Dolly the sheep, the first mammal to be cloned from adult cells | yes | Pennisi (1997) |
| 1998 | Accelerating universe, dark matter | no | |
| 1999 | Capturing the promise of youth with stem cells | yes | Vogel (1999) |
| 2000 | Full genome sequencing | yes | Pennisi (2000) |
| 2001 | Nanocircuits or molecular circuit | no | |
| 2002 | RNA interference | yes | Couzin (2002) |
| 2003 | Dark energy | no | |
| 2004 | Spirit rover landed on Mars | no | |
| 2005 | Evolution in action | yes | Culotta and Pennisi (2005) |
| 2006 | Proof of the Poincaré conjecture | no | |
| 2007 | Human genetic variation | yes | Pennisi (2007) |
| 2008 | Cellular reprogramming | yes | Vogel (2008) |
| 2009 | Ardipithecus ramidus | no | |
| 2010 | The first quantum machine | no | |
| 2011 | HPTN 052 clinical trial | yes | Cohen (2011) |
| 2012 | Discovery of the Higgs boson | no | |
| 2013 | Cancer immunotherapy | yes | Couzin-Frankel (2013) |

Breakthrough of the Year in this way, I then proceeded to categorize each step of each backstory as described in the following section.

## 2.3 Categorizing each step of the backstory

In order to compare the different backstories I assigned each step in the backstory to one of three categories: fundamental, applied or technological research. I defined fundamental research as a publication which was necessary to enable the breakthrough, but which did not lead to any immediate practical applications. Applied research was defined as a publication which was important in the selection of the topic as Breakthrough of the Year and which led to immediate practical applications beneficial to society. The last category, technological research, I assigned to publications wherein technologies or methodologies were described that were crucial to the breakthrough. See also Figure 2.1.

As several publications can be said to belong to more than one of these categories, I assigned a publication based on the role it played *with respect to* the breakthrough as described by Science. The rationale for assigning a certain publication to one of these three categories is explained in each case in the results chapter.

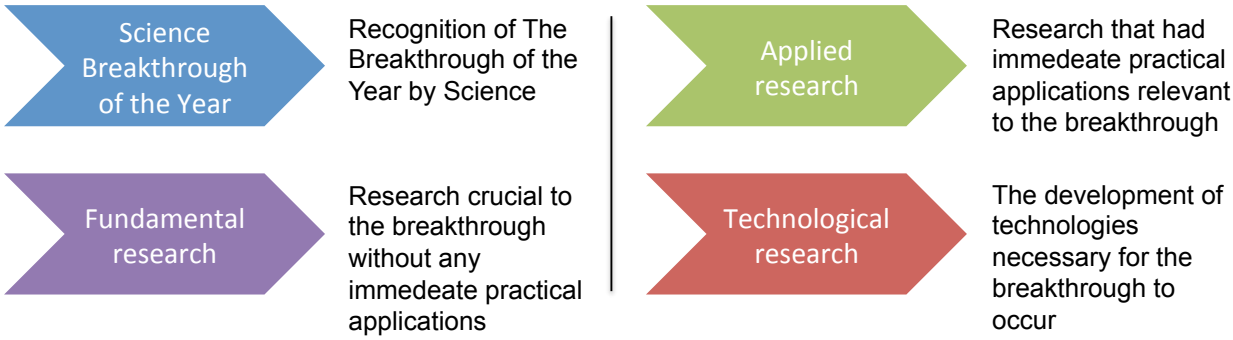


***Figure 2.1:*** *An overview of the categories applied to the developmental steps in the backstory leading to a research breakthrough. The recognition of the breakthrough by Science is rendered in blue. Fundamental research without direct applications but crucial to the breakthrough is rendered in purple. Technological research is applied to research that developed necessary technologies is rendered in red. And Applied research is rendered in green. These arrows are used in the results chapter to visualize the backstory and my categorization of the steps thereof.*

## 2.4 Further analysis

Processing each Breakthrough in this way resulted in 10 discrete developmental backstories, each one consisting of differing combinations of steps categorized as fundamental, applied or technological research. As the final step in my analysis I counted the different steps belonging to each backstory.

# 3. Results

In the following sections, I will describe all of Science's biological breakthroughs of the year I analyzed. Each section will begin with a short explanation of the breakthrough itself. I then describe what the backstory to each breakthrough is and explain my rationale for assigning different categories to each of the publications. Each section ends with the identification of the developmental route the breakthrough belongs to.

## 3.1 1996: New hope in HIV

The first Breakthrough of the Year ever to be selected by Science were newly developed potent drugs against Human Immunodeficiency Virus (HIV) and new insights into how it enters cells in 1996[8]. HIV is a retrovirus that causes acquired immunodeficiency syndrome (AIDS). Retroviruses use RNA as their hereditary material. Upon infecting a cell, their RNA first has to be transcribed into DNA so it can hijack the cells protein production capabilities to create new virus particles. To transcribe its RNA into DNA, an HIV particle carries a protein called reverse transcriptase that can perform this function[18]. The new viral proteins that are then produced by the cell first have to be cleaved by protease before they are functional[19].

For many years, the only available treatments for HIV were reverse transcriptase inhibitors such as AZT. However, in 1995 the United States Food and Drug Administration (FDA) approved a new drug that could target HIV protease called Saquinavir[8]. Two similar drugs, also targeting protease, called Ritonavir and Indinavir were approved quickly thereafter in March 1996. At that point Triple therapy, consisting of two transcriptase inhibitors such as AZT and one protease inhibitor, quickly became the norm in treating HIV.

Analysis of the HIV-1 genome showed that it encodes an aspartic protease[20]. Aspartic proteases are proteases that use an aspartate residue in their active sites to cleave the target peptide. Since its discovery in HIV-1, the protease has been a target for drug development. Protease inhibitor drugs were developed using a process called structure-assisted drug design and discovery process[21]. This process utilizes crystallography to visualize the structure of the target protein, in this case HIV protease. Using the three dimensional structure of the protein, researchers can localize and target the functional parts of a protein to develop drugs in a structured way.

Hot on the heels of protease inhibitors was a fundamental discovery clarifying how HIV enters cells. Researchers had known for a long time that the CD4 receptor is involved in the entry of HIV, but also knew that CD4 alone is not sufficient[22]. An important step in solving this riddle was taken

on 15 December 1995, when researchers identified three related chemokines (signaling proteins) that inhibit HIV replication[23]: RANTES, MIP-1$\alpha$, and MIP-1$\beta$. They did so by building on the work of Jay Levy, who discovered in 1986 that specific white blood cells could suppress HIV[24;25]. Three years later Levy and his colleagues demonstrated that the mechanism of suppression was tied to soluble factors that the cells excrete[26]. Since then, researchers have been trying to identify these factors until Gallo and Lusso succeeded in 1995.

The next step was the subsequent discovery that these chemokines bind to the receptors that are required for HIV to enter the cell: CCR5 and CXCR4[27;28]. Thus, in 1996 researchers finally had a good picture of how HIV enters the cell and what factors suppress this process.

Since Science identifies two related but separate developments, I've also categorized two different developmental routes. My categorizations are shown in Figures 3.1 and 3.2.

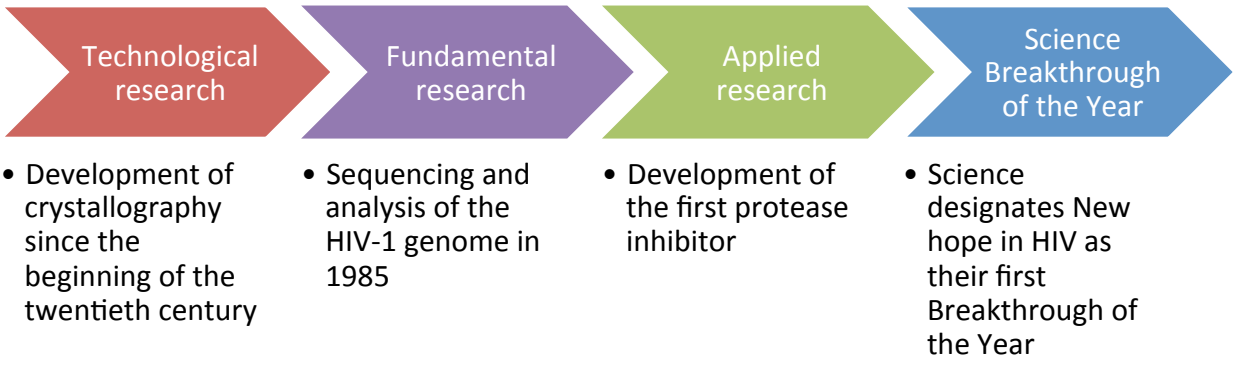


***Figure 3.1:*** *Developmental route of New hope in HIV, focusing on the development of protease inhibitors.*

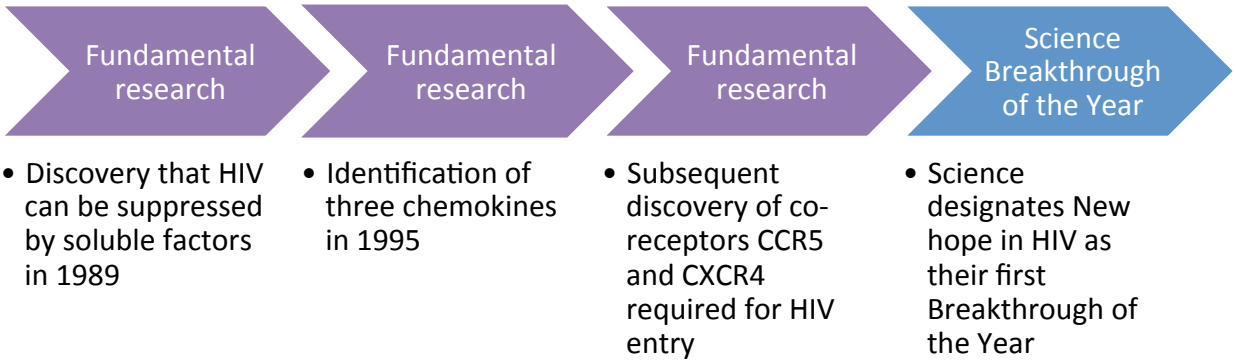


***Figure 3.2:*** *Developmental route of New hope in HIV, focusing on the discovery of the chemokines suppressing HIV and co-receptors required for HIV entry into the cell..*

First, I will explain my reasoning on the development of protease inhibitors:

1. Development of crystallography since the beginning of the twentieth century. Although the development of crystallography involved a lot of fundamental work on the nature of matter and light, I categorize it as technological research. It was the availability of working crystallography technology which enabled the structure-assisted drug design.

2. Analysis of the HIV-1 genome showed that it encodes a protease. This knowledge enabled researchers to understand HIV better and target the protease as a potential weak spot.

3. Applied research uses crystallography to create the first protease inhibitor. Other inhibitors soon follow.

4. Science chooses the FDA approval of the inhibitors as its first Breakthrough of the Year.

I also focus on the separate discovery of the chemokines RANTES, MIP-1$\alpha$, and MIP-1$\beta$. Discovery of these chemokines and their receptors was categorized in the following way:

1. Jay Levy and his colleagues discover that specific white blood cells can suppress HIV in 1986 and determine that this effect is due to soluble factor in 1989. At the time, this had no practical applications and all. Therefore, I categorize this as fundamental research.

2. The factors RANTES, MIP-1$\alpha$, and MIP-1$\beta$ are identified by Gallo and Lusso et al in 1995. At the time, this merely provided the answer to a riddle posed in 1989 by Levy et al.

3. In 1996, the subsequent discovery of the receptors that these chemokines and HIV both bind to. This clarified the mechanism of HIV entry into the cell. No practical applications were developed, so I categorize this as a fundamental discovery as well.

4. Science chooses the new insights into HIV entry into the cell as its first Breakthrough of the Year.

## 3.2 1997: Dolly the sheep

Dolly the sheep, born on July 5, 1996, was the first mammal cloned from an adult cell. In 1997 Science designated her birth as the Breakthrough of the Year[9]. Cloning is the process of using an individual's genetic material to create another organism which has the same DNA, effectively creating an identical twin. This is done using a technique called somatic cell nuclear transfer (SCNT), by which the nucleus of a cell of the individual to be cloned is transfered to a cell that has been stripped of its nucleus. When used for reproductive cloning, the nucleus is implanted in an egg cell, which is then allowed to proliferate into an embryo which can be implanted into the uterus of a carrier mother. The carrier mother then has a normal pregnancy, after which a new individual, genetically identical to the original individual is born. It was using this procedure that Dolly the sheep was created[29].

Although Dolly was the first mammal to be cloned from an adult cell, reproductive cloning was, even at the time, not a new technique. Researchers have already been working applying SCNT on frog cells to clone tadpoles in 1952[9]. Even mammals have been cloned before: in 1986 Willadsen cloned a sheep embryo using an embryonic nucleus and donor[30]. However, until 1996, the only successful clonings had been using a nucleus from embryo's. Thus researchers came to believe that it was impossible for adult nuclei to regress to the point were they could generate a new individual altogether. The cloning of Dolly changed that paradigm. At the Roslin Institute, Wilmut et al. induced a quiescent state in the cell that would donate the nucleus, by starving it. They did this in the hope that it would reset the nucleus to a state were it would once again be able to generate a

new individual. The theories of Wilmut et al. proved to be successful and with the birth of Dolly the sheep they altered the reigning paradigm.

Besides being widely used in research to create identical test subjects, cloning can be useful in other areas as well. Animals with desired traits, such as high milk production or lean meat could be cloned. Or to preserve species threatened with extinction, researchers are considering cloning endangered animals[31].

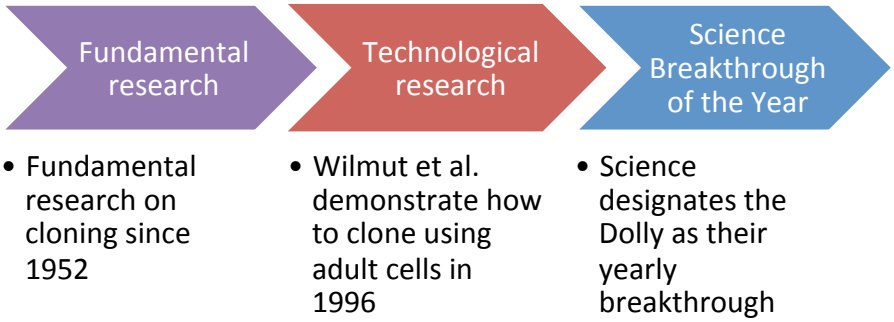


***Figure 3.3:*** *Developmental route of Dolly the sheep.*

I show my timeline of the development of Dolly the sheep in Figure 3.3. Based on the original Science article describing the breakthrough, I have made the following categorizations:

1. Fundamental research on cloning since 1952, which had no practical applications at that point in time.
2. Pivotal research by Wilmut et al. which allows cloning using adult cells in 1996. Their technique of inducing a quiescent state not only changes the reigning paradigm, but also allows much broader use of reproductive cloning. Because they developed the technique used to clone Dolly I categorize this as technological research.
3. Science designates Dolly as their 1997 Breakthrough of the Year.
4. Possible applications to clone desired animals because of their usefulness in research, agriculture or preservation of the environment.

## 3.3 1999: Stem cells

Stem cells are cells that retain the ability to differentiate into several different cell types while maintaining their own undifferentiated state over the course of numerous cycles of cell division[32]. In adult organisms, stem cells replenish specific cells by proliferating and subsequently differentiating into the cell type that is required. In developing organisms, stem cells divide to provide cells that can differentiate into different tissues. Different types of stem cells exist. The best known are embryonic stem cells, which naturally only exist in embryos and which can differentiate into all cell types. But pluripotent stem cells, which can differentiate into a subset of cell types also exist, for example hematopoietic stem cells which can differentiate into different types of blood cells. In

1999 Science designated Stem cells, in an article titled "Capturing the Promise of Youth", as the Breakthrough of the Year[10].

In 1998 both Thomson et al. and Shamblott et al. succeeded at keeping human embryonic stem cells in their undifferentiated states, thereby creating the first human embryonic stem cell lines[33;34]. This paved the way for a lot of publications on the properties of these cells in 1999, prompting Science to select it as their breakthrough[35–38]. Hints of the existence of stem cells were first seen in the 1960s, when self-renewing cells were found in mouse bone marrow and a bone marrow transplant was performed between siblings to successfully treat severe combined immunodeficiency (SCID)[39;40]. In 1981 Evans et al. and Martin independently managed to isolate and maintain embryonic stem cells for the first time, albeit in mice[41;42]. After the first hints of the existence of stem cells in the 1960s, the isolation of embryonic stem cells in mice by Evans et al. and by Martin in 1981 and in humans by Thomson et al. and Shamblott et al. in 1998 classify as important breakthroughs.

Since their isolation, stem cells have been a valuable research tool. But besides being a tool, stem cells hold the incredible potential to regenerate damaged or faulty tissue and thus be of great value in restorative medicine as well. In the past decades, there has been numerous research on both these issues[43–45]. Reason enough to designate the first human embryonic stem cell lines and the subsequent research that they spurred as the 1999 Science Breakthrough of the Year.

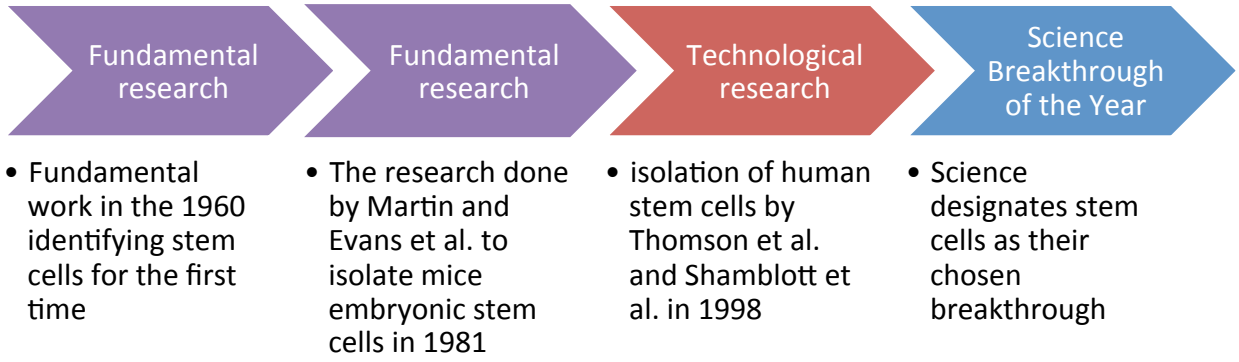


***Figure 3.4:*** *Developmental route of Stem cells.*

Figure 3.4 shows my categorization of the Stem cells breakthrough. I applied the categories in the following way:

1. Fundamental work done in the 1960s identifying stem cells for the first time.
2. The fundamental research done by Martin and Evans et al. to isolate mice embryonic stem cells. At the time, it was predominantly a demonstration that embryonic stem cells could be maintained without differentiating into different stem cells.
3. The important breakthrough of isolating and maintaining human stem cells by Thomson et al. and Shamblott et al. in 1998. Their methods to maintain human stem cells have been used in research and medicine since.
4. Science selects stem cells as theirs chosen breakthrough.

## 3.4 2000: Whole genome sequencing

Since the discovery of DNA as the bearer of genetic information in 1952[46], researchers have longed to read the the full text of the genetic textbook. In 2000, Science marked that this dream had finally come to fruition[11], as whole genome sequencing was finally coming of age. DNA consists of long chains of the same four molecules, called bases, each one represents a single letter of genetic text. Combinations of these four letters spell out the recipes for all the myriad proteins that life consists of. DNA sequencing is the process of analyzing the bases to discover in which order they occur, essentially reading the textbook. Although DNA sequencing has been possible since 1975[47;48], it was only feasible to decipher small sections of the genetic code due to the painstaking nature of these early methods.

Newer methods made it possible to sequence increasingly larger stretches of DNA, until the complete genome of the bacterium *Haemophilus influenzae* was sequenced in 1995[49], making it the first complete genome. In the following years sequencing techniques improved, allowing the sequencing of bigger genomes. Milestones worth mentioning here are the genome of the first multicellular eukaryote, *Caenorhabditis elegans* in 1998[50] and two first drafts of the human genome in 2001[51;52].

The sequencing technique that allowed all these milestones to occur is known as shotgun sequencing. Shotgun sequencing works by breaking up a longer DNA strand into several shorter strands. These shorter strands are then read after which the entire sequence is reconstructed using computer programs. These programs exploit the fact that DNA consists of two complementary strands to recombine the reads into the right order.

Although shotgun sequencing was first described since in 1979 by R. Staden[53], it could only be used to sequence smaller sequences (4000-7000 basepairs). It was only after several improvements to this technology that it became possible to sequence large genomes. The first one of these was pairwise end sequencing, also known as double-barrel shotgun sequencing, which sequences DNA by targeting both ends of the strand. The first published use of pairwise end sequencing was by Edwards et al. in 1990[54]. A final innovation took place in 1995, when Roach et al. showed it was possible to use strands of varying lengths, thereby greatly increasing the genome size that could be sequenced[55]. This final innovation spurred the whole genome sequencing that took place in the following years, starting with the sequencing of the complete genome of *Haemophilus influenzae* in 1995 and the sequencing of the human genome in 2001.

Whole genome sequencing has allowed us to read the full textbook of life for the first time. Since the availability of whole genomes of several species, we are beginning to see how much our genetic code shares with other species. Showing us how we are different, and perhaps more important, how much we have in common. Whole genome sequencing has opened up large new areas of biological research which will continue to generate new insights for years to come. Since this breakthrough in 2000, scientists, doctors and biotechnology companies have also been working on applying these techniques to the promise of personalized medicine.

Figure 3.5 shows my categorization of the development of Breakthrough of the Year celebrating whole genome sequencing in the following way:

1. The discovery of DNA as the bearer of hereditary information in 1952. This fundamental discovery showed us for the first time how our hereditary information is stored.

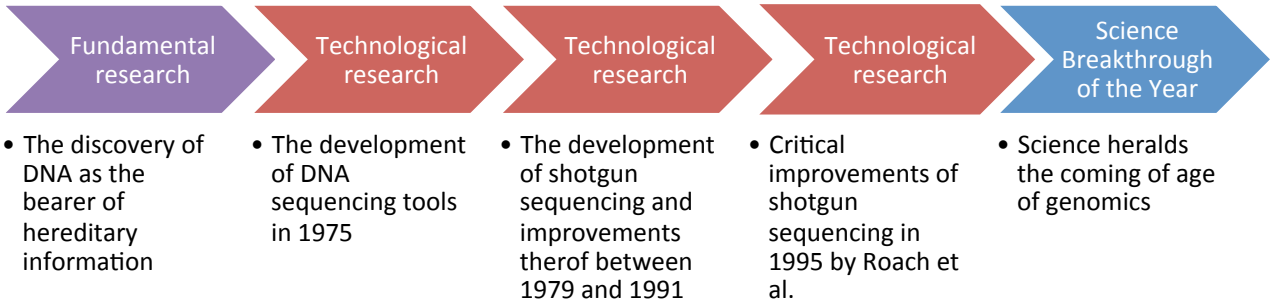


***Figure 3.5:*** *Developmental route of Whole genome sequencing.*

2. The development of the first DNA sequencing technologies in 1975 and the development of shotgun sequencing by R. Staden in 1979. It was the development of these technologies that enabled any sequencing project since.

3. Subsequent improvements on shotgun sequencing, with the turning point of Roach et al. in 1995 being the one that directly led to whole genome sequencing.

4. Science selecting whole genome sequencing as the Breakthrough of the Year.

## 3.5 2002: RNA interference

RNA interference (RNAi) is a cellular process by which RNA molecules alter the expression of genes in Eukaryotes. Prior to the discovery RNAi, it was thought that only the interactions between DNA and specialized proteins such as transcription factors control gene expression, via the well-known processes of transcription and translation. In these processes RNA molecules function only as the carriers of information, having no influence on the actual expression levels of genes. In 1990 it became apparent that RNA might have more functions than only functioning as a messenger between DNA and the ribosome for the first time. Researchers discovered that some small RNA's could influence the transcription of genes in plants[56;57]. In 1998 Andrew Fire and his colleague Craig C. Mello discovered new RNA functionality which allows the molecule to alter gene expression[58]. Their discovery of RNAi opened up a new avenues of research and changed our fundamental understanding of the inner workings of the cell. This prompted Science to designate the discovery of RNAi as the 2002 Breakthrough of the Year[12].

RNA interference is usually initiated by the creation of a long double stranded RNA (dsRNA) sequence targeting the gene to be silenced. The dsRNA sequence is then cleaved into shorter dsRNA fragments by the enzyme Dicer. The dsRNA fragments are then unwound into single stranded RNAs (ssRNA) and incorporated into RNA-induced silencing complex (RISC). The ssRNA in the RISC will then bind to the messenger RNA (mRNA), the molecule that transports the sequence information of the gene to be expressed to the machinery that will create the actual gene product. When the ssRNA binds to the mRNA, a protein called Argonaute cleaves the mRNA, preventing its eventual translation into the gene product[59]. Other functions of interfering RNAs are known as well; it was shown in 2006 that RNAi can also activate genes for example[60].

Nowadays, besides providing useful insight into the functioning of the cell and gene expression, RNAi is being used as a valuable research tool to silence genes[61]. And in years to come it is likely that we will also see applications of RNAi used in treatments of disease[62].

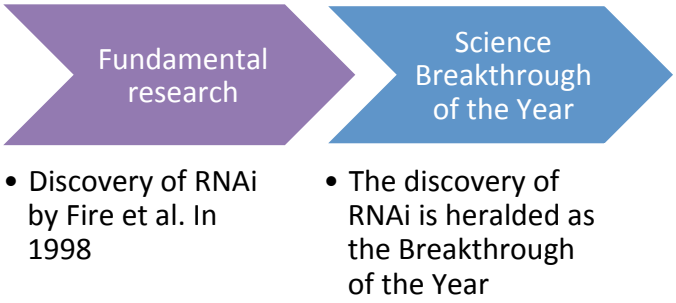


***Figure 3.6:*** *Developmental route of RNA interference.*

In Figure 3.6 I have shown the developmental route of the RNA interference breakthrough. I have applied the categories as follows:

1. The pivotal work by Fire et al. in 1998 showing the existence of the RNAi process and thereby providing us with fundamental insight into the cell's control systems. At the time it predominantly provided insight into the workings of the cell. RNAi would later be applied as a research tool.
2. In 2002 Science heralded RNAi as their yearly breakthrough.

## 3.6 2007: Human genetic variation

Although the first full human genomes were sequenced around 2001[11], the small number of individual sequences were insufficient to provide insight into the genetic differences between individuals and their associated traits. In 2007, new techniques finally allowed researchers to start making leaps in discovering how our genetic makeup differs from person to person. These insights are helping clarify the evolutionary history of our species and give researchers the ability to link genetic differences between people to specific traits such as eye color[63] or disease[64]. The incredible insights promised by these discoveries caused Science to nominate Human genetic variation the 2007 Breakthrough of the Year[14].

These new results become available through Genome Wide Association (GWA) studies[65]. GWA studies typically compare a large number of single-nucleotide polymorphisms (SNPs), mutations consisting of one base-pair mutation, between two groups of people. One group is known to have a certain trait worth studying, such as a specific disease, while the other group functions as a control. Using statistical analyses it is possible to associate specific mutations with specific traits. After associating a gene with a trait, researchers can further determine why this mutation has a certain effect, increasing our understanding of countless diseases and other human traits.

There are two important factors that made the GWA studies possible[14]. The first is the availability of large catalogs such as the International HapMap Project, which provide a map of the SNPs that vary

most between individuals[66]. Such a catalog provides the ability to focus on relatively small parts of the human genome, instead of having to sequence the entire genome for each subject, greatly reducing the time and money needed for GWA studies. The second factor is the reduced cost and increased capacity of the gene chips that allow for quick scanning of more than 100,000 SNPs[65].

The first GWA study to test a large number of SNPs identified a gene linked to Age-Related Macular Degeneration (AMD)[14;67]. In their 2005 article, Klein et al. describe in the supplementary materials that they used SNP information from the HapMap project and gene chips and a corresponding protocol developed by Affymetrix, a biotechnology company[68]. The International HapMap project officially started with a meeting on October 27-29 in 2002[69]. The gene chip technology developed by Affymetrix builds on the earlier developed micro array technology, which was first described by Tse-Wen Chang in 1983[70]. Thus, the insights into Human genetic variation were largely made possible by technological advancements as described in this paragraph.

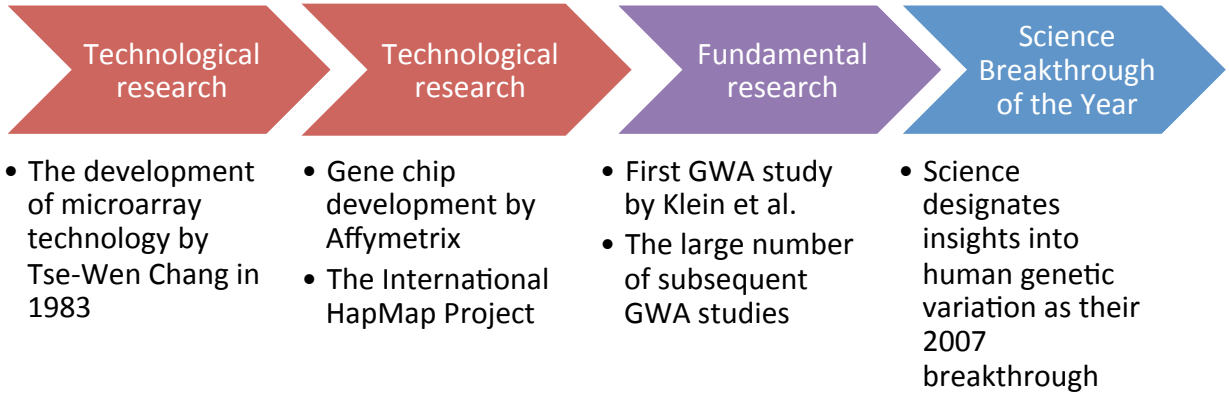


***Figure 3.7:*** *Developmental route of the Human genetic variation.*

In Figure 3.7 I show the developmental steps and my categorizations. My rationale to categorize the steps in this way is as follows:

1. The development of micro array technology by Tse-Wen Chang in 1983. Micro array technology is a technology broadly used to map genomes. GWA studies could not have been done without this enabling technology.

2. Further development of gene chips by Affymetrix in 2004 and the cataloging of SNPs by the International HapMap project that started in 2002. It was the combination of these two developments which enabled the first GWA studies to be done.

3. the first application thereof in a GWA study by Klein et al. in 2005 and the large number of GWA studies following these developments. The first GWA study was a fundamental study, providing insight into the underlying genomic differences that contribute to AMD.

4. Science designates insights into human genetic variation as their 2007 breakthrough.

## 3.7 2008: Reprogramming cells

In 2008, two research breakthroughs caused Science to recognize "reprogramming cells" as the Breakthrough of the Year[15]. Conventional wisdom holds that there is no going back once a cell is differentiated. The developmental path of cells can be visualized as a tree with lots of branches and one thick trunk on which possible paths only move upward: a cell cannot move downward to become less differentiated than it is, or jump to another branch to become a different cell type. However, around 2008 research groups did just those things.

In November 2007, a Japanese and a US group announced that they had succeeded in dedifferentiating human skin cells into induced pluripotent cells (iPS), a type of stem cell, effectively pushing cells down the tree towards the trunk[71;72]. And in October 2008 Zhou et al. published in Nature that they changed adult pancreatic exocrine cells into insulin-producing beta cells in mice[73]. This was the first time adult cells jumped a branch and turned into a different cell type.

Stem cells hold incredible potential, both for research purposes and for treating disease. Because most human cells are difficult to keep alive in the lab, they are hard to study. Therefore, using stem cells to generate new differentiated cells to study is an important research tool. In medicine, the hope has always been that stem cells will one day be used to treat disease by restoring damaged or faulty tissue. However, before the discoveries mentioned in this section, stem cells could only be obtained by harvesting embryonic material, a practice that has always raised ethical debate. The breakthroughs mentioned here have already given research a spectacular new tool for research and might in the future help cure disease while at the same time obviate the usage of embryonic material.

Researchers have known about stem cells since the early 20th century, and the first human stem cell line was derived 1998[32;33]. However, all the breakthroughs responsible for Science's Breakthrough of the Year build on pivotal research by Takahashi et al. in 2006[74]. In that year, the Japanese researchers demonstrated for the first time that it is possible to induce pluripotent stem cells by introducing four factors: Oct3/4, Sox2, c-Myc, and Klf4. In their study, they pluripotent stem cells from both mouse embryonic and adult fibroblasts. All subsequent research has used those same factors or a variation thereof to induce stem cells in different organisms or from different cell types.

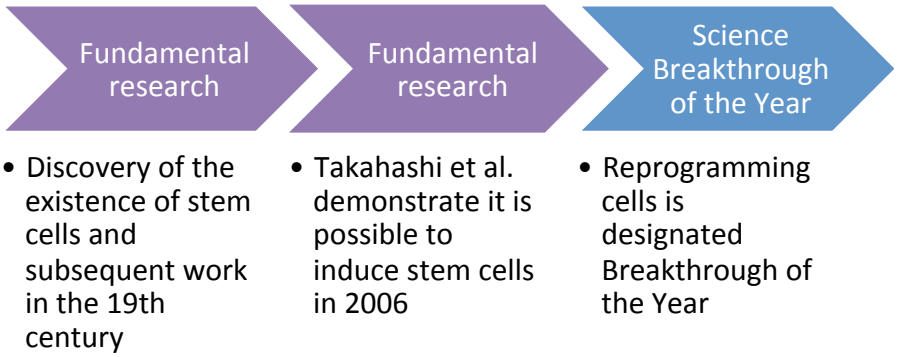


***Figure 3.8:*** *Developmental route of Reprogramming cells.*

Figure 3.8 shows my categorization of the development of the reprogramming cells breakthrough in the following way:

1. Previous fundamental discoveries on the existence of stem cells in the 20th century.

2. The pivotal work by Takahashi et al. in 2006 demonstrating for the first time that it is possible to generate stem cells from adult cells. With this discovery they change the ruling paradigm and our understanding of the basic rules of differentiating cells. Although this discovery would later lead to applications, the publication by Takahashi et al. is fundamental in nature as it only shows that it is possible to reprogram cells into stem cells.

3. Science designates reprogramming cells as their 2008 breakthrough.

## 3.8 2011: HIV treatment as prevention

The human immunodeficiency virus, or HIV, causes acquired immunodeficiency syndrome (AIDS). (For more information on HIV, see section 3.1 on page 8) HIV is a retrovirus which can be treated (but not cured) using antiretroviral drugs (ARVs). ARVs usually consists of multiple drugs to hinder several functions of the virus at the same time, making it more difficult for the virus to evolve resistance. In 2011 Science chose the discovery that ARVs not only treat existing infections, but also function as prevention of HIV transmission as the Breakthrough of the Year[16].

The HIV Prevention Trials Network (HPTN) is an international organization that develops trials to test interventions designed to prevent the transmission of HIV. In 2005, the network started a trial called HPTN 052 to test the hypothesis that treatment with ARVs would also prevent virus transmission. The study enrolled 1763 "discordant" couples; couples in which one person is already infected with HIV and the other person is not. The couples were randomly assigned to one of two groups: in one group the HIV-positive person would immediately receive ARV treatment and in the other group the HIV-positive person would only receive treatment when CD4-count (a standard measure of progression of an HIV infection) drops to 200 or lower. Since HIV is sexually transmitted, the number of partners with the same HIV strain as their initially infected partner in both groups gives an indication of the effectiveness of ARVs as prevention.

The trial was planned to continue until 2015. However, in April 2011 the independent monitoring board recommended to offer ARVs to all infected participants immediately due to the success of the ARV treatment in preventing transmission. When using ARVs the risk of transmission is reduced by 96%[75]. This discovery marked a turning point in the handling of the AIDS epidemic, granting it the Science breakthrough of the year. The question is no longer whether ARV treatment works as prevention, but how to apply the knowledge that it does as soon as possible.

Figure 3.9 shows my categorization of the development of the HPTN 052 breakthrough in the following way:

1. The successful trial by Cohen et al. demonstrating for the first time that ARVs prevent HIV transmission. Even during the process of their discovery, their insights were applied to the second group of test subjects.

2. In 2011 Science designates the trial as their yearly breakthrough.

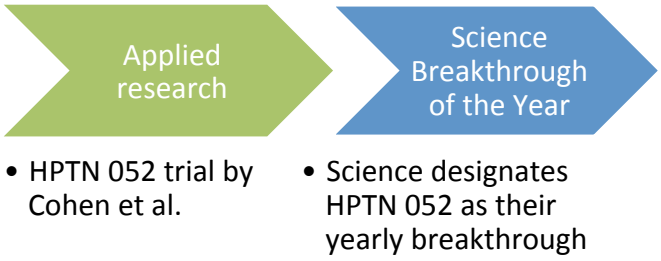


***Figure 3.9:*** *Developmental route of the HPTN 052 HIV Prevention Trial.*

## 3.9 2013: Cancer immunotherapy

In the year 2013, Science designated the discovery of cancer immunotherapy as its Breakthrough of the Year[17]. Cancer immunotherapy is a series of medical techniques designed to activate the immune system to seek out and destroy cancer cells. Although the eventual fate of cancer immunotherapy remains unclear, its discovery and recent successes have already succeeded in changing the outlook on the future of cancer treatment, warranting its designation as the 2013 Breakthrough of the Year. Development of the techniques and research into cancer therapy currently continues at a rapid pace.

Cancer immunotherapy works by introducing antibodies that target immuno-supressant proteins into the patients bloodstream, thereby evoking a powerful T-cell response to tumors. The first developed therapies work by inhibiting the cytotoxic T-lymphocyte antigen 4 (CTLA-4) and programmed death 1 (PD-1) proteins, which function as inhibitory regulators of T cells, in effect freeing the T cells to target tumors. So far several successful drug trials have proven that cancer immunotherapy is beneficial to society. In addition to that, the recent developments in immunotherapy have spurred new avenues of research and treatment that might be a boon for cancer patients in years to come.

In 1987, French researchers who were interested in the workings of the human immune response identified a new protein receptor on the surface of T cells, CTLA-4, and managed to clone the encoding gene[76]. Similarly PD-1 was discovered in 1992 by Ishida et al.[77]. For both proteins, the next step was determining their function as elements regulating the immune response. For CTLA-4 this came in 1995, when two groups showed that mice without CTLA-4 develop overactive immune systems, revealing CLTA-4's nature as an immunesupressor[78;79]. PD-1's role became clear in 1999, when Nishimura et al. discovered its role as a negative regulator of immune response[80].

The leap towards therapies to treat cancer came in 1996, when James Allison demonstrated that inhibiting CTLA-4 with an antibody could lead to tumor regression in mice[81]. After Allison's success he and two biotechnology companies Medarex and Bristol-Myers set to work on developing the first drugs, which eventually led to drug trials and successful treatments.

Also related to cancer immunotherapy is chimeric antigen receptor therapy (CAR therapy), a personalized treatment using genetically altered T cells. In 2010, Rosenberg and colleagues published results which involved modifying a patient's T cells to make them target tumors[82]. However, CAR therapy remains experimental for the time being, while therapies based on antibodies against CTLA-4 and PD-1 are going mainstream. Therefore I have disregarded CAR-therapy in this analysis.

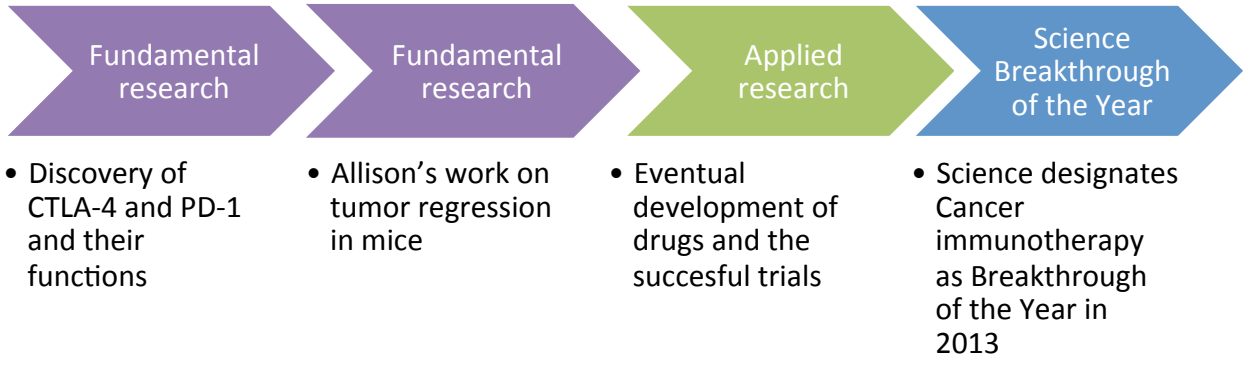


***Figure 3.10:*** *Developmental route of Cancer immunotherapy.*

Figure 3.10 shows my categorization of the development of the cancer immunotherapy breakthrough in the following way:

1. Two fundamental discoveries of the proteins CTLA-4 and PD-1 and the subsequent discoveries on their respective functions. At the time, no one was thinking about applying knowledge of these proteins to the treatment of cancer.
2. The leap of Allison to translate these fundamental insights into a demonstration of tumor regression in mice. Although crucial to the eventual development of cancer therapies, the work by Allison is fundamental in nature.
3. The subsequent work of Allison and two biotechnology companies to turn this discovery into effective drugs, which is clearly an application of earlier research.
4. Science recognizes the Breakthrough in 2013.

## 3.10 Summary and analysis of results

Table 3.1 shows all the Breakthroughs of the Year I analyzed and the number of steps within its developmental route that were assigned to each research category. This summary shows that there is much variation in the type of research leading up to a breakthrough. The variation seen here is two-fold: both within the routes and between routes no single pattern can be found. The different steps leading to a research breakthrough vary from case to case. But in almost all cases there are at least two different categories of research steps present. Between the different routes there is also no single pattern that dominates. This becomes more apparent when one takes the order of the different research steps into account, which are not shown in the Table.

So even this simplified case study tremendous variation in the research leading up to a breakthrough can be found. It is safe to assume that if the full complexity of the developmental paths had been captured, the amount of variation would have been even higher.

***Table 3.1:*** *Table listing all the research breakthroughs I analyzed. Each row represents the developmental paths of one of the breakthroughs. The last three numbers in each row represent the number steps I categorized as fundamental (FR), applied (AR) or technological (TR) research respectively.*

| **Year** | **Page** | **Breakthrough** | **FR** | **AR** | **TR** |
|---|---|---|---|---|---|
| 1996 | 8 | Understanding HIV - Protease inhibitors | 1 | 1 | 1 |
| 1996 | 8 | Understanding HIV - Chemokines & co-receptors | 3 | 0 | 0 |
| 1997 | 10 | Dolly the sheep, the first mammal to be cloned from adult cells | 1 | 0 | 1 |
| 1999 | 11 | Capturing the promise of youth with stem cells | 2 | 0 | 1 |
| 2000 | 13 | Full genome sequencing | 1 | 0 | 3 |
| 2002 | 14 | RNA interference | 1 | 0 | 0 |
| 2007 | 15 | Human genetic variation | 1 | 0 | 2 |
| 2008 | 17 | Cellular reprogramming | 2 | 0 | 0 |
| 2011 | 18 | HPTN 052 clinical trial | 0 | 1 | 0 |
| 2013 | 19 | Cancer immunotherapy | 2 | 1 | 0 |
| | | **Total** | **14** | **3** | **8** |

# 4. Conclusion & discussion

With this thesis I tried to provide an answer to the question I asked in the introduction:

*How have recent research breakthroughs that were beneficial to society been reached?*

To this end I analyzed nine cases from Science's Breakthroughs of the Year. Besides providing my answer to this question in this chapter I will also list gained insights and consider policy recommendations. Furthermore, I also discuss possible future directions that could be explored to expand the work in this thesis. Finally, I will describe how this work adds to previous research already undertaken in this area.

## 4.1 Insights & policy recommendations

**The routes leading to societally relevant breakthroughs are highly varied.** My analysis shows that the developmental routes, when categorized into three different types of research, show much variety. The different steps leading to a single breakthrough often belong to different categories. But there is also a lot variety between routes, with no single pattern dominating. This belies the commonly held belief that "real" breakthroughs can only come about by fundamental research. This point is the most important answer to my research question.

**This thesis can be seen as a critical case.** The case study I performed contained only few cases and the developmental routes where simplified to linear paths. It is safe to assume that the variety I witnessed would have been much greater when the full complexity of research breakthroughs would have been considered.

**The ecology of different routes require different funding instruments.** The different routes described in this thesis can be viewed as an ecology of research that requires diverse funding instruments to achieve all desirable results. The developmental routes can each be stimulated using different types of funding. Policy makers would be wise to consider this fact when deciding on how and what types of research to fund.

**Requirements to demonstrate societal benefits in research proposals should be lessened.** As I have shown here, fundamental research has its place in the diverse ecology leading to societal benefits. This does not mean however that these benefits are always apparent when the research

is first conceived. It would be fair to lessen the requirements for some funding instruments, so that fundamental research which might eventually contribute to a great breakthrough also has its chance to acquire funds.

## 4.2 Discussion & future improvement and additions

**My first attempts to study funding of research breakthroughs proved to be impractical.** Measuring the societal benefits of research funded in a single year by two comparable funding programs (directed and undirected) would have had the benefit of a relatively unbiased selection of scientific work to compare. However, selecting two programs and tracing and quantifying the societal benefits of the funded research proved to be both time-consuming and difficult because of poorly available data. It soon proved that the time alloted to this thesis would be insufficient to complete the study using this approach. However, future researchers with more time could still pursue such a strategy.

**More complex developmental routes.** A simple way to extend this research would be to allow for more complex developmental routes.

**It has proven difficult to know where to draw the line when tracing the developmental histories for each case.** When analyzing the backstory to each research breakthrough, there has to be a point where to stop. I have tried to determine this point as best I can, but a case for different points can certainly be made.

**Categories can be applied differently.** The categories I created to apply to the different steps of each developmental route can be applied differently. I tried to resolve this issue by applying these categories *with respect to* the original Science articles describing the breakthroughs.

**There are relatively few Science Breakthroughs of the Year, introducing a bias.** A bigger sample of research breakthroughs to analyze would certainly improve the analysis and insights this thesis offers. Another benefit of more cases would be the possibility to perform a statistical analysis on the different developmental routes, which might lead to more precise insight into scientific breakthroughs.

## 4.3 Outlook

**With this thesis, I show the variety of research required to enable a breakthrough.** This thesis shows the diversity of routes that contribute to research breakthroughs beneficial to society. Earlier work has not shown this at this point. In 2009 Heinze et al. compared research breakthroughs in two different fields: nanotechnology and human genetics[83]. For a number of breakthroughs they studied the organizational structures such as group composition and funding. In contrast to this study, their approach focused on the groups that achieved the breakthrough, and not on the developmental route leading to the breakthrough. A study by Hollingsworth in 2002 is more comparable to this one[5], but also focuses on institutional factors.

# Bibliography


[1] Ernst-Jan Hamel and Xander Bronkhorst. 'De wetenschap werkt niet zoals het moet', 2013. URL http://www.dub.uu.nl/artikel/achtergrond/wetenschap-werkt-niet-zoals-moet.html.

[2] Steven Hill. Research funding in a time of austerity, April 2013. URL http://www.theguardian.com/science/political-science/2013/apr/24/research-funding-austerity.

[3] Science in Transition. Over Science in Transition, 2013. URL http://www.scienceintransition.nl/over-science-in-transition.

[4] Michael Polanyi. Tacit knowledge. *Knowledge in organizations*, pages 135–146, 1997.

[5] JR Hollingsworth. *Research organizations and major discoveries in twentieth-century science: a case study of excellence in biomedical research*. 2002. URL http://www.miroslavradman.com/content_files/Hollingsworth.pdf.

[6] Wikipedia. Breakthrough of the Year, . URL http://en.wikipedia.org/wiki/Breakthrough_of_the_Year.

[7] JB Reece. *Campbell biology*. 2011. ISBN 10: 0321739752. URL http://199.120.181.101/syllabi/PDF/BIOL/BIOL1231.pdf.

[8] Michael Balter. AIDS Research: New Hope in HIV Disease. *Science*, 274(5295):1988, December 1996. doi: 10.1126/science.274.5295.1988. URL http://www.sciencemag.org/content/274/5295/1988.1.short.

[9] Elizabeth Pennisi. CLONING: The Lamb That Roared. *Science*, 278(5346):2038–2039, December 1997. doi: 10.1126/science.278.5346.2038. URL http://www.sciencemag.org/content/278/5346/2038.short.

[10] Gretchen Vogel. Capturing the Promise of Youth. *Science*, 286(5448):2238–2239, December 1999. doi: 10.1126/science.286.5448.2238. URL http://www.sciencemag.org/content/286/5448/2238.short.

[11] Elizabeth Pennisi. Genomics Comes of Age. *Science*, 290(5500):2220–2221, December 2000. doi: 10.1126/science.290.5500.2220. URL http://www.sciencemag.org/content/290/5500/2220.short.

[12] Jennifer Couzin. Small RNAs Make Big Splash. *Science*, 298(5602):2296–2297, December 2002. doi: 10.1126/science.298.5602.2296. URL http://www.sciencemag.org/content/298/5602/2296.short.

[13] Elizabeth Culotta and Elizabeth Pennisi. Evolution in Action. *Science*, 310(5756):1878–1879, December 2005. doi: 10.1126/science.310.5756.1878. URL http://www.sciencemag.org/content/310/5756/1878.short.

[14] Elizabeth Pennisi. Human Genetic Variation. *Science*, 318(5858):1842–1843, December 2007. doi: 10.1126/science.318.5858.1842. URL http://www.sciencemag.org/content/318/5858/1842.short.

[15] Gretchen Vogel. Reprogramming Cells. *Science*, 322(5909):1766–1767, December 2008. doi: 10.1126/science.322.5909.1766. URL http://www.sciencemag.org/content/322/5909/1766.1.short.

[16] Jon Cohen. HIV Treatment as Prevention. *Science*, 334(6063):1628, December 2011. doi: 10.1126/science.334.6063.1628. URL http://www.sciencemag.org/content/334/6063/1628.short.

[17] Jennifer Couzin-Frankel. Cancer Immunotherapy. *Science*, 342(6165):1432–1433, December 2013. doi: 10.1126/science.342.6165.1432. URL http://www.sciencemag.org/content/342/6165/1432.short.

[18] Yong Hui Zheng, Nika Lovsin, and B. Matija Peterlin. Newly identified host factors modulate HIV replication. In *Immunology Letters*, volume 97, pages 225–234, 2005. ISBN 0165-2478. doi: 10.1016/j.imlet.2004.11.026.

[19] Ashraf Brik and Chi-Huey Wong. HIV-1 protease: mechanism and drug discovery. *Organic & biomolecular chemistry*, 1:5–14, 2003. ISSN 14770520. doi: 10.1039/b208248a.

[20] L Ratner, W Haseltine, R Patarca, K J Livak, B Starcich, S F Josephs, E R Doran, J A Rafalski, E A Whitehorn, and K Baumeister. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature*, 313:277–84, 1985. ISSN 0028-0836. doi: 10.1038/313277a0. URL http://www.ncbi.nlm.nih.gov/pubmed/2578615.

[21] A Wlodawer and J Vondrasek. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annual review of biophysics and biomolecular structure*, 27:249–284, 1998. ISSN 1056-8700. doi: 10.1146/annurev.biophys.27.1.249.

[22] David C. Chan and Peter S. Kim. HIV entry and its inhibition, 1998. ISSN 00928674.

[23] F. Cocchi, A. L. DeVico, A. Garzino-Demo, S. K. Arya, R. C. Gallo, and P. Lusso. Identification of RANTES, MIP-1alpha, and MIP-1beta as the Major HIV-Suppressive Factors Produced by CD8+ T Cells, 1995. ISSN 0036-8075.

[24] C M Walker, D J Moody, D P Stites, and J A Levy. CD8+ lymphocytes can control HIV infection in vitro by suppressing virus replication. *Science (New York, N.Y.)*, 234:1563–1566, 1986. ISSN 0036-8075. doi: 10.1126/science.2431484.

[25] Michael Balter. Elusive HIV-Suppressor Factors Found. *Science*, 270(5242):1560–1561 CR – Copyright © 1995 American Ass, December 1995. ISSN 00368075. doi: 10.2307/2887886. URL http://www.jstor.org/stable/2887886.

[26] C M Walker and J A Levy. A diffusible lymphokine produced by CD8+ T lymphocytes suppresses HIV replication. *Immunology*, 66:628–630, 1989. ISSN 0019-2805.

[27] M. Dean, M. Carrington, C. Winkler, G. A. Huttley, M. W. Smith, R. Allikmets, J. J. Goedert, S. P. Buchbinder, E. Vittinghoff, E. Gomperts, S. Donfield, D. Vlahov, R. Kaslow, A. Saah, C. Rinaldo, R. Detels, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study, and S. J. O'Brien. Genetic Restriction of HIV-1 Infection and Progression to AIDS by a Deletion Allele of the CKR5 Structural Gene, 1996. ISSN 0036-8075.

[28] C K Lapham, J Ouyang, B Chandrasekhar, N Y Nguyen, D S Dimitrov, and H Golding. Evidence for cell-surface association between fusin and the CD4-gp120 complex in human cell lines. *Science (New York, N.Y.)*, 274:602–605, 1996. ISSN 0036-8075. doi: 10.1126/science.274.5287.602.

[29] I Wilmut, A E Schnieke, J McWhir, A J Kind, and K H Campbell. Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385:810–813, 1997. ISSN 1536-2302. doi: 10.1089/clo.2006.0002.

[30] S M Willadsen. Nuclear transplantation in sheep embryos. *Nature*, 320(6057):63–65, March 1986. URL `http://dx.doi.org/10.1038/320063a0`.

[31] Nationale Human Genome Research Institute. Cloning factsheet. URL `http://www.genome.gov/25020028`.

[32] Wikipedia. Stem cell, . URL `http://en.wikipedia.org/wiki/Stem\_cell\#Key\_research\_events`.

[33] J. A. Thomson. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science*, 282 (5391):1145–1147, November 1998. doi: 10.1126/science.282.5391.1145. URL `http://www.sciencemag.org/content/282/5391/1145`.

[34] M J Shamblott, J Axelman, S Wang, E M Bugg, J W Littlefield, P J Donovan, P D Blumenthal, G R Huggins, and J D Gearhart. Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proceedings of the National Academy of Sciences of the United States of America*, 95:13726–13731, 1998. ISSN 0027-8424. doi: 10.1073/pnas.95.23.13726.

[35] Booma D Yandava, Lori L Billinghurst, and Evan Y Snyder. "Global" cell replacement is feasible via neural stem cell transplantation: Evidence from the dysmyelinated shiverer mouse brain. *Proceedings of the National Academy of Sciences*, 96(12):7029–7034, June 1999. doi: 10.1073/pnas.96.12.7029. URL `http://www.pnas.org/content/96/12/7029.abstract`.

[36] B E Petersen, W C Bowen, K D Patrene, W M Mars, A K Sullivan, N Murase, S S Boggs, J S Greenberger, and J P Goff. Bone Marrow as a Potential Source of Hepatic Oval Cells. *Science*, 284(5417):1168–1170, May 1999. doi: 10.1126/science.284.5417.1168. URL `http://www.sciencemag.org/content/284/5417/1168.abstract`.

[37] Oliver Brüstle, Kimberly N Jones, Randall D Learish, Khalad Karram, Khalid Choudhary, Otmar D Wiestler, Ian D Duncan, and Ronald D G McKay. Embryonic Stem Cell-Derived Glial Precursors: A Source of Myelinating Transplants. *Science*, 285(5428):754–756, July 1999. doi: 10.1126/science.285.5428.754. URL `http://www.sciencemag.org/content/285/5428/754.abstract`.

[38] Christopher R R Bjornson, Rodney L Rietze, Brent A Reynolds, M Cristina Magli, and Angelo L Vescovi. Turning Brain into Blood: A Hematopoietic Fate Adopted by Adult Neural Stem Cells in Vivo. *Science*, 283(5401):534–537, January 1999. doi: 10.1126/science.283.5401.534. URL `http://www.sciencemag.org/content/283/5401/534.abstract`.

[39] A J BECKER, E A McCULLOCH, and J E TILL. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature*, 197:452–454, 1963. ISSN 0028-0836. doi: 10.1038/197452a0.

[40] R A Gatti, H J Meuwissen, H D Allen, R Hong, and R A Good. Immunological reconstitution of sex-linked lymphopenic immunological deficiency. *Lancet*, 2:1366–1369, 1968. ISSN 01406736. doi: 10.1016/S0140-6736(68)92673-1.

[41] M. J. Evans and M. H. Kaufman. Establishment in culture of pluripotentital cells from mouse embryos. *Nature*, pages 154–156, 1981. ISSN 0028-0836. doi: 10.1038/292154a0.

[42] G R Martin. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 78:7634–7638, 1981. ISSN 0027-8424. doi: 10.1073/pnas.78.12.7634.

[43] K R Thomas and M R Capecchi. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*, 51:503–512, 1987. ISSN 00928674. doi: 10.1016/0092-8674(87)90646-5.

[44] Robert E Schwartz, Jonathan L Linehan, Matthew S Painschab, Wei-Shou Hu, Catherine M Verfaillie, and Dan S Kaufman. Defined conditions for development of functional hepatic cells from human embryonic stem cells. *Stem cells and development*, 14:643–655, 2005. ISSN 1547-3287. doi: 10.1089/scd.2005.14.643.

[45] Hans S Keirstead, Gabriel Nistor, Giovanna Bernal, Minodora Totoiu, Frank Cloutier, Kelly Sharp, and Oswald Steward. Human embryonic stem cell-derived oligodendrocyte progenitor cell transplants remyelinate and restore locomotion after spinal cord injury. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25:4694–705, 2005. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.0311-05.2005. URL `http://www.ncbi.nlm.nih.gov/pubmed/15888645`.

[46] A.D. Hershey and Martha Chase. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *The Journal of General Physiology*, pages 39–56, 1952. ISSN 0022-1295.

[47] F Sanger and A R Coulson. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology*, 94:441–448, 1975. ISSN 00222836. doi: 10.1016/0022-2836(75)90213-2.

[48] A M Maxam and W Gilbert. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74:560–564, 1977. ISSN 0027-8424. doi: 10.1073/pnas.74.2.560.

[49] R D Fleischmann, M D Adams, O White, R A Clayton, E F Kirkness, A R Kerlavage, C J Bult, J F Tomb, B A Dougherty, and J M Merrick. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science (New York, N.Y.)*, 269:496–512, 1995. ISSN 0036-8075. doi: 10.1126/science.7542800.

[50] C. elegans sequencing consortium. Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium [published erratum appears in Science 1999 Jan 1;283(5398):35]. *Science*, 282:2012–2018, 1998. URL `http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/referer?http://www.sciencemag.org/cgi/content/full/282/5396/2012`.

[51] J C Venter, M D Adams, E W Myers, P W Li, R J Mural, G G Sutton, H O Smith, M Yandell, C A Evans, R A Holt, J D Gocayne, P Amanatides, R M Ballew, D H Huson, J R Wortman, Q Zhang, C D Kodira, X H Zheng, L Chen, M Skupski, G Subramanian, P D Thomas, J Zhang, G L Gabor Miklos, C Nelson, S Broder, A G Clark, J Nadeau, V A McKusick, N Zinder, A J Levine, R J Roberts, M Simon, C Slayman, M Hunkapiller, R Bolanos, A Delcher, I Dew, D Fasulo, M Flanigan, L Florea, A Halpern, S Hannenhalli, S Kravitz, S Levy, C Mobarry, K Reinert, K Remington, J Abu-Threideh, E Beasley, K Biddick, V Bonazzi, R Brandon, M Cargill, I Chandramouliswaran, R Charlab, K Chaturvedi, Z Deng, V Di Francesco, P Dunn, K Eilbeck, C Evangelista, A E Gabrielian, W Gan, W Ge, F Gong, Z Gu, P Guan, T J Heiman, M E Higgins, R R Ji, Z Ke, K A Ketchum, Z Lai, Y Lei, Z Li, J Li, Y Liang, X Lin, F Lu, G V Merkulov, N Milshina, H M Moore, A K Naik, V A Narayan, B Neelam, D Nusskern, D B Rusch, S Salzberg, W Shao, B Shue, J Sun, Z Wang, A Wang, X Wang, J Wang, M Wei, R Wides, C Xiao, C Yan, A Yao, J Ye, M Zhan, W Zhang, H Zhang, Q Zhao, L Zheng, F Zhong, W Zhong, S Zhu, S Zhao, D Gilbert, S Baumhueter, G Spier, C Carter, A Cravchik, T Woodage, F Ali, H An, A Awe, D Baldwin, H Baden, M Barnstead, I Barrow, K Beeson, D Busam, A Carver, A Center, M L Cheng, L Curry, S Danaher, L Davenport, R Desilets, S Dietz, K Dodson, L Doup, S Ferriera, N Garg, A Gluecksmann, B Hart, J Haynes, C Haynes, C Heiner, S Hladun, D Hostin, J Houck, T Howland, C Ibegwam, J Johnson, F Kalush, L Kline, S Koduru, A Love, F Mann, D May, S McCawley, T McIntosh, I McMullen, M Moy, L Moy, B Murphy, K Nelson, C Pfannkoch, E Pratts, V Puri, H Qureshi, M Reardon, R Rodriguez, Y H Rogers, D Romblad, B Ruhfel, R Scott, C Sitter, M Smallwood, E Stewart, R Strong, E Suh, R Thomas, N N Tint, S Tse, C Vech, G Wang, J Wetter, S Williams, M Williams, S Windsor, E Winn-Deen, K Wolfe, J Zaveri, K Zaveri, J F Abril, R Guigó, M J Campbell, K V Sjolander, B Karlak, A Kejariwal, H Mi, B Lazareva, T Hatton, A Narechania, K Diemer, A Muruganujan, N Guo, S Sato, V Bafna, S Istrail, R Lippert, R Schwartz, B Walenz, S Yooseph, D Allen, A Basu, J Baxendale, L Blick, M Caminha, J Carnes-Stine, P Caulk, Y H Chiang, M Coyne, C Dahlke, A Mays, M Dombroski, M Donnelly, D Ely, S Esparham, C Fosler, H Gire, S Glanowski, K Glasser, A Glodek, M Gorokhov, K Graham, B Gropman, M Harris, J Heil, S Henderson, J Hoover, D Jennings, C Jordan, J Jordan, J Kasha, L Kagan, C Kraft, A Levitsky, M Lewis, X Liu, J Lopez, D Ma, W Majoros, J McDaniel, S Murphy, M Newman, T Nguyen, N Nguyen, M Nodell, S Pan, J Peck, M Peterson, W Rowe, R Sanders, J Scott, M Simpson, T Smith, A Sprague, T Stockwell, R Turner, E Venter, M Wang, M Wen, D Wu, M Wu, A Xia, A Zandieh, and X Zhu. The sequence of the human genome. *Science (New York, N.Y.)*, 291:1304–1351, 2001. ISSN 0036-8075. doi: 10.1126/science.1058040.

[52] J D McPherson, M Marra, L Hillier, R H Waterston, A Chinwalla, J Wallis, M Sekhon, K Wylie, E R Mardis, R K Wilson, R Fulton, T A Kucaba, C Wagner-McPherson, W B Barbazuk, S G Gregory, S J Humphray, L French, R S Evans, G Bethel, A Whittaker, J L Holden, O T McCann, A Dunham, C Soderlund, C E Scott, D R Bentley, G Schuler, H C Chen, W Jang, E D Green, J R Idol, V V Maduro, K T Montgomery, E Lee, A Miller, S Emerling, Kucherlapati, R Gibbs, S Scherer, J H Gorrell, E Sodergren, K Clerc-Blankenburg, P Tabor, S Naylor, D Garcia, P J de Jong, J J Catanese, N Nowak, K Osoegawa, S Qin, L Rowen, A Madan, M Dors, L Hood, B Trask, C Friedman, H Massa, V G Cheung, I R Kirsch, T Reid, R Yonescu, J Weissenbach, T Bruls, R Heilig, E Branscomb, A Olsen, N Doggett, J F Cheng, T Hawkins, R M Myers, J Shang, L Ramirez, J Schmutz, O Velasquez, K Dixon, N E Stone, D R Cox, D Haussler, W J Kent, T Furey, S Rogic, S Kennedy, S Jones, A Rosenthal, G Wen, M Schilhabel, G Gloeckner, G Nyakatura, R Siebert, B Schlegelberger, J Korenberg, X N Chen, A Fujiyama, M Hattori, A Toyoda, T Yada, H S Park, Y Sakaki, N Shimizu, S Asakawa, K Kawasaki, T Sasaki, A Shintani, A Shimizu, K Shibuya, J Kudoh, S Minoshima, J Ramser, P Seranski, C Hoff, A Poustka, R Reinhardt, and H Lehrach. A physical map of the human genome. *Nature*, 409:934–941,

2001. ISSN 0028-0836. doi: 10.1038/35057157.

[53] R Staden. A strategy of DNA sequencing employing computer programs. *Nucleic acids research*, 6:2601–2610, 1979. ISSN 0305-1048. doi: 10.1093/nar/6.7.2601.

[54] a Edwards and C Caskey. Closure strategies for random DNA sequencing. *Methods*, 3(1): 41–47, August 1991. ISSN 10462023. doi: 10.1016/S1046-2023(05)80162-8. URL http://linkinghub.elsevier.com/retrieve/pii/S1046202305801628.

[55] J C Roach, C Boysen, K Wang, and L Hood. Pairwise end sequencing: a unified approach to genomic mapping and sequencing. *Genomics*, 26:345–353, 1995. ISSN 08887543. doi: 10.1016/0888-7543(95)80219-C.

[56] C Napoli, C Lemieux, and R Jorgensen. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant cell*, 2:279–289, 1990. ISSN 1532-298X. doi: 10.1105/tpc.2.4.279.

[57] Alexander R. van der Krol, Leon A. Mur, Pieter de Lange, J. N M Mol, and Antoine R. Stuitje. Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect. *Plant Molecular Biology*, 14:457–466, 1990. ISSN 01674412. doi: 10.1007/BF00027492.

[58] Andrew Fire, SiQun Xu, Mary K Montgomery, Steven A Kostas, Samuel E Driver, and Craig C Mello. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, 391(6669):806–811, February 1998. ISSN 0028-0836. URL http://dx.doi.org/10.1038/35888http://www.nature.com/nature/journal/v391/n6669/suppinfo/391806a0_S1.html.

[59] Wikipedia. RNA Interference, . URL http://en.wikipedia.org/wiki/RNA_interference.

[60] Long-Cheng Li, Steven T Okino, Hong Zhao, Deepa Pookot, Robert F Place, Shinji Urakami, Hideki Enokida, and Rajvir Dahiya. Small dsRNAs induce transcriptional activation in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, 103: 17337–17342, 2006. ISSN 0027-8424. doi: 10.1073/pnas.0607015103.

[61] Derek M Dykxhoorn and Judy Lieberman. The silent revolution: RNA interference as basic biology, research tool, and therapeutic. *Annual review of medicine*, 56:401–423, 2005. ISSN 0066-4219. doi: 10.1146/annurev.med.56.082103.104606.

[62] Daniel Kim and John Rossi. RNAi mechanisms and applications. *BioTechniques*, 44:613–616, 2008. ISSN 0736-6205. doi: 10.2144/000112792.

[63] Patrick Sulem, Daniel F Gudbjartsson, Simon N Stacey, Agnar Helgason, Thorunn Rafnar, Kristinn P Magnusson, Andrei Manolescu, Ari Karason, Arnar Palsson, Gudmar Thorleifsson, Margret Jakobsdottir, Stacy Steinberg, Snaebjorn Palsson, Fridbert Jonasson, Bardur Sigurgeirsson, Kristin Thorisdottir, Rafn Ragnarsson, Kristrun R Benediktsdottir, Katja K Aben, Lambertus A Kiemeney, Jon H Olafsson, Jeffrey Gulcher, Augie Kong, Unnur Thorsteinsdottir, and Kari Stefansson. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*, 39(12):1443–1452, December 2007. ISSN 1061-4036. URL http://dx.doi.org/10.1038/ng.2007.13http://www.nature.com/ng/journal/v39/n12/suppinfo/ng.2007.13_S1.html.

[64] PR Burton, DG Clayton, and LR Cardon. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145):661–678, June 2007. ISSN 0028-0836. URL http://dx.doi.org/10.1038/nature05911http://www.nature.com/nature/journal/v447/n7145/suppinfo/nature05911_S1.htmlhttp://www.nature.com/nature/journal/v447/n7145/abs/nature05911.html.

[65] Jennifer Couzin and Jocelyn Kaiser. Closing the Net on Common Disease Genes. *Science*, 316(5826):820–822, May 2007. doi: 10.1126/science.316.5826.820. URL http://www.sciencemag.org/content/316/5826/820.short.

[66] RA Gibbs, JW Belmont, P Hardenbol, and TD Willis. The international HapMap project. *Nature*, 426(6968):789–796, December 2003. ISSN 0028-0836. URL http://www.nature.com/nature/journal/v426/n6968/suppinfo/nature02168_S1.htmlhttp://www.nature.com/ng/journal/v36/n11s/full/doifinder/10.1038\%2Fnature02168.

[67] Robert J Klein, Caroline Zeiss, Emily Y Chew, Jen-Yue Tsai, Richard S Sackler, Chad Haynes, Alice K Henning, John Paul SanGiovanni, Shrikant M Mane, Susan T Mayne, Michael B Bracken, Frederick L Ferris, Jurg Ott, Colin Barnstable, and Josephine Hoh. Complement Factor H Polymorphism in Age-Related Macular Degeneration. *Science*, 308(5720):385–389, April 2005. doi: 10.1126/science.1109557. URL http://www.sciencemag.org/content/308/5720/385.abstract.

[68] Hajime Matsuzaki, Shoulian Dong, Halina Loi, Xiaojun Di, Guoying Liu, Earl Hubbell, Jane Law, Tam Berntsen, Monica Chadha, Henry Hui, Geoffrey Yang, Giulia C Kennedy, Teresa A Webster, Simon Cawley, P Sean Walsh, Keith W Jones, Stephen P A Fodor, and Rui Mei. Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nat Meth*, 1(2):109–111, November 2004. ISSN 1548-7091. URL http://dx.doi.org/10.1038/nmeth718http://www.nature.com/nmeth/journal/v1/n2/suppinfo/nmeth718_S1.html.

[69] Wikipedia. International HapMap project, . URL http://en.wikipedia.org/wiki/International_HapMap_Project.

[70] T W Chang. Binding of cells to matrixes of distinct antibodies coated on solid surface. *Journal of immunological methods*, 65:217–223, 1983. ISSN 00221759. doi: 10.1016/0022-1759(83)90318-6.

[71] Kazutoshi Takahashi, Koji Tanabe, Mari Ohnuki, Megumi Narita, Tomoko Ichisaka, Kiichiro Tomoda, and Shinya Yamanaka. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5):861–72, November 2007. ISSN 0092-8674. doi: 10.1016/j.cell.2007.11.019. URL http://www.ncbi.nlm.nih.gov/pubmed/18035408.

[72] Junying Yu, Maxim A Vodyanik, Kim Smuga-Otto, Jessica Antosiewicz-Bourget, Jennifer L Frane, Shulan Tian, Jeff Nie, Gudrun A Jonsdottir, Victor Ruotti, Ron Stewart, Igor I Slukvin, and James A Thomson. Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.)*, 318(5858):1917–20, December 2007. ISSN 1095-9203. doi: 10.1126/science.1151526. URL http://www.ncbi.nlm.nih.gov/pubmed/18029452.

[73] Qiao Zhou, Juliana Brown, Andrew Kanarek, Jayaraj Rajagopal, and Douglas A Melton. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*, 455(7213):627–32, October 2008. ISSN 1476-4687. doi: 10.1038/nature07314. URL http://www.ncbi.nlm.nih.gov/pubmed/18754011.

[74] Kazutoshi Takahashi and Shinya Yamanaka. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, 126(4):663–676, June 2006. doi: 10.1016/j.cell.2006.07.024. URL http://www.cell.com/cell/abstract/S0092-8674(06)00976-7.

[75] Myron S Cohen, Ying Q Chen, Marybeth McCauley, Theresa Gamble, Mina C Hosseinipour, Nagalingeswaran Kumarasamy, James G Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, Jose H S Pilotto, Sheela V Godbole, Sanjay Mehendale, Suwat Chariyalertsak, Breno R Santos, Kenneth H Mayer, Irving F Hoffman, Susan H Eshleman, Estelle Piwowar-Manning, Lei Wang, Joseph Makhema, Lisa A Mills, Guy de Bruyn, Ian Sanne, Joseph Eron, Joel Gallant, Diane Havlir, Susan Swindells, Heather Ribaudo, Vanessa Elharrar, David Burns, Taha E Taha, Karin Nielsen-Saines, David Celentano, Max Essex, and Thomas R Fleming. Prevention of HIV-1 infection with early antiretroviral therapy. *The New England journal of medicine*, 365(6):493–505, August 2011. ISSN 1533-4406. doi: 10.1056/NEJMoa1105243. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3200068\&tool=pmcentrez\&rendertype=abstract.

[76] Jean-Francois Brunet, Francois Denizot, Marie-Francoise Luciani, Magali Roux-Dosseto, Marie Suzan, Marie-Genevieve Mattei, and Pierre Golstein. A new member of the immunoglobulin superfamily - CTLA-4. *Nature*, 328(6127):267–270, July 1987. URL http://dx.doi.org/10.1038/328267a0.

[77] Y Ishida, Y Agata, K Shibahara, and T Honjo. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *The EMBO journal*, 11(11):3887–95, November 1992. ISSN 0261-4189. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=556898\&tool=pmcentrez\&rendertype=abstract.

[78] Elizabeth A. Tivol, Frank Borriello, A.Nicola Schweitzer, William P. Lynch, Jeffrey A. Bluestone, and Arlene H. Sharpe. Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity*, 3 (5):541–547, November 1995. ISSN 10747613. doi: 10.1016/1074-7613(95)90125-6. URL http://www.sciencedirect.com/science/article/pii/1074761395901256.

[79] P. Waterhouse, J. M. Penninger, E. Timms, A. Wakeham, A. Shahinian, K. P. Lee, C. B. Thompson, H. Griesser, and T. W. Mak. Lymphoproliferative Disorders with Early Lethality in Mice Deficient in Ctla-4. *Science*, 270(5238):985–988, November 1995. ISSN 0036-8075. doi: 10.1126/science.270.5238.985. URL http://www.sciencemag.org.proxy.library.uu.nl/content/270/5238/985.

[80] Hiroyuki Nishimura, Masato Nose, Hiroshi Hiai, Nagahiro Minato, and Tasuku Honjo. Development of Lupus-like Autoimmune Diseases by Disruption of the PD-1 Gene Encoding an ITIM Motif-Carrying Immunoreceptor. *Immunity*, 11(2):141–151, August 1999. ISSN 10747613. doi: 10.1016/S1074-7613(00)80089-8. URL http://www.sciencedirect.com/science/article/pii/S1074761300800898.

[81] D R Leach, M F Krummel, and J P Allison. Enhancement of antitumor immunity by CTLA-4 blockade. *Science (New York, N.Y.)*, 271(5256):1734–6, March 1996. ISSN 0036-8075. URL http://www.ncbi.nlm.nih.gov/pubmed/8596936.

[82] James N Kochenderfer, Wyndham H Wilson, John E Janik, Mark E Dudley, Maryalice Stetler-Stevenson, Steven A Feldman, Irina Maric, Mark Raffeld, Debbie-Ann N Nathan, Brock J

Lanier, Richard A Morgan, and Steven A Rosenberg. Eradication of B-lineage cells and regression of lymphoma in a patient treated with autologous T cells genetically engineered to recognize CD19. *Blood*, 116(20):4099–102, November 2010. ISSN 1528-0020. doi: 10.1182/blood-2010-04-281931. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2993617\&tool=pmcentrez\&rendertype=abstract.

[83] Thomas Heinze, Philip Shapira, Juan D. Rogers, and Jacqueline M. Senker. Organizational and institutional influences on creativity in scientific research. *Research Policy*, 38(4):610–623, May 2009. ISSN 00487333. doi: 10.1016/j.respol.2009.01.014. URL http://linkinghub.elsevier.com/retrieve/pii/S0048733309000195.