

## The Impact of Sequencing Human Genome on Genetically Inherited Diseases

**Hameed Khan\***

Department of Genetics & Robotics, Senior Scientist, NCMRR (National Center for Medical Rehabilitation Research), National Institutes of Health (NIH), Adjunct Professor NYLF, Bethesda, Maryland, USA

**\*Corresponding author:** Hameed Khan, Department of Genetics & Robotics, Senior Scientist, NCMRR (National Center for Medical Rehabilitation Research), National Institutes of Health (NIH), Adjunct Professor NYLF, Bethesda, Maryland, USA

**Citation:** Khan H. (2023) The Impact of Sequencing Human Genome on Genetically Inherited Diseases. J Can Ther Res. 3(1):1-29.

**Received:** November 23, 2022 | **Published:** January 11, 2023

**Copyright** © 2023 by Khan H. All rights reserved. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

This abstract attempts to explore if genetically inherited diseases could be prevented before conception. Sequencing Egg and Sperm before conception and comparing with the Reference Sequence will identify any deleterious mutations which will predict the onset of any serious illness in the embryo. Male and female differ genetically by their respective sex chromosome composition, that is, XY as male and XX as female. Although both X and Y chromosomes evolved from the same ancestor pair of autosomes, the Y chromosome harbors male-specific genes, which play pivotal roles in male sex determination, germ cell differentiation, and masculinization of various tissues. While the X-chromosome of the mother is made of 164 million AT-GC nucleotide base pairs and carries 1,144 genes, the Y-chromosome of the father is made of 59 million A-T-GC base pairs and carries 231 genes. The next generation of Nanopore Sequencer will sequence the X and Y-chromosome cheaper, faster and with great precision and accuracy. To this date, more than 6,000 genetic disorders are identified, and new genetic disorders are constantly being added in the medical literature. Some of the confirmed disorders include Down syndrome (Trisomy 21), Fragile X syndrome, Klinefelter syndrome, Turner syndrome, Cystic Fibrosis, Thalassemia, Sickle Cell Anemia, Huntington's Disease, Duchenne's Muscular Dystrophy, Tay-Sachs Disease etc. Although some of these genetic diseases are treatable, most of them remain under investigation. Once a mutation related to a disease is identified, our next challenge is to design drugs to shut off that mutated gene. Since the language of life of all living

Aziridine, 2-4, Dinitrophenyl Benzamide (CB1954) to treat a solid aggressive tumor like the Walker carcinoma 256 in Rats and how similar rational was used to design Di-aziridine benzoquinone Carbamate (AZQ)(US Patent 4,233,215) to treat solid aggressive brain tumor like Glioblastoma in humans

### **Keyword**

Sequencing Genome; Mutations; Genetic Diversity; Genome diseases; Inbreeding; AZQ

### **A Note to my readers**

The Impact of Sequencing Human Genomes are a series of lectures to be delivered to the scholars of the National Youth League Forum (NYLF) and the International Science Conferences. NYLF scholars are the very best and brightest students selected from all over the USA and the world brought to Washington by Envision, an outstanding organization that provides future leaders of the world. I am reproducing here part of the lecture which was delivered at the International Science Conference that was PCS 6th Annual Global Cancer Conference held on November 15-16, 2019, in Athens, Greece.

### **Special notes**

I am describing below the use of highly toxic lethal chemical weapons (Nitrogen Mustard) which was used during WWI and its more toxic analogs developed as more toxic weapons during WWII. I described the use of Nitrogen Mustard as anti-cancer agents in a semi-autographical way to accept the responsibility of its use. When we publish research papers, we share the glory with colleagues and use the pronoun "We" but only when we share the glory not the misery. In this article by adding the names of my coworkers, the animal handlers, I will share only misery. The Safety Committee is interested to know who generated the highly lethal Chemical Waste, How much was it generated and how was it disposed. I accept the responsibility. The article below sounds semi-auto-biographical, it is, because I am alone responsible for making these compounds of Nitrogen Mustard, Aziridines and Carbamate. To get a five-gram sample for animal screening, I must start with 80 grams of initial chemicals for a four-step synthesis. To avoid generating too much toxic chemical waste, instead of using one experiment with 80 grams, I conducted 80 experiments with one gram sample, isolating one crystal of the final product at a time. The tiny amount of waste generated at each experiment was burned and buried at a safe place according to safety committee rules.

### **Introduction**

At the dawn of new millennium, we embarked on three great revolutions in science. A Quantum Revolution, a Computer Revolution, and a Genetic Revolution. These revolutions deal with matter, mind, and life.

## Quantum Revolution

The first is the Quantum revolution. In Quantum Revolution, we solved the mystery of matter. We found that total matter in the Universe is made of atoms. The quantum revolution allows us to understand the world of the very small and the fundamental properties of the matter. Our deepest understanding of the atomic world comes from the advent of the quantum theory. Our Universe is made of matter and matter is made of atoms. To understand the Universe, we study a single atom; we found that the atom is made of a positively charged Proton with Neutron in the central nucleus around which revolve the negatively charged electrons in a specific orbits carrying a unit amount of energy called the quantum energy. It is the nucleus that carries tremendous amount of energy. In fission reactions when we break the nucleus, we could release that energy? The nucleus is made of positively charged protons and neutral neutrons. Since positively charged protons repel each other, they are fused together by nuclear forces that hold the protons tightly together. When Hydrogen atoms fused together, they form Helium atoms releasing high energy neutron which travels at a speed of about 10 million meters per second or about three percent of the speed of light. The highly charged neutrons bombard a radioactive atom like Uranium atom, it undergoes fission, which releases energy and several more neutrons, which then split more Uranium atoms and release more energy and neutrons. This sets a nuclear chain reaction. A chain reaction refers to a process in which neutrons released in fission produce an additional fission in at least one further nucleus. This nucleus in turn produces more neutrons, and the process repeats until all atoms are broken releasing tremendous amount of energy at once. With a ten-pound pure Uranium-235, we could generate so much energy that a city like Hiroshima can be turned to ashes in minutes. Or if we control the chain-reaction by placing a rod of Boron or Cadmium which absorbs neutrons to control the chain-reaction, we could convert matter into the nuclear energy that could be converted into electricity. The process may be controlled (nuclear power) or uncontrolled (nuclear weapons).

Our deepest understanding of the atomic world comes from the advent of the quantum theory. It is the quantum revolution nearly a century ago that led to the 20th-century technological revolution, which was based on the transistor, laser, and atomic clock. These inventions gave us computers, optical fiber communication, cell phones and the global positioning system, all of which are vital to the world economy. The greatest achievement we made in Quantum revolution is that we learned to convert matter into energy. Energy that runs the engine of modern society.

## Computer revolution

The second is the computer revolution. It was the revolution of mind. It brought the information age. Using a binary code (zero and one), we wrote programs to store all the information we generated from the dawn of human civilization to the present day. We can not only store information, but also, we could retrieve, cut, and paste and move the information around with the speed of light. Our greatest accomplishment is that we have capture space time. We can communicate with anyone around the world within seven second.

## Genetic revolution

The third is the genetic revolution, the greatest revolution of them all. It solves the mystery of life. Now, we know with fair certainty how life could have evolved at some remote corner of the Earth about three

and a half billion years ago and how it crawled on evolutionary path to fly at will or stay still. In a three-billion-year journey across time, it reached us and help us develop our language our consciousness and our mind. We are the most intelligent of all living creatures on Earth. We are so intelligent that we ask questions about ourselves who we are and where have we all come from? What was it that made us this way? While hundreds of experiments were conducted to answer the first two questions, a single experiment answers the last question. It was the greatest experiment ever conceived by human mind. It will answer the most fundamental question we have asked ourselves since the dawn of human civilization; Who are we? Where have we all have come from and what was it that made us this way. The answers to these questions are embedded in the nucleus of every cell of our body. Our book of life is written in the nucleus of each cell. Reading the total genetic information that makes us is called the Human Genome. We decided to read our entire genome letter by letter, word by word and sentence by sentence under a project entitled The Human Genome Project. The goal of the human genome project is to determine the total number of base pairs nucleotides and the order in which they are arranged that make us human. A string of nucleotides called the Deoxyribose Nucleic Acid(DNA) and identifying how DNA got together to make genes and mapping, and sequencing all of the genes of the human genome from both a physical and a functional standpoint.

The Sequencing of the Human Genome which is not only reading the entire book of life of human being letter, by letter, word by word and sentence by sentence, chapter by chapter but also the order in which these letters are arranged called sequencing, is the greatest discovery of all times. The sequencing of the Human Genome will answer the most fundamental questions, we have asked ourselves since the dawn of human civilization; what it means to be human; what the nature of our memory and our conscientiousness is; our development from a single cell to a complete human being; the biochemical basis of our senses; the process of our aging; the scientific basis of our similarity and dissimilarity. Similarities that all living creatures from a tiny blade of grass to the mighty Elephants including man, mouse, monkey, mosquitos, and microbes are all made of the same chemical building blocks, the nucleotides, and yet we are so diverse that no two individuals are alike, even identical twins are not exactly identical; they grow up to become two separate individuals.

To understand the basis of all diseases, we must read and understand the total genetic information that makes us humans that is to read the genome of a healthy human being (called the Reference Sequence) and compare with the sick person's genome to identify the spelling errors in the genome responsible for causing the disease. That is how the story of our book of life begins: As we all know that we are the loving union of our parents. Our mother's egg receives our father's sperm, and we are conceived. The fertilized egg carries complete information to make us. More than seventy years ago, the Nobel Laureate, Irvin Schrödinger, was the first person to propose that the hereditary molecule must contain a "code-script" that determined "the entire pattern of the individual's future development and of its functioning in the mature state". This was the first clear suggestion that genes contained some kind of "code". Now, we know that the essence of life is information and genes are the bearers of that information, carrying it in a tiny, complex code inside every cell of our bodies. If we examine for comparison, the fertilized egg of a human, mouse, and monkey under a microscope. We observed that all fertilized eggs look the same and yet first fertilized egg carries the instructions to make a man, the second carries the information to make

a mouse and third carries the information to make a monkey. We are certain that there exists a secret code within those fertilized eggs to make different species; Schrodinger called that secret code, the Script Code, now known as the Genetic Code. To understand the secret code, we must examine the internal structure of the fertilized egg. We propose that we examine three "C". The first C stands for the chromosome, the coloring bodies present inside each cell. The traits to make man, mouse or monkey must be located on the Chromosomes. These traits must be held together tightly by the second C, the covalent bonds. As the living cells grow, they must have the ability to copy the instructions accurately that copying is the third C. The Genetic Code to make man, mouse or monkey must be written on the chromosomes. Based on this information Crick and Watson broke the Genetic Code and unlocked the secret of life. If we unlock the secret of life, we will understand how evolution puts the traits together over millennia to separate man, from mouse and mouse from monkey. By unlocking the secret of life, we can understand how the normal cells work and how the normal cells become abnormal leading to all diseases including cancerous.

On further examination, we found that the chromosomes are made of four chemicals and information is located on them and these four molecules are called nucleotide bases. These bases are made of Deoxy Ribonucleic Acid (DNA). DNA is made of a string of nucleotides. It is a storehouse of information and is made of the same four nucleotide bases and they are: Adenine (A), Thymine (T), Guanine (G), and Cytosine (C). According to Crick's Central Dogma, [1] the information flows from the DNA which is transcribed into RNA which is translated in Ribosome into proteins. RNA is converted into an active form and is transcribed into mRNA (or messenger RNA after splicing out noncoding DNA) and by converting Thymine to active form Uracil (U) and from a double stranded DNA to a single stranded RNA and where the sugar Deoxy Ribose is replaced by sugar Ribose. The mRNA is translated by Ribosome into proteins. Gene Expression begins in Ribosome when a 4-letter genetic text is converted to a three-letter Codon which code for a single amino acid. By comparing Gene Profiles of normal genes with mutated genes, one can identify with precision and accuracy the exact location of mutated (altered or damage) nucleotide responsible for causing the disease. Comparing Gene Profiles is an excellent diagnostic method which helps us design drugs to specifically shut off the mutated genes.

Seventy years ago, Schrödinger predicted secret code of life using such a poor resolution microscope that we don't even use in our high school today. Instead, we use electron microscope. We can magnify the same fertilized egg to a million times of its original size, almost the size of a house. What we observe inside the fertilized egg is very analogous to the house. The house has a kitchen; the cell has a nucleus. Suppose your kitchen has a shelf which contains 46 volumes of cookbooks which contain 24,000 recipes which carry instructions to cook food for your breakfast, lunch, and dinner. The nucleus in the fertilized egg contains 46 chromosomes; (23 from our mother and 23 from our father), which carry 24,000 chapters called genes. Genes are units of inheritance which code for all 20 amino acids. Hundreds of amino acids join to form a protein and thousands of proteins interact to make a cell. Millions of cells interact to make an organ and several organs interact to make a man or a mouse or a monkey. The number and the order of the nucleotides determine the composition of a species. [2-3].

If the cookbook in your kitchen is written in English language, it uses 26 letters, but the book of life of all

living creatures is written in 4 letters and they are A, T, G and C. These are the initials of four chemicals called nucleotides found the nucleus of all living cells. Nucleotides are made of sugar Ribose (Deoxy Ribose in DNA and Ribose in RNA), a phosphate group and one of the four Nitrogen bases, two Purines and two Pyrimidines and the Thymine is converted to Uracil in RNA. These molecules are found in the nucleus of all living cells from a tiny blade of Grass to mighty elephant including man, mouse, and monkey. The total genetic information to make any living creature is based on the above four-letter text and out of these four letters, only three letter Codon which carries the Genetic Code for an amino acid (such as GUU is for amino acid Valine, GCU is for Alanine, GAA is for Glutamine etc.) the building blocks for all proteins. Sixty-four codons code for 20 amino acids and codons for all 20 amino acids have been decoded. All living creatures use the same genetic code. A string of these nucleotides is called the DNA (Deoxy Ribonucleic Acid). Reading the number and the order of nucleotides are called genome sequencing [4-5].

As I said above, a gene is a piece of DNA. Out of four nucleotides text, three letters code for an amino acid called codon. A gene is made of several hundred codons. On a piece of DNA, a gene is identified by a start and a stop codon. The start codon is AUG which codes for the amino acid Methionine and there are three stop codons, and they are UAG, UGA, and UGG. The extension of DNA synthesis stops once the one of the three stop codons appears. A gene codes for a protein. We found that the smallest gene in bacterial genome and the largest gene is in Duchenne Muscular Dystrophy genome. To code for a protein, between start and stop codon, a gene has accumulated several hundreds to several thousand codons. A single mutation in the coding region of a single codon will alter the gene function. Mutation is caused by either exposure to radiations, chemical or environmental pollution, genetic inheritance, viral infections or DNA deletion, insertion relocation or inversion. Mutation can be good, bad, or neutral. A good mutation can convert a single cell organism to a multi cellular creature resulting in evolution. A bad mutation is responsible for coding for a wrong amino acid responsible for causing diseases. A neutral mutation can serve as a gene marker identifying its presence close to a good or a bad gene. By sequencing Human Genome, we have identified 16,000 good genes, 6,000 bad genes and 2,000 pseudo or neutral genes. Less than 2% of our genome codes for proteins. The remaining 98% of our genome carry pieces of DNA which serves as switches, promoters, enhancers etc. The greatest Darwinian transformation is controlled by switches. By switching on and off a gene, the body plan gene called the FOX gene can bring the evolutionary changes in the body. Genes code for protein, it is the switches that turn genes on or off.

Gene Expression begins in Ribosome when a 4-letter genetic text is converted to a three-letter Codon. By comparing Gene Profiles of normal genes with mutated genes, one can identify with precision and accuracy the exact location of mutated (altered or damage) nucleotide responsible for causing the diseases. Comparing Gene Profiles is an excellent diagnostic method which helps us design drugs to specifically shut off the mutated genes. Delivering drugs from injection site to the target site is the essential way of treating diseases. We cannot design novel drugs unless we find the abnormal mutations responsible for causing that disease. The reading of the total genetic information that make us human is called the Human Genome. The reading the entire book of our life is authorized by the US Congress under The Human Genome Project.

In 1990, US Congress authorized three billion dollars to NIH to decipher the entire Human Genome under

the title, "The Human Genome Project." We found that our genome contains six billion four hundred million nucleotides base pairs half comes from our father and another half comes from our mother. Less than two percent of our Genome contains genes which code for proteins. The other 98 percent of our genome contains switches, promoters, terminators etc. The 46 Chromosomes present in each cell of our body are the greatest library of the Human Book of Life on planet Earth. The Chromosomes carry genes which are written in nucleotides base pairs. Before sequencing (determining the number and the order of the four nucleotides on a Chromosomes), it is essential to know how many genes are present on each Chromosome in our Genome. The Human Genome Project has identified not only the number of nucleotide base pairs on each Chromosome, but also the number of genes on each chromosome [6].

The following list provide the details composition of each Chromosome including the number of nucleotides and the number of genes on each Chromosome:

We found that the Chromosome-1 is the largest Chromosome carrying 263 million A, T, G and C nucleotides bases and it has only 2,610 genes. The Chromosome-2 contains 255 million nucleotides bases and has only 1,748 genes. The Chromosome-3 contains 214 million nucleotide bases and carries 1,381 genes. The Chromosome-4 contains 203 million nucleotide bases and carries 1,024 genes. The Chromosome-5 contains 194 million nucleotide bases and carries 1,190 genes. The Chromosome-6 contains 183 million nucleotide bases and carries 1,394 genes. The Chromosome-7 contains 171 million nucleotide bases and carries 1,378 genes. The Chromosome-8 contains 155 million nucleotide bases and carries 927 genes. The Chromosome-9 contains 145 million nucleotide bases and carries 1,076 genes. The Chromosome-10 contains 144 million nucleotide bases and carries 983 genes. The Chromosome-11 contains 144 million nucleotide bases and carries 1,692 genes. The Chromosome-12 contains 143 million nucleotide bases and carries 1,268 genes. The Chromosome-13 contains 114 million nucleotide bases and carries 496 genes. The Chromosome-14 contains 109 million nucleotide bases and carries 1,173 genes. The Chromosome-15 contains 106 million nucleotide bases and carries 906 genes. The Chromosome-16 contains 98 million nucleotide bases and carries 1,032 genes. The Chromosome-17 contains 92 million nucleotide bases and carries 1,394 genes. The Chromosome-18 contains 85 million nucleotide bases and carries 400 genes. The Chromosome-19 contains 67 million nucleotide bases and carries 1,592 genes. The Chromosome-20 contains 72 million nucleotide bases and carries 710 genes. The Chromosome-21 contains 50 million nucleotide bases and carries 337 genes. Chromosome-22 contains 56 million nucleotides and carries 701 genes. Finally, the sex chromosome of all females called the (X) contains 164 million nucleotide bases and carries 1,141 genes. The male sperm chromosome contains 59 million nucleotide bases and carries 255 genes.

If you add up all genes in the 23 pairs of Chromosomes, they come up to 26,808 genes and yet we keep on mentioning 24,000 genes needed to keep us function normally. Out of 24,000 genes, we have identified 16,000 good genes, 6,000 bad genes and 2,000 Pseudo genes. As I said above, a gene codes for a protein, not all 24,000 genes code for proteins. It is estimated that less than 19,000 genes code for protein. Because of the alternative splicing, each gene codes for more than one protein. All functional genes in our body make less than 50,000 protein which interact in millions of different ways to give a single cell. Millions of cells interact to give a tissue, hundreds of tissues interact to give an organ and several organs interact to make a human.

Not all genes act simultaneously to make us function normally. Current studies show that a minimum of 2,000 genes are enough to keep human function normally; the remaining genes are backup support system, and they are used when needed. The non-functional genes are called the Pseudo genes. For example, millions of years ago, humans and dogs shared some of the same ancestral genes; we both carry the same olfactory genes, only in dogs they still function to search for food. Since humans don't use these genes to smell for searching food, these genes are broken and lost their functions, but we still carry them. We call them Pseudo genes. Recently, some Japanese scientists have activated the Pseudo genes, this work may create ethical problem in future as more and more Pseudo genes are activated. Nature has good reasons to shut off those Pseudo genes.

On April 3, 2003, we sequenced the entire Human Genome. We not only read the entire script of our genome, letter by letter, word by word, sentence by sentence, but also we read the number of letters and the order in which they are arranged (sequence) called under the title, "The Human Genome Project". We found that less than two percent of the Gene in our Genome codes for proteins and the rest is the non-coding regions which contains switches to turn the genes On or Off, pieces of DNA which act as promoters and enhancers of the genes. Using restriction enzymes (which act as molecular scissors), we can cut, paste, and copy genetic letters in the non-coding region which could serve as markers and which has no effect, but a slight change in the coding region makes a normal cell become abnormal or cancerous. Recent studies showed that mutations in switches, promoters and enhancers which are present in the non-coding regions are also responsible for some unusual diseases. We need to go back and look at these regions more carefully.

Our Genome provides the genetic road map of all our genes, past, present and future. For example, it can tell us how many good or bad genes we inherit from our parents and how many of those gene we are going to pass on to our children. If a family has too many bad genes, and have a family history of serious illnesses, you can break off the information flow and stop having children or stop donating mutated eggs and sperms.

Our search for unknown diseases has come to a closure:

There are two most powerful implications of the human Genome Sequencing. One of them is that we have come to closure. What it means is that we have the catalog of all genes in the Human Genome, we can search the entire genome and locate the desired gene. we will not wonder in the wilderness anymore. Everything there is to know about human health and traits are written on these genes in nucleotide sequences. Our Genomes provides the catalog of all genes.

We can scan the whole genome (Reference Sequence) for its response to a given situation. When we look at a normal cell and compare with an abnormal cell, we see the differences. Or when we compare their gene expression looking for a specific proteins, from a specific genes and for a specific nucleotide sequence, we can identify a specific disease. In the olden days, before sequencing human genome, a physician would order several tests and would say to his patient, I don't know what is wrong with you, I will see if any of these tests show if my guess is right and if it is wrong, he will recommend few more tests to see if he could identify the illness. The guesswork and the trial-and-error days are over. Now, after



sequencing the human genome, the physician would say I don't know what is wrong with you, but I do know where to find it. It is written in your Genome. It would be easy for a Physician to scan the patient entire genome and compare against the Reference Sequence to identify the mutations responsible for causing the disease. We will take a small blood sample of the patient, separate his WBC, extract DNA, sequence his Genome and compare with the Reference Sequence letter by letter, word by word by word and sentence by sentence, we can easily identify the mutations responsible for causing the disease. The result will provide the best diagnostic method to identify a disease.

Our Genome is not just a diagnostic road map of our genes, it also tells us to clone the good genes and shut off the bad genes. Using the good genes, it also tells us to make its large-scale protein for worldwide use such as Insulin and Human growth hormone. On the other hand, identifying the bad genes and tell us to design drugs to shut off bad genes responsible for causing Cancers, Cardiovascular disease, and Alzheimer. We have already demonstrated that using the genetic engineering techniques, we can cut, paste, copy, and sequence a good gene for industrial scale preparation such as Insulin to treat 300 millions of diabetic around the world. Similarly Human Growth Hormone, once available in minute quantities from the pituitary glands of humans' cadaver can now be produced in large amount in the Labs using the same genetic engineering method. Many valuable medicines are produced including Interleukin-II, for the treatment of the Kidney cancer, Factor-VIII for treating Hemophilia, Hepatitis B vaccine, Eleuthero protein for anemia, Whooping cough vaccine, and for Somatotropin for treating dwarfism.

Once the good and bad genes are identified, we know that the good genes code for good proteins which keep us healthy, and the bad genes produce bad protein that make us sick. Genome sequencing of bad genes start a new era of Genomic Medicine which is based on the genetic make-up of the individuals. The next step would be to design drugs to shut off the mutated genes. Gene Therapy will work if the disease is caused by a single gene mutation. Drug Therapy will work if multiple genes are responsible for causing diseases such as Cancers, Cardiovascular diseases, and Alzheimer.

### **Genetic Engineering and Recombinant Technology**

It involves cutting, pasting, copying, and sequencing DNA from the genome. As I said above, in humans, less than two percent of the Genome codes for proteins and the remaining 98 percent of Genome contains non-coding regions which carry switches, enhancers, promoters, inhibitors etc. Only two percent of the Genome transcribes into RNA. The non-coding regions of the RNA is spliced out into mRNA which carry three letter Codons. Each codon codes for a single amino acid except in the alternative splicing. It is the Codons which is translated in the Ribosome into proteins. The protein carries out our body functions as soon as it folds and becomes three dimensional. As the Central Dogma of Francis Crick describes that double stranded DNA replicates (makes its own copies) in the nucleus and it transcribes into the single stranded RNA as it leaves nucleus as mRNA in the cytoplasm (after splicing out non-coding sequence) which is translated in the Ribosomes into proteins. Information flows from both Good genes and bad genes from nucleus into the cell keeping the organism healthy or sick. Good proteins from good genes keep us healthy and bad proteins from mutated genes produce bad proteins that make us sick. The flow of information is continuous and uninterrupted. Information flows from both Good genes and bad genes from nucleus into the cell keeping the organism healthy or sick. Good proteins from good genes keep us

healthy and bad proteins from mutated genes produce bad proteins that make us sick.

One of the greatest challenges of the 21st century medicine is to cut out and purify a single gene from the string of genes from the entire human genome and insert this single human gene into a biological system making a recombinant fragment called Vector. Vectors protect genes from enzymatic destruction. Many microbial life form could serve as Vectors (such as bacteria or plasmids for smaller genes, and for larger genes using Phagemids, or Cosmids (Cosmid vectors are hybrids between plasmid and phage  $\lambda$  vectors. Cosmid vectors are designed to clone large fragments of DNA and to grow their DNA as a virus or as a plasmid. Cosmid. Vectors are used in homologous recombination between two different plasmids in the same cell. To scale up, the Vectors are harvested in Yeast or in bacteria. Later, large number of pure genes are isolated by breaking the Vectors with restriction enzymes.

Genes that carry instructions to made proteins are expressed in a two-step process and they are Transcription and Translation. First, the amino acids Codons are spliced or inserted into a double stranded DNA which is later transcribed into a single stranded m-RNA by removing non-coding nucleotides. As I said above, it is the m-RNA which is translated in the Ribosomes into all 20 amino acids. The Cells decode m-RNA in groups of three nucleotides called Codons which carry instructions to produce the amino acids. It is also worth repeating that when double stranded DNA is transcribed into a single stranded m-RNA, the nucleotide Thiamin is converted to Uracil. The Methyl group of Thiamine is replaced by a more water-soluble Hydroxyl group forming the Uracil. The nucleotide T for Thiamin is replaced by U for the Uracil. The m-RNA is translated into amino acids in Ribosomes. The gene expression has a Start Codon (AUG) which codes for amino acid Methionine and there are three Stop Codon which are UGG, UAG and UGA. Once the Stop Codon appears at the tail end of the DNA, amino acids synthesis stops. The Codons for each essential amino acid and their alternative codons are described below:

Valine (GTT, GTC, GTA, GTG), Leucine (CTT, CTC, CTA, CTG; TTA, TTG), Isoleucine (ATT, ATC, ATA), Phenylalanine (TTT, TTC), Tryptophan (TGG), Lysine (AAA, AAG), arginine (CGT, CGC, CGA, CGG; AGA, AGG), Histidine (CAT, CAC), Methionine (ATG), Threonine (ACT, ACC, ACA, ACG). In the past, all that repetition made it impossible to assemble some chopped-up pieces of DNA in the correct order. It's like having identical puzzle pieces – scientists didn't know which went where, leaving big gaps in the genomic picture. Now, with arrival of new sequencers, we read the book of our life cheaper and faster. The key advances included rapid improvements in the next generation gene sequencing machines made by Oxford Nanopore Technologies and Pacific Biosciences. The new Nanopore machine can accurately read a million letters of DNA at a time. Pacific Biosciences introduced a new sequencing machine which generate long-read sequencing that could read DNA sequence with greater than 99 percent accuracy. As the cost of sequencing goes down, all of us should have our genome sequenced. This genetic profile will serve as a part of our medical record. It will provide instant medical help.

## Genomic Medicine

(What to do with good genes? cut, paste, and copy the gene)

Out of 24 thousand genes in our genome, we carry sixteen thousand good genes; six thousand bad genes or mutated genes which are responsible for six thousand diseases and two thousand Pseudo genes which

are broken and have lost their functions. We also learned that about a minimum of 2,000 genes are essential to maintain a life and the remaining genes are backup or supporting genes to provide immediate help if needed. Using the Restriction Enzymes, like EcoR1, we could cut out all good and bad genes from the human genome to sequence and clone them by PCR. We need to study bad genes as well as good genes to design drugs to shut off defected genes. Diseases are caused by a single defected gene (called Mendelian disorders, for example, occur when specific mutations in single genes — called germline mutations — are inherited from either of one's two parents. Well-known examples of single genetic defects called the Mendelian diseases include cystic fibrosis, sickle cell disease, and Duchenne muscular dystrophy or multiple genetic defects (such as Cancers, Cardiovascular diseases or Alzheimer).

## Gene Therapy

In Gene Therapy a single bad gene is replaced by a good gene. With single genetic defects, we could cut, paste, and copy good gene in a Vector such as flu virus and infect patients' WBC with the transgenic virus which replace bad gene with the good gene. In a single genetic defect, Gene Therapy works supremely well as in the treatment of SCID (Severe Combined Immune Deficiency Syndrome) which is caused by a single genetic defect.

The supremely successful example of Gene Therapy is the treatment of Severe Combined Immune Deficiency Syndrome (SCID). SCID is also known as the Bubble Boy Syndrome. Because of a single genetic defect. patients lack the ability to fight off environmental microbes. To protect them from any infection, they are provided purified air in an enclosed environment. French Anderson and Mike Blaise at our Institute (NIH), using genetic engineering technique replaced the defected gene with a healthy gene. After harvesting in a harmless virus, infected the WBC obtained from the patient. The trans-genic WBC in vitro, returned the modified WBC to the same patient. Patients recovered and lives a normal life. With multiple genetic defects such as in cancers, Gene Therapy will not work, but Drug Therapy will work. We could design drugs to shut off multiple genes and prevent it from producing bad protein or cancers.

The first step is to cut the human genome at a specific site with a specific enzyme (prepare a Restriction Site Map) at the specific sites using restriction enzymes (which serves as molecular scissors such as EcoR1) first accomplished by El Salvador Luria, Max Delbruck, and Hamilton Smith. The fragment of human DNA (a single gene) if not protected will be destroyed by antibody. A naked gene is a piece of DNA (which has a start codon AUG and after a few thousand nucleotide (codons) end at one of the three stop codons UAG, UGA or UGG). If a gene is not protected by recombinant technology (making a hybrid) that is by recombining with the DNA of Virus, or Plasmids, or Chloroplasts (for plants) which serves as Vectors, will be destroyed by enzymes. One can store the fragments or genes in the Vectors once the human DNA fragment is stabilized in Vectors by recombinant technology; we can not only purify this fragment (genes), but also, we can make millions of copies of this fragment of DNA by transferring into the host cells such as Bacteria, mammalian cells or Yeast cell which autonomously replicates to produce library of genes. Each Library contains millions of copies of identical genes that produce same protein. Before the genetic revolution, Insulin is extracted from pancreas of the slaughtered animals which is used to treat old diseases such as diabetes; a tiny fragment of impurity could set anaphylactic shock and kill the patients. Now, highly pure human Insulin is produced by Genetic Engineering Technique and is used to treat 300

million diabetic patients worldwide without the loss of a single life using the same recombinant technology. Other products of Genomic Medicine such as Growth hormones and hormone proteins to treat Hemophilia by factor VIII protein are being developed as genomic medicines by recombinant technology.

## **Drug Therapy**

(What to do with bad Genes? Either to replace by Gene Therapy or to design drugs to shut off bad genes). The essence of life is information, and the information is located on the four nucleotide bases A-T and G-C. According to Central Dogma of Crick and Watson, the information on DNA is transcribed on RNA which is translated in Ribosome to protein. Attempts are being made to design drugs to attack cancer cells on all three level that is DNA, RNA and Protein. Herceptin, a novel class of drug, has been successful in attacking protein. Craig Milo has designed double stranded RNA to shut off gene and prevents its translation into protein. Attack on DNA to shut off a gene presents the greatest challenge because of the number of mutations on cancer cells.

### **How to design drugs to shut off bad gene variants?**

We design drugs to shut off genes responsible for causing cancers because largest amount of funds is available to the National Cancer Institute (NCI) about \$5B per year. Cancer is the leading cause of death and has surpassed the death of cardiovascular diseases. Over 636,000 people died of cancer; 1.9 million new cases will be diagnosed this year including 78,000 Prostate Cancer, 40,000 Breast cancer, 16000 Lung and Bronchus Cancer and 15,000 Colon and Rectal Cancer. Once diagnosed mutation by Gene Sequencing, the next step is to design drug to shut off those genes. I describe below my own work how I designed drugs to shut off cancer causing genes in animals and then translated the work in humans. (see special note-II above).

### **The rational drug design to treat cancers**

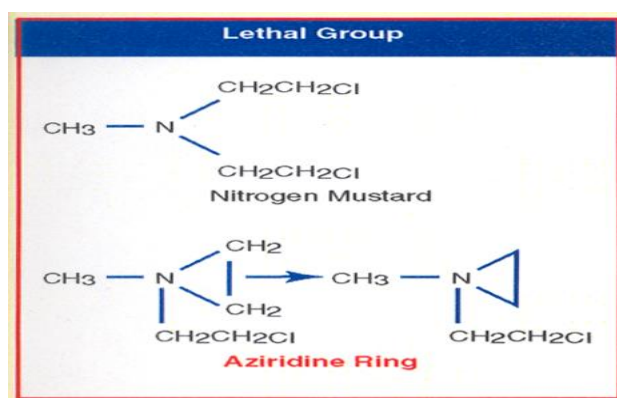
All three old age diseases that is Cancer, Cardiovascular Diseases and Alzheimer carry multiple mutated genes responsible for causing these diseases. In each of the above three diseases, it is the harmful mutated genes that code for wrong protein which causes these diseases. If we design drugs to shut off mutated genes in one disease, using the same rationale, we should be able to shut off bad genes in all three old age diseases. Although Coronary Artery disease is a complex disease, researchers have found about 60 genomic variants that are present more frequently in people with coronary artery disease. Most of these variants are dispersed across the genome and do not cluster at one specific chromosome. During WWI, highly toxic chemicals are used as chemical weapons. These chemicals attack the DNA shutting off gene. After the war, scientists modified these chemicals to shut off malignant genes to treat diseases. To reduce toxicity, prodrugs are designed to seek out the specific malignant genes which replicate faster producing acids. Aziridines and Carbamate moieties are prodrugs, stable in neutral and basic media, but sensitive to acid. Drugs carrying Aziridines, and Carbamate moieties are broken down in acidic media generating Carbonium ions which attack DNA shutting off genes. Only the acid producing genes will be attacked no matter where they are located. It does not matter whether they are clustered or dispersed across genome.

## Shutting off mutated genes by cross-linking double stranded dna

### (By Nitrogen Mustard)

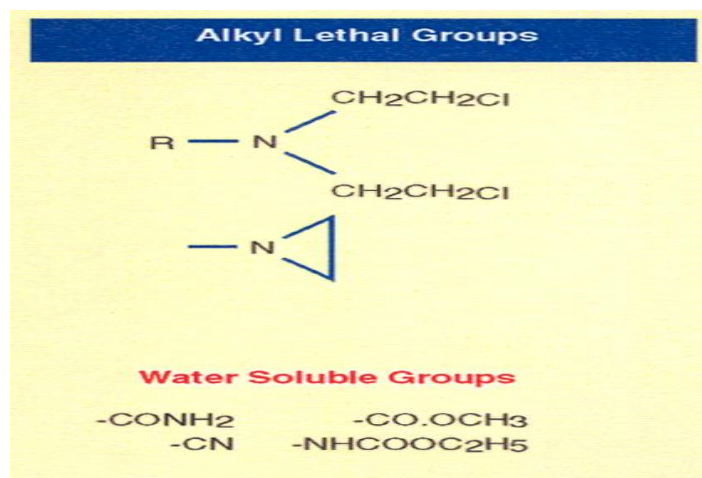
The supreme intellect for drug design is Ross, an Englishman, who is a Professor of Chemistry at the London University, England. Professor WCJ Ross is also the Head of Chemistry Department at the Royal Cancer Hospital, a post-graduate medical center of the London University. Ross was the first person who designed drugs for treating Cancers. He designed drugs to cross-link both strands of DNA that we inherit one strand from each parent. Cross-linking agents such as Nitrogen mustard are extremely toxic and were used as chemical weapon during the First World War (WWI). More toxic derivatives were developed during the Second World War (WWII). Using data for the toxic effect of Nitrogen Mustard on soldiers during WWI, Ross observed that Soldiers exposed to Nitrogen Mustard showed a sharp decline of White Blood Cells (WBC) that is from 5,000 cell/CC to 500 cells/CC. He immediately realized that children suffering from Childhood Leukemia have a very high WBC count that is over 90,000 cells/CC. In sick children, most of the WBCs are premature, defected, and unable to defend the body from microbial infections. Ross rationale was that cancer cells divide faster than the normal cells, by using Nitrogen Mustard to cross linking both strands of DNA, one can control and stop the abnormal WBC cell division in Leukemia patients. It was indeed found to be true. Professor Ross was the first person to synthesize many derivatives of Nitrogen Mustard. By using an analog of Nitrogen Mustard, called Chlorambucil, he was successful in treating Childhood Leukemia. In America, two Physicians named Goodman and Gilman from the Yale University were the first to use Nitrogen Mustard to treat cancer in humans. Nitrogen Mustards and its analogs are highly toxic. Ross was a Chemist, over the years, he synthesized several hundred derivatives of Nitrogen Mustard to modify toxicity of Nitrogen Mustard [6-8].

Although analogs of Nitrogen Mustard are highly toxic, they are more toxic to cancer cells and more cancer cells are destroyed than the normal cells. Toxicity is measured as the Chemotherapeutic Index (CI) which is a ratio between toxicity to Cancer cells versus the toxicity to Normal cells. Higher CI means that the drugs are more toxic to cancer cells. Most cross-linking Nitrogen Mustard have a CI of 10 that is they are ten times more toxic to cancer cells. Some of the Nitrogen Mustard analogs Ross made over the years are useful for treating cancers such as Chlorambucil for treating childhood leukemia (which brought the WBC level down to 5,000/CC). Children with Childhood Leukemia treated with Chlorambucil showed no sign of Leukemia even after 20 to 25 years later. Chlorambucil made Ross one of the leaders of the scientific world. He also made Melphalan and Myrophine for treating Pharyngeal Carcinomas [9- 13].



### DNA binding aziridine group

Nitrogen Mustard neither have selectivity nor specificity. They attacked all dividing cells including normal cells. During the study of the mechanism of action of radiolabeled Nitrogen Mustard on DNA, it was discovered that the two arms of Nitrogen Mustard do not bind to the double stranded DNA simultaneously. It binds to one strand of DNA at a time. The carbonium ion of the other arm of Nitrogen mustard attacks its own Nitrogen atom forming a stable three-member aziridinium ion. (see above chart). We were unable to isolate the aziridinium ion as growing tumor which produces acid which break down aziridinium ion to produce a second carbonium ion which attacks the second strand of DNA. We were able to isolate cross-linking DNA product. This study showed that to attack a single strand of DNA, we must synthesize Aziridine in the Lab. Synthesis of Aziridine analogs will give two advantages over Nitrogen Mustard: first, instead of cross-linking both strands of DNA, Aziridine binds to one strand of DNA, reducing its toxicity of double strand Nitrogen Mustard by half. Second, it gives selectivity, the Aziridine ring opens only in the acidic medium. Once the active ingredient Aziridine was determined to attack DNA, the next question was what drug delivery method should be used to deliver Aziridine at the tumor site.



The above structures are Nitrogen Mustard (2-bischloroethyl methyl amine) and Aziridine.

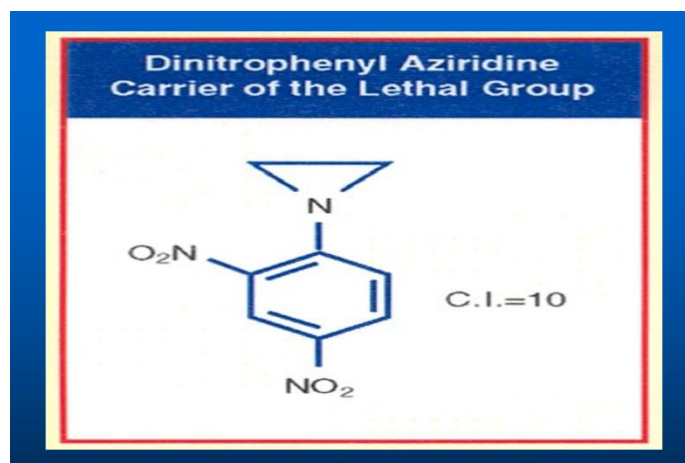
### DNA binding lethal groups

As a part of my doctoral thesis, I was assigned a different path. Instead of cross-linking DNA by Nitrogen Mustard, I am to design drugs to attack only one strand of DNA by making Aziridine analogues. We decided to use Aziridine moiety that would be an excellent active component to shut off a gene by binding to a single strand of DNA. To deliver Aziridine to the target site DNA, we decided to use Dinitrophenyl moiety as a delivery agent because its analog Dinitrophenol disrupt the Oxidative Phosphorylation of the ATP (Adenosine Triphosphate) which provides energy to perform all our body functions. To provide energy to our body function, the high energy phosphate bond in ATP is broken down to ADP (Adenosine Diphosphate) which is further broken down to AMP (Adenosine Mono Phosphate), the enzyme Phosphokinase put the inorganic phosphate group back on the AMP giving back the ATP. This cyclic

process of Oxidative Phosphorylation is prevented by Dinitrophenol. I decided to use Dinitrophenol as drug delivery method for the active ingredient Aziridine. Dinitrophenol also serves as a dye which stains a tumor called the Walker Carcinoma 256, a solid and most aggressive tumor in Rat. The first molecule I made by attaching the C-14 radiolabeled Aziridine to the dinitrophenol dye. The Dinitrophenyl Aziridine was synthesized using Dinitrochlorobenzene with C-14 radiolabeled Aziridine in the presence of Triethyl amine which removes the Hydrochloric Acid produced during the reaction. When the compound Dinitrophenyl Aziridine was tested against the implanted experimental animal tumor, the Walker Carcinoma 256 in Rats, it showed a TI (Therapeutic Index) of ten. The TI was like most of the analogs of Nitrogen Mustard. Since this Aziridine analog was not superior to Nitrogen Mustard, it was dismissed as unimportant.

### Structure Activity Relationship

Reexamination of the X-ray photographs showed that most of the radioactivity was concentrated at the injection site. Very little radioactivity was observed at the tumor site. It was obvious that we need to make derivatives of Dinitrophenyl Aziridine to move the drug from the injection site to the tumor site. Because of the lack of an effective drug delivery method, Dinitrophenyl Aziridine stays at the injection site. A very small amount of radioactivity was found on the tumor site.

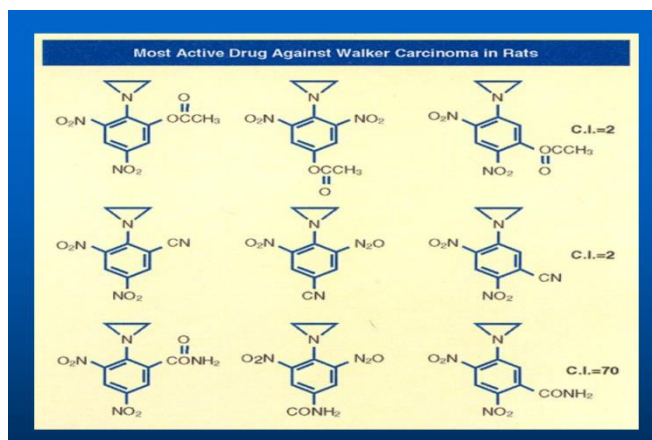


Dinitrophenyl benzamide a novel drug delivery molecule for Aziridine

I immediately realized that by making water and fat-soluble analogs of Dinitrophenyl Aziridine, I should be able to move the drug from the injection site to the tumor site. To deliver 2,4-Dinitrophenylaziridine from the injection site to tumor site, I could alter the structure of 2,4-Dinitrophenylaziridine by introducing the most water-soluble group such as ethyl ester to least water-soluble group such as Cyano- group or to introduce an intermediate fat/water double Amido group.

An additional substituent in the Dinitrophenyl Aziridine could give three isomers, Ortho, Meta, and Para substituent. Here conformational chemistry plays an important role in drug delivery. Ortho substituent always give inactive drug. Model building showed that because of the steric hinderance, Aziridine could not bind to DNA shutting off the genes. On the other hand, Meta and Para substituents offer no steric hinderance and drug could be delivered to DNA. The following chart showed that I synthesized all nine C-14 radio labeled analogs of 2,4-Dinitrophenyl aziridines and tested them against implanted Walker

## Carcinoma 256 in Rats.



Derivatization of Dinitro phenyl Benzamide based on Partition Coefficient

### The most water-soluble substituent

The first three compounds on top line of the above chart carry all three isomer of most water-soluble Ethyl Ester group attached to 2,4-Dinitrophenyl aziridine. The compound in vivo is hydrolyzed ester to produce most water-soluble carboxylic group. Within 24 hours of injection, the entire radioactive compound was extracted from the Rat's urine washed down from the cages. Since the Ortho position was not available for DNA binding, it showed no biological activity, but the third compound in which one of the Ortho position was free to bind to DNA showed some activity.

### The least water-soluble substituent

On the other hand, when the least water-soluble Cyano-group was attached to all three isomers of the 2,4-Dinitrophenyl aziridine compound as shown in the second line of the above chart, most of the compound stayed at the injection site. Only the last Cyano-derivative attached to DNA showed some activity.

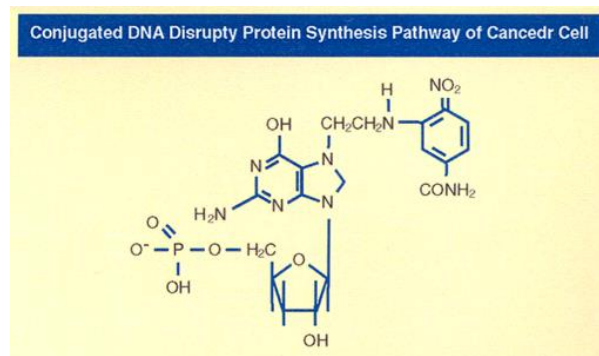
### The moderately soluble substituent

The last line of the above chart showed that the first two Amido groups were sterically hindered and did not bind to DNA and showed no biological activity, but the last compound presents the perfect drug delivery method. The entire drug was delivered from the injection site to the tumor site. The drug 1-Aziridine, 2,4-dinitro, 5-benzamide (CB1954) showed the highest biological activity. It has a CI of seventy; it is seventy times more toxic to cancer cells, highest toxicity ever recorded against Walker Carcinoma 256 in Rats. [14-16].

Nitrogen Mustards are highly toxic because they have neither specificity nor selectivity. They attack all dividing cells whether they are normal or abnormal. On the other hand, the analogs of Aziridines and Carbamates serve as prodrug and remain inactive in the basic and neutral media. They become activated only in the presence of acid producing cancer cells. Aziridine attacks DNA in acidic medium, particularly the N-7 Guanine. The dye Dinitro benzamide has great affinity for Walker Tumor. The Aziridine Dinitro benzamide (CB1954) stain the tumor. As the tumor grows, it uses Glucose as a source of energy. Glucose is broken down to Lactic Acid. It is the acid which activates the Aziridine ring. The ring opens to generate

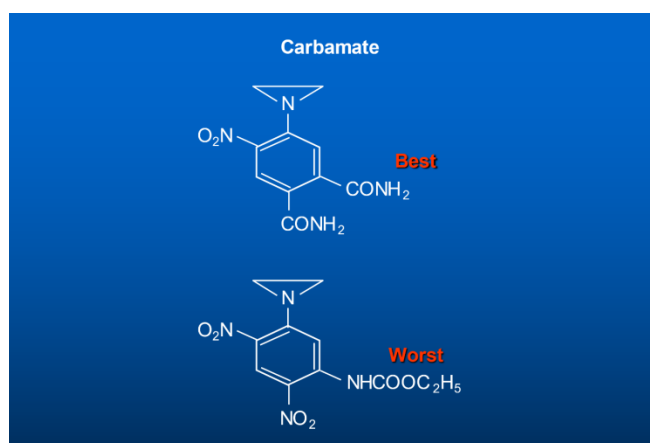


a carbonium ion which attacks the most negatively charged N-7 Guanine of DNA (as shown below) shutting off the Walker Carcinoma gene in Rat. The following conjugate structure show how CB1954 binds to a single stranded of DNA shutting off the gene.



Conjugated DNA Disrupting Protein Synthesis Pathway of Cancer Cell

For the discovery of CB1954, The University of London, honored with the Institute of Cancer Research (ICR) post-doctoral award to synthesize more analogs of CB1954. To improve drug delivery method, over the years, I made over a hundred additional analogs of Dinitro phenyl aziridine. To continue my work, I was honored with the Institute of Cancer Research Post-Doctoral Fellowship Award of the Royal Cancer Hospital of London University. To increase the toxicity of CB1954 to Walker Carcinoma, I made additional 20 analogs as a postdoctoral fellow. When I attached one more Carbonium ion generating moiety, the Carbamate moiety to the Aziridine Dinitrobenzene, the compound Aziridine Dinitro benzamide Carbamate was so toxic that its Therapeutic Index could not be measured. Although it is one of the most toxic compounds ever made, it is most useful in identifying and removing the resistant cell lines from the Tumor Bank. Further work in London University was discontinued for safety reason. [17-18].



The Best and the Worst Dinitro phenyl Aziridine Analogs

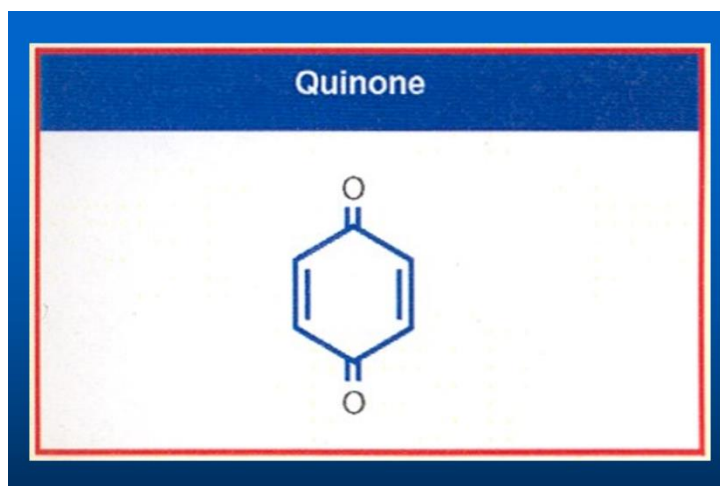
I continued my work on the highly toxic Aziridine/Carbamate combination in America when I was offered the Fogarty International Fellowship Award to continue my work at the National Cancer Institute (NCI) of the National Institutes of Health (NIH). I brought the idea from London University of attacking one strand of DNA using not only Aziridine, but also Carbamate without using the same dye Dinitro benzamide. My

greatest challenge at NCI is to translate the animal work which I did in London University to humans.

## Translating Animal Work to Humans

### Designing drugs to treat Glioblastoma: the human brain cancers

One day, I heard an afternoon lecture at the NIH in which the speaker described that radio labeled Methylated Quinone crosses the Blood Brain Barrier (BBB) in mice. When injected in mice, the X-ray photograph showed that the entire radioactivity was concentrated in the Mice's brain within 24 hours. I immediately realized that Glioblastoma multiforme, the brain tumor in humans, is a solid aggressive tumor like Walker Carcinoma in Rats. I decided to use Quinone moiety as a novel drug delivery molecule to cross BBB delivering Aziridine rings to attack Glioblastomas. By introducing an additional Carbamate moiety, I could increase its toxicity several folds. I planned to use this rationale to translate animal work to human by introducing multiple Aziridine and Carbamate moieties to the Quinone to test against Glioblastomas in humans.



The Structure of a non-toxic and non-addictive Quinone used for crossing the Blood Brain Barrier (BBB)

Glioblastoma (GBM) is a *primary* type of brain cancer which originates in the brain, rather than traveling to the brain from other parts of the body, such as the lungs or breasts. GBM is also called glioblastoma multi for me which is the most common type of primary brain cancer in adults. Attaching Nitrogen Mustard group to Quinone will produce highly toxic compound which will have neither specificity nor selectivity. Such a compound will attack all dividing cells whether they are normal or abnormal. On the other hand, the analogs of Aziridines and Carbamates are pro-drugs and remain inactive in the basic and neutral media. They become activated only in the presence of acid producing cancer cells.

### DNA Binding Aziridines

I continued my work on the highly toxic Aziridine/Carbamate combination in America. I brought the idea from London University of attacking one strand of DNA using not only Aziridine, but also Carbamate without using the same dye Dinitro benzamide. I decided to use Quinone moiety as a carrier for Aziridine rings to attack Glioblastomas. By introducing an additional Carbamate moiety, I could increase its toxicity by several folds. I planned to use this rationale to translate animal work to human by introducing multiple Aziridine and Carbamate moieties to the Quinone to test against Glioblastomas in humans. By attaching

two Aziridines and two Carbamate moieties to Quinone, the most useful Diaziridine Dicarbamate Quinone, I named this novel compound AZQ. Over the years, I made 45 analogs of AZQ. They were all considered valuable enough to be patented by the US Government (US Patent 4,233,215). By treating brain cancer with AZQ, we observed that Glioblastoma tumor not only stops growing, but it also starts shrinking. I could take care of at least one form of deadliest old age cancers, Glioblastomas. Literature search showed that AZQ is extensively studied.

As I said above, Glioblastomas, the brain cancers, is a solid and aggressive tumor and is caused by mutations on several chromosomal DNA. Deleterious mutations of DNA are the result of damaging DNA nucleotides by exposure to radiations, chemical and environmental pollution, viral infections, or genetic inheritance. The other factors responsible for causing DNA mutations are due to the fast rate of replication of DNA. For example, the bacteria E-coli grows so rapidly that within 24 hours, a single cell on a petridish containing nutrients forms an entire colony of millions when incubated on the Agar Gel. Mistakes occur in DNA during rapidly replication such as Insertion of a piece of DNA, Deletion, Inversion, Multiple Copying, Homologous Recombination etc. When an additional piece of nucleotide is attached to a DNA string, it is called Insertion, or a piece of DNA is removed from the DNA string; it is called Deletion or structural Inversion, translocation of DNA is also responsible for mutations. Since the gene in a DNA codes for Proteins, Insertion and Deletion on DNA have catastrophic effects on protein synthesis.

With the Quinone ring as a carrier across BBB, I could introduce different combinations of Aziridine rings and Carbamate moieties to Quinine and could create havoc for Glioblastomas. My major concern was how toxic this compound would be to the normal human brain cells. Fortunately, brain cells do not divide, only cancer cells divide. Glioblastomas represent such an example. In Glioblastomas, three major changes occur on Chromosomes (C-7, C-9 & C-10) and two minor changes occur on Chromosomes (C-1 & C-19). These mutations are responsible for causing brain cancers in humans. In a normal human cell, Chromosome-7 which is made of 171 million nucleotide base pairs and carries 1,378 genes. When Insertion occurs on Chromosome-7. Ninety-seven percent of Glioblastoma patients are affected by this mutation. On the other hand, a different mutation occurs on Chromosome-9 which is made of 145 million nucleotide base pairs and carries 1,076 genes. A major Deletion of a piece of DNA occurs on Chromosome-9 which results in eighty- three percent patients who are affected by this mutation. A minor Deletion of DNA also occurs on Chromosome-10 which is made of 144 million base pairs, and carries 923 genes. Although it is a minor deletion of a piece of DNA and yet it contributes to ninety-one percent patients with Glioblastoma. To a lesser extent, small mutation occurs on Chromosome-1 (the largest Chromosome in our Genome). It is made of 263 million nucleotide base pairs and carries 2,610 genes) and Chromosome-19 (it is made of 67 million base pairs and carries 1,592 genes) is also implicated in some forms of Glioblastomas.

All known Glioblastomas causing genes are located on five different chromosomes and carries a total of 9,579 genes. It appears impossible to design drugs to treat Glioblastomas since we don't know which nucleotide on which gene and on which chromosome is responsible for causing the disease. Besides identify the site of binding of DNA with C-14 radiolabeled studies, we can also confirm by comparing with the mega sequencing project. With the completion of 1,000 Human Genome Project, it becomes easier. By simply comparing the patient's genome with the sequencing of one thousand genomes, letter by letter,

word by word and sentence by sentence, we could identify the difference called the variants with precision and accuracy, the exact variants, or mutations responsible for causing the disease. Once the diagnosis is confirmed, the next step is how to treat the disease.

One of the greatest challenges of nanotechnology is to seek out the very first abnormal cell in the presence of billions of normal cells of our brain and shut off the genes before it spread. I worked on this assignment about a quarter of a century; conducted over 500 experiments which resulted in 200 novel drugs. They were all tested against experimental animal tumors. Forty-five of them were considered valuable enough to be patented by the US Government (US Patent 4, 146, 622 & 4,233,215). One of them is AZQ which not only stop the growth of Glioblastoma, but also the tumor start shrinking. For the discovery of AZQ, I was honored with, "The 2004 NIH Scientific Achievement Award." One of America's highest Award in Medicine. I was also honored with the India's National Medal of Honor, "Vidya Ratna" a Gold Medal. [see Exhibits 1-6].

### **Exhibit # 1**

2004 NIH Scientific Achievement Award

Presented to

Dr. Hameed Khan

By

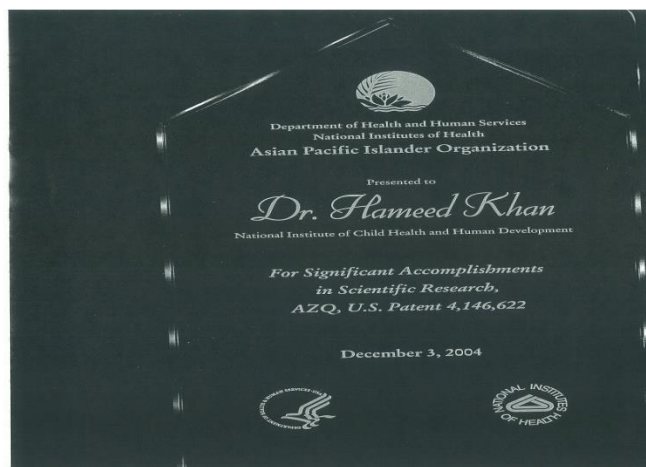
Dr. Elias Zerhouni,  
The Director of NIH

During the NIH/APAO Award Ceremony held on December 3, 2004.



Dr. Khan is the Discoverer of AZQ (US Patent 4,146,622), a Novel Experimental Drug Specifically Designed to shut off a Gene that causes Brain Cancer for which he receives a 17-year Royalty for his invention (License Number L-019-01/0). To this date, more than 300 research papers have been published on AZQ. The award ceremony was broadcast live worldwide by the Voice of America (VOA). Dr. Khan is the first Indian to receive one of America's highest awards in Medicine.

### **Exhibit # 2**



### Exhibit # 3

His Excellency, Dr. A.P.J. Abdul Kalam, The President of India Greeting. Dr. A. Hameed Khan.

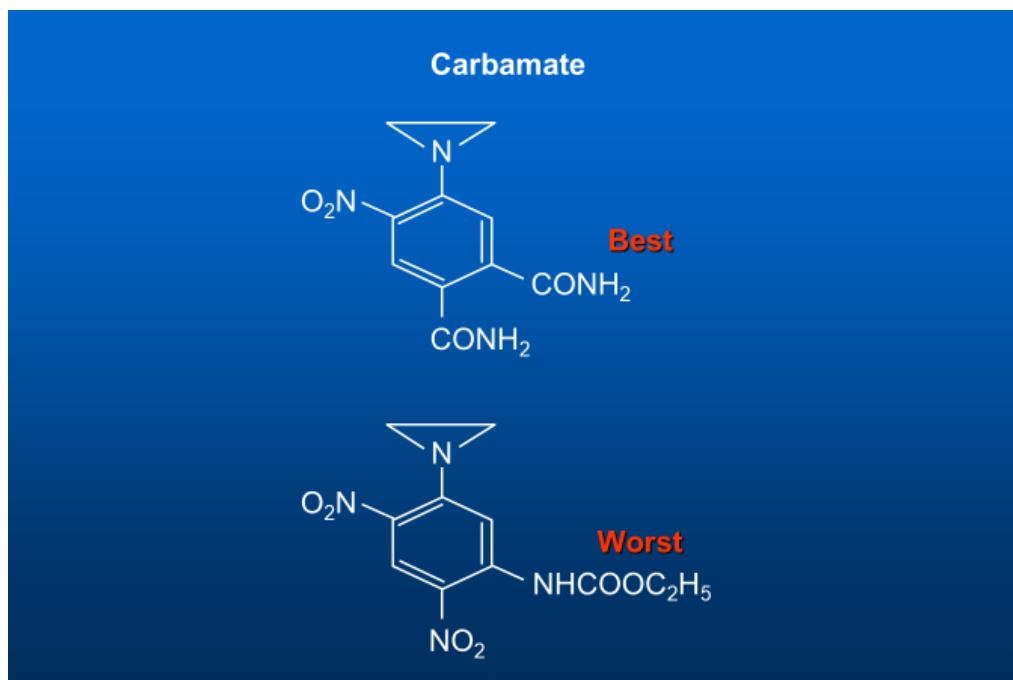


Discoverer of anti-cancer AZQ, after receiving 2004, Vaidya Ratna, The Gold Medal, One Of India's Highest Awards in Medicine. At The Rashtrapathi Bhavan (Presidential Palace), in Delhi, India, during a Reception held on April 2, 2004.

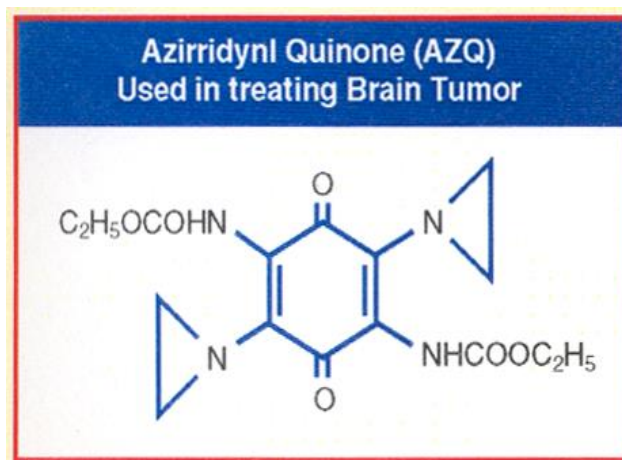
### India's National Medal of Honor

#### Exhibit # 4

Single Strand DNA Binding Aziridine and Carbamate.



### DNA Single Strand Binding Agents



**U.S.**

Structure of AZQ for treating Glioblastoma

**Exhibit # 5**

**Gold Medal for Dr. Khan**



Dr. A. Hameed Khan, a Scientist at the National Institutes of Health (NIH) USA, an American Scientist of Indian Origin was awarded on April 2, 2004. Vaidya Ratna; The gold Medal, one of India's Highest Awards in Medicine for his Discovery of AZQ (US Patent 4,146,622) which is now undergoing Clinical Trials for Treating Bran Cancer. For Discovering AZQ Dr. Khan was Awarded A Gold Medal

## Exhibit # 6

*The Royals of Travancore*



Dr. Hameed Khan of NIH was invited to give the "Maharaja Thramal Memorial Award Lecture" "On the Impact of Genetic Revolution on our lives during 21st Century and Beyond" at the University of Travandrum. After the lecture, His Royal Highness Sree Padmanabha Dasa Marthanda Vama (the brother-in-law) of Her Royal Highness Maharani Travancore (on his left) invited Dr. Hameed Khan and Mrs. Vijayalakshmi Khan for the Tea at the Pattom Palace at Thiruvanthapuram on May 12, 1999. Standing on Dr. Khan's right is the Son-in-law of Her Royal Highness, the Maharani.

While Genome Center at NIH sequence and map the Genomes, my Institute NICHD (National Institute of Child Health and Human Development), provides funds to new investigators to identify Gene Markers that is to identify deleterious mutations responsible for causing various genetic diseases. Identification of the genetic make-up for causing diseases will help us design drugs to shut off those genes. How could we help new investigators to design novel drugs to treat old diseases? The following example will illustrate rational approach to treat Breast Cancer. Future genomic medicine.

## Based on The Genetic Make-up

### What other cancers should be explored next?

Of all cancers, the largest killer of women is the Breast Cancer. Despite the use of highly advanced treatment methods such as Chemotherapy, Radiation therapy and Surgery, within three years, the tumor returns as metastatic cancer and kill the patient. On the rational basis, I propose the following approach to develop novel drug to treat Breast Cancer. Although mutations on BRCA1 gene responsible for causing Breast Cancer located on Chromosome-17 has been identified years ago, so few drugs were designed on rational grounds. Now, we have sequenced Chromosome-17. We found that it is made of 92 million nucleotide bases pairs carrying 1,394 genes. By comparing with the Reference Sequence, we can easily identify which nucleotide on which gene of the Chromosome-17 is responsible for causing Breast Cancer. As I said above, Genomic medicine is a predictive medicine. By MRI (Magnetic Resonance Imaging which takes three-dimensional images) and gene sequencing, we should be able to predict if the abnormal changes in the cellular DNA will lead to Breast Cancer. Without this knowledge, it has been so difficult to design drugs on rational basis to treat Breast Cancer. By the time the Breast Cancer diagnosis is confirmed in a patient, the BRCA1 gene has accumulated more than three thousand mutations. Genotyping of the blood sample would also show the existence of many cells carrying mutated cells responsible for creating secondary deposits. It is also found in some cases when not detected earlier, by the time Breast Cancer is confirmed, metastatic cancer cells have already been spread from Liver Lung on their way to Brain [20].

As a Fogarty International Postdoctoral Fellow at the NCI, I was given the chance to work on any cancer, I was pleased. Since all other organs including Breast and Liver could be removed and replaced by organ transplant except Brain, I thought that protecting Brain is utmost important to save life. For years, I work on the development of AZQ. Once the AZQ was developed to protect the Brain Cancer, I could focus on the Breast and Prostate Cancers. Recent, Radiolabeled studies in mice showed that male hormone Testosterone has great affinity for female organs like Breast, Ovary, and Fallopian tube cells. On the other hand, Estrogen, the female hormone, has great affinity for male Prostate gland. By attaching multiple Aziridine rings and Carbamate ions to both Hormones, I could design novel drugs to attack both the Breast and the Prostate cancers. Now, I found that I could increase its toxicity several folds to abnormal cells by attaching more than four Aziridine and Carbamate moieties to both Male and Female Hormones.

In a Breast tumor, within the start and stop codon, BRCA1 gene has captured over two hundred thousand nucleotide bases. The BRCA1 gene carries about three thousand mutations. These mutations are caused by exposure to radiations, chemical or environmental pollutants, viral infection or genetic inheritance. To attack the mutated nucleotides among the three thousand mutations in BRCA1 gene, we could use male hormone, Testosterone, and bind multiple radio labeled Aziridine and Carbamate ions to attack BRCA1 mutations. By using three dimensional MRI, we could show how many radio-labeled nucleotides were bound to which mutations. Out of seventeen positions available for substitutions on Testosterone ring system. There are only three positions that is [1][3] and [17] are available for substitution on Testosterone ring system [19].

Carl Djerassi [C. Djerassi et al. J. Amer. Chem. Soc. 72. 4534 (1950)] had demonstrated that we could activate additional positions for substitutions on hormone ring system such as the position 9 and 10 by



reacting with Bromo-acetamide which introduce a Bromo on position 10 which could be de-brominated by Collidine to introduce a [9-10] double bond which we could be further brominated to produce [9-10] dibromo compound. These bromo ion could be replaced by additional Aziridines or Carbamate ions. We could increase or decrease the number of Aziridine and Carbamate ions to get maximum benefit by further brominating position 15 and 16 to introduce additional Aziridine and Carbamate moieties.

Similarly, we could use the female hormone Estrogen as a carrier and by attaching multiple Aziridine and Carbamate ions to attack Prostate tumor in Men. Since seventeen positions also available are also available on Estrogen ring as well; again, we could increase or decrease the number of Aziridine and Carbamate ions to get the maximum benefit by using Djerassi' method as we did with Testosterone. The above methods are novel approach to designing drugs to treat Breast and Prostate cancers using genetic make-up of a patient to treat metastatic cancers.

### **Ethical Issues**

Although "The Impact of Sequencing Human Genome on Genetically Inherited Diseases" has provided enormous benefit by comparing egg and sperm with the Reference Sequence before the conception. If parents decide to bring this defected child into this world, scientists could use genetic engineering techniques to keep this child healthy at a very high cost. Somatic cell alteration is acceptable, but germ-line engineering is not permitted at this time, because the changes made in the Egg and Sperm is passed on to the next generations. The other forbidden area includes insertion of oncogenes, or toxins genes in living systems. Currently, no funds are available for such studies. As the cost of next generation Nanopore Sequencer is coming down, we all must sequence our genome which could serve as a part of our medical record. The problem is who should have access to your genome. By looking at your genome sequence, what if you discover that you harbor a potentially deleterious gene in your genome? Who should know your secrets, your employer, your insurance company, your friends, relatives, or your spouse or your fiancé? Before releasing this information to others, you should carefully consider many ethical problems that could emerge with sequencing your genome. By comparing with Reference Sequence, we can easily identify deleterious mutations. The Govt. has passed laws to protect you from your employers and insurance company.

By 2050, Novel Drug Design would have produced new class of medicine to treat all known six thousand mutated genes. We would not only produce new treatment, but also, we would have new food, new fuel, new medicine to treat every disease known to mankind to protect, preserve and prolong human life beyond one hundred years. This section discusses the impact of prolonging human life beyond one hundred years. Our attempt to prolong Human life by shutting off the genes of the old age diseases raises several ethical and moral questions. We face the same population problem when we succeed in shutting off genes of all three old age diseases that is Cancer, Cardiovascular disease, and Alzheimer. In future, because of genetic advances, most people will live longer and happier life. It raises several questions. What happens after we achieve that goal of reaching human age beyond 100 years? What would be the quality of our life? By exercises and good nutrition, if the body mass is not retained, the Centenarians are most likely to be fragile and weak. They need the help of caretakers to perform the daily routine. By 2050, if we increase the age of about a hundred year of about a billion people, we need another billion

caretakers. Will the society be happy with this achievement? I doubt it. The society is hardly likely to accept such a proposal.

To cure diseases to prolong human life, several present and future attempts are described below:

We need to make two rationale approaches: First, to identify rare allele in the Genome of Centenarians responsible for prolonging their lives. Once identified the allele, we need to conduct genetic engineering that is to cut, paste, copy, and splice the allele into the Genome of volunteers to study its function. Second, to design drugs to shut off genes of old age to prolong life. Next attempt to increase human life would be to prevent the loss of Telomeres, the six-letter code (TTAGGG) that shorten our DNA and shorten our lifespan. During replication, each Chromosome loses about 30 Telomeres each year. If we prevent the loss of Telomeres by using the enzyme Telomerase Reverse Transcriptase (TRT), we could slow down the aging process. We have already demonstrated in the worm *C. Elegance* that we could increase its lifespan several fold. Now, we could translate this work in human being; we could try by making a less virulent Flu Virus carrying TRT gene when injected to a volunteer who comes down with a mild Flu. When he recovers from the Flu, the TRT gene would have inserted in the entire genome of every cell in his body (we can confirm the insertion by sequencing). Suppose at each replication only 15 Telomeres are deleted instead of 30 Telomeres. This person is likely to live twice as long. Also suppose the sequencing of his genome would confirm that every cell of his body carries the TRT gene. Since the longevity treatment with the Flu virus is safe, inexpensive and would be easily available to everyone, should we provide the treatment to every man, woman, and child on the face of the Earth?

Such studies are likely to raise two serious ethical questions. First, we must ask ourselves, do people have a right to live and second do we have a right to live if we wish, no matter how old, how weak or how sick we are? The answer to first question is, according to the UN charter, we all have the right to life, liberty, and pursuit of happiness. It is the second question which is troublesome. Do people have a right to extent their lives if they wish? Most people are reluctant to answer this question either No or Yes. Both answers have some support. Those who said No, have a good reason. First, they argue that there are eight billion people live on planet Earth and we are adding 90 million additional people each year. According to UN estimate, by 2050, the population of the world is likely to reach nine billion. Does our planet Earth have all resources to support such a population explosion? Can we provide food, fuel, and medicine to all the people of the world? In poor countries millions are starving now. By extending life, we will have serious problems such as lawlessness, riots, and chaos in the streets. The current population of Earth have polluted the water, polluted the air, and polluted the land. Today, they wonder if the water they drink is safe, the food they eat is safe and the air they breathe is safe. If we continue to pollute the Planet with the current rate which is 110 million ton of pollutant that we release in the atmosphere each year; how much pollutants, we will have accumulated in the atmosphere in ten years or in hundred years.

On the other hand, those who say Yes; we should extend life have good reasons as well. We have only one planet Earth to live. We have no Planet B to save human life. We look up to Heaven to find another home for humanity. To search for a suitable Planet for human life to survive, we need to train an army of Astronauts to travel into deep space with extended life span. They may have to travel for centuries to

find a habitable planet for humans. We do not want them to die on their way to find a new home for humanity. We must continue to search for treatment to prolong human life. The next generation of scientists, my students, will pursue this goal with great zeal and great enthusiasm. [21, -42]. What if we succeed tomorrow in developing treatment of all three old age diseases to double or triple our lifespan? If we don't succeed tomorrow may be the day after tomorrow. Say the treatment is safe, inexpensive, and easily available to every man, woman, and child on the face of the Earth. Who decide that person A will receive the longevity treatment and will live, and person B will not receive the same treatment and therefore will die? We need debate and discussion and come up with guidelines for our society. One person cannot provide answer to all these questions. All I want to do is to raise these questions in your mind. My aim will be fulfilled if I have made you think along these lines. Opinion expressed in this lecture is mine and does not represents NIH policy.

## References

1. Watson JD, Crick FH. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 171(4356) 737-38.
2. Khan H. (2021) The impact of sequencing Human Genome on Genomic Food & Medicine. *Genome*. 9(1):6-19.
3. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. *Nature*. 409(6822): 860-21.
4. Collins FS, Lander ES, Rogers J, Waterson RH. Finishing the euchromatic sequence of the human Genome. *Nature*.
5. <https://pubmed.ncbi.nlm.nih.gov/16341006/>
6. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, et al. (2017) DNA sequencing at 40: past, present and future. *Nature*. 550(7676):345-53.
7. <https://www.rxlist.com/leukeran-drug.htm>
8. Ross WCJ. (1953) The Chemistry of Cytotoxic Alkylating Agents” In *Advances in Cancer Research* by Greenstein, J.P., and Haddow, A., Academic Press, Inc., New York, , 397-449.
9. Ross WCJ. (1962) “Biological Alkylating Agents” Butterworth, London.
10. Ross WCJ. (1949) *Journal of Chemical Society*, 183.
11. Ross WCJ. (1950) *J. Chem. Soc.*, 2257.
12. Ross WCJ. (1964) Mitchley, B.C.V., *Ann. Rep. Brit. Empire Cancer Campn*, 42:70.
13. Facon T, Mary YJ, Hulin C, Benboubker L, Attal M, et al. (2007) *Randomized Controlled Trial*. 370(9594):1209-18.
14. L M Cobb, T A Connors, L A Elson, A H Khan, B C Mitchley, et al. (1969) *Biochem Pharmacol* .
15. Khan AH, Ross WC. (1969) Tumour-growth inhibitory nitrophenylaziridines and related compounds: structure-activity relationships. *Chem Biol Interact*. 1(1):27-47 PART I
16. Khan AH, Ross WC. (1969) Tumour-growth inhibitory nitrophenylaziridines and related compounds: structure-activity relationships. *Chem Biol Interact*. 1(1):27-47 PART II
17. Khan AH, Driscoll JS. (1976) Potential central nervous system antitumor agents. Aziridinylbenzoquinones. *J Med Chem* . 19(2):313-7. PART I
18. Khan AH, Driscoll JS. (1976) Potential central nervous system antitumor agents. Aziridinylbenzoquinones. *J Med Chem* . 19(2):313-7. PART II
19. <https://scivisionpub.com/pdfs/the-impact-of-diagnostic-mri-on-the-early-detection-of-lethal-genes-in-human-genome-and-to-develop-genomic-medicine-to-treat-brain-1697.pdf>

20. Hameed AK. (2021) The Impact of Diagnostic MRI on the Early Detection of Lethal Genes in Human Genome and to Develop Genomic Medicine to Treat Brain Cancers. *J Med Clin Res Rev.* 5(6): 1-9.
21. <https://www.intechopen.com/books/8631>
22. <https://link.springer.com/book/10.1007/978-3-319-48054-1>
23. Hameed AK. (2021) The Impact of Diagnostic MRI on the Early Detection of Lethal Genes in Human Genome and to Develop Genomic Medicine to Treat Brain Cancers. *J Med - Clin Res & Rev.* 5(6):1-9.
24. e-Book, "The Impact of Sequencing Human Genome on Genomic Medicine and the Discovery of AZQ (US Patent 4,146,622) Specifically Designed to shut off genes that cause Brain Cancer." Chapter I, Volume 180, pages 1- 63, 2021: Edited by Leon V. Berhardt, Advances in Medicine and Biology, NOVA Medicine and Health, Publication
25. <https://www.intechopen.com/chapters/73119>
26. Hameed AK. (2021) The Impact of Sequencing Human Genome on Genomic Food & Medicine. *Int J Gen Genomes.* 9(1):6-19.
27. Hameed AK. (2021) The Impact of Sequencing Genomes on The Human Longevity Project. *J Med - Clin Res Rev.* 5(7): 1-12.
28. Hameed AK. (2021) The Impact of Diagnostic MRI on the Early Detection of Lethal Genes in Human Genome and to Develop Genomic Medicine to Treat Brain Cancers. *J Med Clin Res & Rev.* 5(6): 1-9.
29. Khan H. (2021) The Impact of Sequencing Genomes on The Human Longevity Project. *J Med - Clin Res & Rev.* 5(7): 1-12.
30. Hameed AK. The Impact of Sequencing Human Genome on Genomic Food & Medicine. *Int J Genetics Genomics.* 9(1):6-19.
31. Hameed AK. (2022) The Impact of Sequencing Human Genome on Genomic Medicine and the Discovery of AZQ (US Patent 4,146,622) Specifically Designed to Shut off Genes That Cause Brain Cancer. *Int J Genetics Genomics.* 2022:64-78.
32. <https://www.intechopen.com/books/9831>
33. EAS J Biotechnol Genet ISSN: 2663-189X (Print) & ISSN: 2663-7286 (Online) Published By East African Scholars Publisher, Kenya.
34. Mukasa A. (2020) Genome Medicine for Brain Tumors: Current Status and Future Perspectives. *Neurol Med Chir.* 60(11):531-42.
35. *Journal of Cancer Research Reviews & Reports: The Impact of Sequencing Human Genome on Cancer Chemotherapy J Can Res Rev Rep, 2021*
36. *Journal of Cancer Research Reviews & Reports- The Impact of Sequencing Human Genome on Cancer Chemotherapy*
37. Hameed AK. (2021) The Impact of Sequencing Genomes on the Understanding of the Origin of Life on Earth. *Biomed J Sci & Tech Res* 40(2)-2021.
38. Hameed AK (2022). The Impact of Sequencing Human Genome on the Genetically Engineered Life. *J. Cancer Res and Cellular Therap.* 6(1).
39. Hameed AK. (2022). Sequencing the Human Genome to Study the correlation between Genes and Violence, Horizon publishers.
40. Hameed AK (2022) The Impact of Sequencing Human Genome on the 2nd Genesis on Earth, *Journal of Current Trends in Agriculture, Environment and Sustainability*, Katalyst publisher.
41. Hameed AK. (2022) Assessment of Human Genome Sequencing on the New-World Order, *Research Aspects in Biological Sci.* 4:978-93-5547-735-4
42. Hameed AK. (2022) The Impact of Sequencing Human Genome on Novel Drug Delivery Method to Treat Cancers. *Int J Genetics and Genomics.* 10(3):64-78.

43. Hameed AK. (2022) The Impact of Sequencing Human Genome on Aging. *Nov Res Sci.* 12(1). NRS. 000780. 2022.