A History of Genetics and Genomics

Phil McClean
September 2003

Genomics is a recent convergence of many sciences including genetics, molecular biology, biochemistry, statistics and computer sciences. Before scientists even uttered the word genomics, these other fields were richly developed. Of these fields, the history of genetics and molecular biology are particularly relevant to the techniques, experimental designs, and intellectual approaches used in genomics. The development of computers and the internet has provided researchers ready access to the large body of information generated throughout the world. Table 1 is an extensive history of the major developments in these fields. The narrative will try to unify some of these discoveries into major findings.

Mid to Late 19th Century: Evolution, Natural Selection, Particulate Inheritance and Nuclein

Understanding origins is a constant pursuit of man. In the 1858, our understanding of the origin of species and how species variability arose was revolutionized by the research of Darwin and Wallace. They described how new species arose via evolution and how natural selection uses natural variation to evolve new forms. The importance of this discovery was reflected in the now famous quote of Dobzhansky:

"Nothing in biology makes sense except in the light of evolution."
- Theodore Dobzhansky, The American Biology Teacher, March 1973

A few years later, Gregor Mendel, an Austrian monk, summarized his years of research on peas in his famous publication. In that paper, he described the unit of heredity as a particle that does not change. This was in contrast to the prevailing "blending theory of inheritance." Equally important, Mendel formalized the importance of developing pure (genotypically homozygous) lines, keeping careful notes, and statistically analyzing the data. His approach of crossing individuals with variable phenotypes and following them in successive generations is still the only approach utilized to understand the genetic inheritance of a trait. Others in this century were concluding that statistical approaches to biology would help solve problems in biology and inheritance.

Research in the 19th century was often performed in isolation. While Mendel was concluding that inheritance was particulate in nature, others were trying to figure out the physical nature of the particle. Haeckel correctly predicted that the heredity material was located in the nucleus. Miescher showed the material in the nucleus was a nucleic acid. Others observed the behavior of chromosomes and suggested they had a role in heredity. One wonders how concepts might have evolved if information was mobile at that time as it is today.

Early 20th Century: Mendelian Principles are extended and the Chromosomal Theory of Inheritance solidifies

Except for his early adult years, Mendel did not have an active research program. Therefore, his groundbreaking research went largely unnoticed. It was not until 1900 that others, who had performed similar experiments to his, arrived at the same conclusions. Their publications cited his work, leading to a rediscovery of the Mendelian principles. Quickly following the rediscovery, other genetic principles such as linkage, lethal genes, and a bit later, maternal inheritance were described. In each case, the principles provided to be simple extensions of the Mendelian laws, providing further evidence of their importance.

At the beginning of the century, the work on chromosomes coalesced into the *chromosomal theory of inheritance*. This theory focused research on the chromosome as the location of genes. The field of cytogenetics was based on this discovery. The first observations of chromosomal abnormalities (duplications, deletions, translocations, inversions) are reported. Observations such as position effect demonstrate that there is a direct link between chromosome structure and phenotype. All of these discoveries justified research to discover the physical basis of heredity.

Mid 20th Century: DNA is the stuff of life; the preeminence of the Darwinian theory of evolution via natural selection is confirmed

As early as the 1870s, the material in the nucleus was determined to be a nucleic acid. From the 1920s through the mid-1950s, a series of experiments demonstrated that DNA was indeed the genetic material. The transformation experiments of Griffith demonstrated that a factor found in a lethal strain of bacteria could convert a non-lethal strain of the bacteria into a lethal strain. It was the careful experiments of Avery, MacLeod and McCarty that determined DNA, not protein or RNA was the factor responsible for the conversion. This was further confirmed by Hershey and Chase, although their experiments had flaws which prevented them from being definitive. Watson and Crick determined the structure of DNA, and others suggested that DNA contained a genetic code. By the mid 1960s that code was deciphered. Experiments involving the process of transcription and translation led to the development of the “central dogma of molecular biology” concept by Crick.

The experiments of the early 19th century that confirmed that Mendelian principles could be extended to many gene systems became a major component of what was to be called the Modern synthesis (on neo-Darwinism). The experimental demonstration that mutations could be induced was also an important component of the solidification of the concept that natural selection was a major factor in evolution. Finally, the theories embodied in population genetics were also critical. The synthesis states that mutations create variation; recombination develops new forms, the variation extends through the population, and based on environmental constraints the variation is finally acted upon by the forces of natural selection to produce more fit individuals.

Mid-late 20th Century and the Early Days of the 21st Century: The Age of Molecular Genetics; Phylogenetics Studies Intensive; The Information Age; The Emergence of Genomics Science

The discoveries of the mid to late 20th century defined processes that would provide the tools for molecular biology, recombinant DNA technology, and finally the biotechnology industry. The elucidation of the process of DNA replication described the necessary components needed for the widely-used chain termination DNA sequencing procedure. Understanding replication helped determine those tools necessary for the radiolabelling of DNA. The development was necessary to support Southern hybridizations and the early molecular mapping experiments. Understanding replication also defined the role of the ligase enzyme that is so critical for DNA cloning. Restriction enzymes were discovered and used to construct recombinant DNA molecules that contained foreign DNA that could be grown in abundance in bacterial cells. The discovery of reverse transcriptase also enabled cDNA cloning which is essential for the modern EST projects. Cloning is essential for the discovery of gene structure and function. It is also an essential step for all of the genome sequencing projects. The importance of the PCR procedure cannot be emphasized enough.

The advent of protein and DNA sequencing launched a new era of phylogenetics. Species could now be compared at the molecular level. New procedures for the development of phylogenies are developed. The neutral theory of molecular evolution is proposed. This is a direct attack on preeminence of selection as the driving force of evolution. The theory suggests that most mutations are neutral and are fixed by genetic drift and not selection. It is debated whether the evolution of species is driven more by neutral effects or selection. Some feel the two theories are compatible and exert their effects on different genes.

The information age is essential to genomics. The electronic analysis, distribution and storage of genomic data is a hallmark of the science. Critical to this was the development of computers, both large and small, which put computing power in the hands of all scientists. The free distribution of analytical software provided scientists with the tools to study the details of their experiments. The internet spawned the distribution of information from central databases. E-mail connected scientists and fostered the rapid exchange of ideas. The advent of the WWW provided a new medium for the presentation of information.

Whole genome are sequenced for the first time. For other species, the gene content is described using ESTs. Microarray analyses provided the first glimpse of global expression patterns. Proteomics begins to describe the protein component of the genome. Metabolomics is established.

Table 1. An annotated history of genetics and genomics.

Year	Who	Discovery
1858	Charles Darwin Alfred Russell Wallace	These scientists jointly announced ***the theory of evolution via natural selection***. Darwin, an upper class Englishman, had lectured about the topic, but it took the insistence of Wallace, a commoner who independently realized the same concept, for him to publicly state the theory.
1859	Charles Darwin	Publication of ***"The Origins of Species"***, a treatise that formally outlined the theory of evolution via natural selection
1865	Gregor Mendel	The concept of ***particulate (gene) inheritance*** was established. The laws of segregation and independent assortment were demonstrated. The publication is entitled "Experiments in Plant Hybridization" and outlines the famous "pea experiments."
1866	Ernst Haeckel	Proposes the idea that the hereditary material resides in the nucleus.
1871	Friedrich Miescher	The term ***nuclein*** is used for the material found inside the nucleus of a cell. Further experiments (1874) revealed nuclein consisted of a nucleic acid and protein.
1871	Lambert Adolphe Jacques Quetelet	It is shown that statistical analysis can provide important insights in biology. This concept was critical to the development of the field of biometry, the application of statistics to biological phenomenon.
1879-1892	Walther Flemming Eduard Strasburger Edouard van Beneden	The first accurate counting of chromosomes are made. Cell division is observed. Terms ***chromatin***, ***mitosis***, ***cytoplasm***, ***nucleoplasm***, ***prophase***, and ***metaphase*** are coined.
1887	August Weismann	A universal ***theory of chromosome behavior*** is proposed that predicts meiosis in sex cells. Edouard van Beneden confirmed this theory in the same year.
1888	Henrich Wilhelm Gottfried Waldeyer	The term ***chromosome*** is applied to the condensed version of material found in the nucleus.
1889	Francis Galton	The book ***"Natural Inheritance"*** is published. Variation was studied by quantitatively measuring difference among traits. The field of biometry is formally founded.

1894	William Bateson	In the book “Materials for the Study of Variation”, the concept of discontinuous variation is discussed, an important tenet found in Mendel’s work.
1894	Karl Pearson	Methods for the ***analysis of statistical frequency distributions*** developed. These were necessary for the later development of mathematical models of evolution.
1899	William Bateson	The use of hybridization between two individuals is described as a tool of the scientific analysis of heredity. This again was discovered to be an important tenet of Mendel’s work.
1900	Carl Correns Hugo de Vries Erich von Tschermak	***Mendel’s work is rediscovered*** independently. de Vries and Correns were experiments similar to those of Mendel and arrived at similar results. Once they read Mendel’s paper, they recognized its preeminence and made the world aware of it.
1900	Hugo de Vries	The term ***mutation*** is used to describe the apparently spontaneous appearance of new traits in evening primrose (*Oenothera*).
1902	C.E. McClung	The concept that specific chromosomes are responsible for determining sex in a number of animals is presented.
1902	Walter Sutton Theodor Boveri	Within a specific species, each chromosome is described as having unique physical characteristics. It is shown that chromosomes occur in pairs, one parent contributes each member of the pair, and the pairs separate during meiosis. Sutton suggests chromosomes are a physical manifestation on which the unit of heredity resides. This came to be known as the ***chromosomal theory of inheritance***.
1902	Archibald Garrod	The first human disease is described that exhibits Mendelian inheritance. The disease is alkaptonuria. Later (1909) Garrod is the first to discuss the biochemical genetics of man.
1902	William Bateson	The terms ***genetics***, ***homozygote***, ***heterozygote***, ***epistasis***, F_1, F_2, and ***allelomorph*** (shortened later to allele) were first used.
1903	Wilhelm Johannsen	The important concepts of phenotype, genotype, and selection were elucidated. The terms were actually coined later (1909).
1905	William Bateson R.C. Punnett	Experiments performed on sweet pea demonstrated the concept of ***linkage***.

1905	Lucien Claude Cuenot	***Lethal genes*** were discovered in the classic experiment involving a cross between two yellow mice. The population segregated in the following ratio: 2 yellow fur mice to 1 agouti fur mouse.
1907	Friedrick Laibach	The first suggestion of ***Arabidopsis as a model organism***.
1908	G.H. Hardy W. Weinberg	Independently, the ***Hardy-Weinberg principle of genetic equilibrium*** is formulated. This is the unifying theory that underlies population genetics.
1909	G.H. Shull	The use of self-fertilized corn to produce commercial seed is proposed. This is a direct application of the scientific approach utilized by Mendel.
1909	Wilhelm Johannsen	While studying seed size in common bean, it becomes apparent that the outward appearance of the individual and its genetic makeup must be clearly distinguished. The terms ***phenotype*** and ***genotype*** and coined for this purpose. The term ***gene*** is also used for the first time.
1909	H. Nilsson-Ehle	The quantitative variation is seedcoat color in wheat and oat was explained by the interaction of multiple genetic factors. This is the birth of ***quantitative genetics***.
1910	Thomas Hunt Morgan	The first demonstration of ***sex linkage*** in *Drosophila* is published. This suggested genes reside on chromosomes. The era of ***fruit fly as a model organism*** begins.
1913	Alfred Sturtevant	The first ***genetic map*** is developed in *Drosophila*. This solidifies the concept of linkage in this organism.
1914	Calvin Bridges	The observation of ***non-disjunction*** in sex chromosomes provides the tool necessary to definitively provie the chromosome theory of inheritance
1917- 1923	Calvin Bridges	The first observations of ***deficiencies*** (1917), ***duplications*** (1919), and ***translatocations*** (1923) are observed in *Drosophila* chromosomes.
1915- 1917	Frederick Twort Felix D'Herelle	***Bacteriophage***, a virus that attacks bacteria were first described.
1919	Thomas Hunt Morgan	It is shown the ***number of chromosomes equals the number of linkage groups***.
1923	A.E. Boycott C. Diver	***Maternal inheritance*** is demonstrated for shell coiling direction in snail, *Limnea peregra*. Sturtevant suggests the phenomenon is the result of the phenotype of the female ooplasm, and that the phenotype is controlled by the genotype of the mother.

1925	Alfred Strutevant	The *Drosophila* Bar-eye effect is described. This is the first demonstration of ***position effect***.
1926	Alfred Strutevant	The first ***inversion*** is observed in *Drosophila*.
1927	J.B.S. Haldane	The concept of ***homologs*** is presented. This is based on observations of coat color in rodents and carnivores.
1927	B.O. Dodge	***Neurospora*** first used as a genetic organism.
1927	H.J. Muller	***X-rays induce mutations*** in *Drosophila*.
1928	L.J. Stadler	Corn mutations are induced by x-rays, and the relationship between the number of mutations and the ***x-ray dosage effect*** is established.
1928	F. Griffith	***Transformation*** of *Pneumococci* is obtained. This is the critical experiment that leads to the eventual discovery that DNA was the genetic material.
1930	R.A. Fisher	The first formal analysis of selection is provided in the book ***"The Genetical Theory of Natural Selection."***
1930	Arne Tiselius	***Electrophoresis*** is developed as a technique to separate proteins.
1931	Harriet Creighton Barbara McClintock	Using corn as their experimental organism, it was demonstrated that crossing over between two homologous chromosomes involves the physical exchange of genetic material between the two chromosomes. C. Stern also demonstrated this concept using *Drosophila*. The fact that corn has a longer life cycle demonstrated that McClinktock first developed the idea.
1932-1953	R.A. Fisher Theodore Dobzhansky	The ***Modern Synthesis*** is formulated. This couples the laws of Mendelian inheritance and a knowledge of mutation with the Darwinian theory of evolution via natural selection.
1933	T.S. Painter	***Salivary gland*** chromosomes of *Drosophila* first used in cytogenetic studies. These became critical tools to exactly locate genes responsible for phenotypes.
1935	G.W Beadle and B. Ephrussi; and A. Kuhn and A. Butenandt	The biochemical genetics controlling eye-pigment synthesis in *Drosophila* (Beadle and Ephrussi) and *Ephestia* (Kuhn and Butenandt) is described.
1936	Calvin Bridges	The ***first cytogentic map*** of the *Drosophila* salivary gland chromosome is published.
1936	Alfred Sturtevant Theodore Dobzhansky	The use of inversions to describe the phylogeny of a chromosome is described.

1937	Theodore Dobzhansky	***"Genetics and the Origin of Species"***, a seminal book in the field of evolutionary genetics, is published.
1937	Arne Tiselius	The use of ***electrophoresis*** to separate proteins was demonstrated.
1939	E.L. Ellis Max Delbrück	***Phage genetics*** research begins with experiments that showed a virus is taken up in a single step, undergoes a latent period, and then burst the bacterial cell.
1941	George Beadle E. L. Tatum	The ***one gene, one enzyme*** concept is developed based on biochemical studies of *Neurospora*.
1941	K. Mather	Following experiments that demonstrate multiple genes control traits in several organisms, the term ***polygenes*** is coined.
1943	S.E. Luria Max Delbrück	Experiments are reported that demonstrate that bacteria are capable of undergoing ***spontaneous mutations***. This begins the field of ***bacterial genetics***.
1944	Oswald T. Avery Colin M. MacLeod Maclyn McCarty	Extending the experiments of Griffith (1929), it is first shown that ***DNA is the genetic material***. This fact is often lost, and this discovery is often afforded to Hershey and Chase (1953).
1945	S.E. Luria	It is demonstrated that mutations can occur in bacterial viruses.
1945	John von Neumann	The concept of a ***computer*** is developed.
1946	J. Lederberg E.L. Tatum	***Bacterial genetic recombination*** is demonstrated. This involves movement of DNA from one bacteria to another
1946	Max Delbrück and W.T. Bailey A.D. Hershey	***Bacteriophage genetic recombination*** is demonstrated.
1948	H.J. Muller	The phenomenom of ***dosage compensation*** is described.
1948	J. Lederberg and N. Zinder B.D. Davis	The technique of using penicillin selection to discover ***biochemical mutants*** of bacteria is described. The two groups make this discovery independently.
1948	Barbara McClintock	Following an intense study of the genes controlling seed coat color in corn, the concept of ***transposable elements*** is proposed to explain the spontaneous induction of mutations.
1949	J.V. Neel	***Sickle-cell anemia*** is inherited as single gene, recessive Mendelian trait.

1950	Erwin Chargaff	***Adenine=thymine and guanine=cytosine***. It is demonstrated that within all DNA molecules, the number of adenines equals the number of thymine, and the number of guanine equals the number of cytosine.
1952	F. Sanger Colleagues	The complete amino acid ***sequence of insulin*** is determined, and it is shown that the molecule is a dimer held together by disulphide bonds.
1952	A.D. Hershey M. Chase	The classic ***"blender experiment"*** is reported that shows that phage DNA enters (along with a little protein) and leads to the eventual rupture of the cell. This is often, and mistakenly, considered the definitive experiment proving that DNA is the genetic material.
1953	James Watson Francis Crick	A ***structural model of DNA*** is presented that states it consists of two anti-parallel chains held together by hydrogen bonds. The model suggests a model of DNA replication.
1954	George Gamow	It is suggested that ***DNA contains a code*** that is responsible for the production of proteins. This concept supported later research that discovered the actual code.
1955	Seymor Benzer	The fine structure of the *rII* region of the T4 bacteriophage of *E. coli* is worked out. This will eventually lead to the concept that the nucleotide is the unit of mutation. The term ***cistron*** is coined.
1955	Severo Ochoa	An enzyme, ***RNA polymerase***, is described that can synthesize RNA. This is a key discovery relating to the mechanism of ***transcription***.
1955	Newton Morton	***LOD score*** method of determining linkage distance in humans developed.
1956	Arthur Kornberg	***DNA polymerase I*** is purified from *E. coli*. This enzyme is shown to be a component of DNA ***replication***. The understanding of the process was critical for the discovery of key enzymes and mechanisms critical to recombinant DNA technology.
1956	F. Jacob E.L. Woolman	***Bacterial conjugation*** is shown to involve the physical exchange of DNA between two bacterial strains.
1957	Francis Crick	The ***central dogma of molecular biology*** is proposed. This is a first elucidation of the link between the sequences in the DNA molecule and the production of proteins.

1957	V.M. Ingram	The normal and mutant sickle cell hemoglobin molecules are distinguished by a mutation in a single amino acid.
1958	F. Jacob E. Woolman	It is demonstrated that the *E. coli* chromosome is circular, and the F factor disturb the linear order of gene by its insertion into the chromosome.
1958	Francis Crick	Crick predicts the existence of ***tRNA*** when he suggests that amino acids are brought to a template mRNA by a nucleic acid ***adaptor molecule*** (tRNA), and the molecule actually fits on the mRNA.
1958	Matthew Meselsohn F.W. Stahl	***Semiconservative DNA replication*** is proven by the use of density equilibrium centrifugation.
1959	R.L. Sinsheimer	The bacteriophage ΦX174 of *E. coli* is shown to contain a single-stranded DNA molecule.
1961	F. Jacob J. Monod	The publication of “Genetic Regulatory Mechanisms in the Synthesis of Proteins” describes the ***Lac Operon*** controlling network in *E. coli*.
1961	Marshall Nirenberg	The concept that each amino acid corresponds to a ***triplet code*** was developed. The first correspondence was found between the triplet AAA and the amino acid phenylalanine.
1961	Sydney Brenner Francois Jacob Matthew Meselsohn	***Ribosomes*** are described as the site of protein synthesis. It is also proven that mRNA exists and binds ribosomes.
1965	Margaret Dayhoff	The first ***“Atlas of Protein Sequence and Structure”*** is published. It contains the sequenceof 65 proteins and is considered the first publication in ***bioinformatics***.
1965	Emile Zuckerandl Linus. Pauling	The concept of a ***molecular clock*** is introduced. The theory suggests that the rate of amino acid substitutions are linear over time.
1966	Marshall Nirenberg H. Gobind Khorana	The ***genetic code*** that correlates the triplet code in the mRNA with a specific amino acid is completed.
1967	Waclaw Szybalski W. Summers	It is shown, that of the two strands that make up the DNA molecule only one is used during transcription.
1967	W.M Fitch E. Margoliash	The first ***phylogentic tree*** was developed. It consisted of a comparison of the amino acid sequence of cytochrome C from twenty species that ranged from fungi to mammals.

1968	M. Kimura	The ***neutral theory of molecular evolution*** is presented. It suggests that genetic drift is responsible for evolution. To some, this diminishes the importance of natural selection in evolution. The theory is later synthesized in the 1983 publication ***"The Neutral Theory of Molecular Evolution"***.
1969	Arpanet	The ***Arpanet*** is developed that links computers at four universities. This is thought to be the beginning of the internet.
1969	Jonathan Beckwith	The first bacterial gene is isolated.
1969	Bell Laboratories	The UNIX operating system is created. It has evolved into many flavors, but is the most robust OS for everything from database work to WWW page delivery.
1970	Werner Arber Hamilton Smith	The first ***restriction enzyme***, *Hin*dII is isolated and shown to cut DNA at a specific sequence.
1970	David Baltimore Howard Temin	The enzyme ***reverse transcriptase***, that can make a DNA copy of RNA was discovered independtly. The enzyme is an essential element of ***cDNA cloning***.
1970	U.K. Laemmli	The use of a stacking gel and SDS in electrophoresis is introduced. This innovation made protein characterization much easier.
1971	Lynn Margulis	The ***endosymbiotic theory*** is proposed. The theory suggests that organelles arouse when a pre-eukaryotic cell took up energy transducing bacteria.
1971	Ray Tomlinson	***E-mail***, the first killer application for the internet, is created.
1972	Paul Berg	The ***first recombinant DNA molecule*** was created by splicing together bacterial and viral DNA. This was described as a general approach of mixing together two different DNA molecules.
1973	Herbert Boyer Stanley Cohen	***Recombinant DNA transformation of E. coli*** is achieved when a kanamycin gene is inserted into a plasmid, and the gene functions properly. When genes from toad were later inserted and replicated via the same bacterial plasmid, it is demonstrated that any gene can be cloned in this manner.
1973	Brookhaven National Laboratory	The Brookhaven ***Protein Database*** is created.

1973	Robert Metcalfe	The principles that support ***ethernet*** are described. This will eventually provided the ability of fast delivery of data over fiber optic cable.
1974	Vint Cerf Robert Kahn	The concept of the ***internet*** is developed. The ***TCP*** protocol that supports that transfer of data packets is developed.
1974	Charles Goldfarb	The ***Standardized General Markup Language*** (SGML) is created. The hypertext markup language (*HTML*), the language used to create WWW pages, is a subset of SGML.
1974	Goodman et al.	The ***parsimony procedure*** for developing phylogentic trees is described.
1975	P. O'Farrell	***Two-dimensional electrophoresis*** is developed. This will become an important component of the field of proteomics.
1975	Erwin Southern	A technique is described to transfer and immobilize DNA on a filter memberane and subsequently hybridize a radiolabelled probe to the membrane. ***Southern hybridizations*** quickly become a key element of molecular biology research and were a key component of genetic mapping using RFLPs.
1975	King Wilson	It is suggested for the first time that the sequences of the human and chimpanzee are very similar.
1977	Frederick Sanger Walter Gilbert	The chain-termination and chemical methods of ***DNA sequencing*** are developed. The chain-termination (or Sanger techniqu)e was less onerous and became widely adapted.
1977	Roger Staden	The ***Staden Programs*** that are used to analyze DNA sequence data begin to be developed.
1977	Phillip Sharp Rich Roberts	Independently, it is shown that a mammalian viral gene is interrupted by DNA sequences not found in the mature mRNA. These sequences are called ***introns*** by Walter Gilbert in 1978. It was soon shown that introns are common feature of eukaryotic genes.
1978	David Botstein	The ***RFLP*** (restriction fragment length polymorphism) technique is described and launches a new generation of genetic mapping. The technique launches ***map-based cloning*** of genes.

1979	Goodman et al.	The ***parsimony procedure*** for developing phylogenetic trees is first described
1980	Sanger Group	The ***first complete genome sequence*** is published. The genome was of the bacteriophage ΦX174 of *E. coli*
1980	K. Wuthrich	Protein structure is determined using ***multi-dimensional NMR***.
1981	Felsenstein	The ***maximum likekihood procedure*** to develop phylogenetic trees is published.
1982	GenBank	***GenBank established*** to be a database of all DNA sequences. Initially housed at the Los Alamos National Laboratory, it was transferred to National Center for Biotechnology Information (NCBI) in 1988.
1983	Kary Mullis	The procedure to amplify large amounts of DNA is described. The procedure was later improved by the use of a DNA polymerase from the *Thermus aquaticus* bacterial. The polymerase chain reaction or ***PCR procedure*** will prove one of the most widely applied procedures in all of molecular biology.
1986	Leroy Hood Lloyd Smth Michael Hunkapiller Tim Hunkapiller	The ***first automated DNA sequencer*** is released. This development, based on fluorescent labeling of nucleotides and the Sanger sequencing technique, made it possible to sequence genomes in a reasonable amount of time. Applied Biosystems released the first sequencer.
1986	SWISS-PROT	The initial version of the ***SWISS-PROT*** protein database is released.
1987	Eric Lander et al.	***MAPMAKER***, a computer program for developing genetic linkage maps from molecular marker data is released.
1988	H.D. Higgins P.M. Sharp	The ***Clustal multiple sequence alignment*** approach is published.
1989	Jean Thierr-Mieg Richard Durbin	The ***ACeDB (A C. elegans DataBase)*** is developed. This is the first database developed for genomic information.
1990	US Government	The 15-year ***Human Genome Program*** is launched. The goal is to "find all the genes on every chromosome in the body and to determine their biochemical nature."
1990	Altshul et al.	***BLAST***, a computational approach to aligning two DNA sequences is described. An updated version is released in 1997.

1990	Tim Berners-Lee	Version 1.0 of ***HTML*** is published. This language is used to delivery WWW pages.
1991	Craig Venter	The first ***EST (expressed sequence tagged) sequences*** are published. These short sequences prove to be important for gene discovery.
1993-4	Genthon	The first ***high density linkage map*** of the human genome is published.
1995	Celera Corp Craig Venter	The entire ***H. influenza genome sequence (1.8 Mbp)*** is published. This bacterium is the first free-living organism that is sequenced. The result demonstrated the feasibility of the ***shotgun genome sequencing*** approach.
1995	Pat Brown Ron Davis	The first ***microarray*** system is described. This system uses cDNA sequences printed on glass slides.
1996	Yeast Genome Consortium	The entire ***S. cerevisiae (yeast) genome sequence (12.1 Mbp)*** is published. This is the first eukaryotic genome to be sequenced. The genome is reported to contain 6,250 genes, later (2003) revised to 5,700.
1997	deRisi et al.	The first study of regulatory pathway using microarray technology is published.
1997	*E. coli* genome project	***E. coli (4.7 Mbp) is sequenced*** and shown to contain 4,500 genes.
1998	*C. elegans* genome project	The ***C. elegans genome sequence (97 Mbp)*** is reported. This is the first genome of a multicellular organism to be sequenced. The genome contains 19,100 genes.
1998	Eisen et al.	Cluster analysis approaches are described for the analysis of microarray data.
1999	Affymetrix	The ***oligonucleotide microarray*** system is developed.
2000	*Arabidopsis* Genome Initiative	The ***Arabidopsis thaliana genome (125 Mbp) is sequenced***. This is the first publishing of a flowering plant genome. The genome is shown to contain 25,500 genes.
2000	*Drosophila* Genome Initiative	The ***Drosophila genome sequence (123 Mbp)*** reveals it contains 13,400 genes. This is the first shotgun approach to the sequencing of a eukaryotic organism.
2000	Steven Tanksley	The first gene controlling a quantitative trait is cloned.

2001	International Human Genome Sequencing Consortium Celera Corp	The ***human genome sequence (2900 Mbp)*** is published. It is estimated that the genome contains between 35,000 and 40,000 genes. Later (2002) estimates place the number at 30,000 genes.
2001	Sanger Institute Wellcome Trust	The ***Ensembl Genome Browser*** provides easy access to human genome data.
2002	Mouse Genome Sequencing Consortium	The ***mouse genome sequence (2500 Mbp)*** is published. The number of genes is estimated to be 30,000.
2002	Syngenta, Torrey Pines Research Institute and Beijing Genomics Institute	The draft of the ***rice genome sequence (470 Mbp)*** is published. The estimates of the number of genes range from 32,000 – 56,000. The sequences of the *indica* and *japonica* strains were published jointly.
2002	Puffer-fish Genome Sequencing Consortium	The ***puffer-fish (Fugu rubripes) genome (365 Mbp)*** is sequenced. A total of 31,059 genes are observed. This is the smallest animal gene by size but contains a similar number of genes as human and mouse.
2002	Mosquito Sequencing Consortium	The ***malaria-parasite-carrying mosquito genome sequence*** (278 Mbp) is published. It is shown to contain 13,600 genes, similar to the number found in *Drosophila*
2002	Malaria Sequencing Consortium	The sequence of ***Plasmodium falciparum geneome (23 Mbp)*** is published. This is the parasite that causes malaria. The genome consists of 5300 genes.
2003	British Columbia Cancer Agency	The ***SARS-associated coronavirus genome sequence (30 Kbp)*** is released. The genome contains 16 open-reading frames. The sequence is released less than five months after the disease began spreading world-wide.

A History of Genetics and Genomics

Mid to Late 19th Century: Evolution, Natural Selection, Particulate Inheritance and Nuclein

1858
- Darwin and Wallace
- Role of natural variation and natural selection in evolution

"Nothing in biology makes sense except in the light of evolution."
- Theodore Dobzhansky, The American Biology Teacher, March 1973

1865
- Gregor Mendel
- Particulate inheritance

1866
- Ernst Haeckel
- Heredity materials was in the nucleus

1871
- Friedrich Miescher
- Material in the nucleus was a nucleic acid

Early 20th Century: Mendelian Principles are extended and the Chromosomal Theory of Inheritance solidifies

1900
- Correns, de Vries, von Tschermak
- Mendel’s work is rediscovered
- The age of genetics begins

1902
- Walter Sutton and Theodor Boveri
- Chromosomal Theory of Inheritance
- The heredity material resides in chromosomes

1905-1923
- Linkage
- Sex linkage
- Genetic mapping
- Number of linkage groups = number of chromosomes
- Lethal genes
- Maternal inheritance

1908
- Hardy and Weinberg
- Hardy-Weinberg principle of genetic equilibrium

1909
- Nilsson-Ehle
- Theory of quantitative traits and quantitative genetics

Mid 20th Century: DNA is the stuff of life; the preeminence of the Darwinian theory of evolution via natural selection is confirmed

1928

- Griffith
- Transformation experiments

1944

- Avery, MacLeod, McCarty
- Definitive proof that DNA is the genetic material

1953

- Watson and Crick
- DNA structure is defined

1954-1961

- DNA code is determined
- Transcription is described
- Replication is described
- Translation is described
- Operons are discovered

1932-1953

- Fisher and Dobzhansky
- The Modern Synthesis is formulated
- Links Darwinian evolutionary theory and Mendelian genetics

1968

- Kimura
- Neutral Theory of Molecular Evolution is introduced

Mid-late 20th Century and the Early Days of the 21st Century: The Age of Molecular Genetics; Phylogenetics Studies Intensive; The Information Age; The Emergence of Genomics Science

1969

- ARPANET
- Internet comes on line

1970

- Arber and Smith
- First restriction enzyme, HindII, is isolated

1970

- Baltimore and Temin
- Discovery of reverse transcriptase

1972

- Berg
- First recombinant DNA molecule is constructed

1973

- Boyer and Cohen
- First functional recombinant E. coli cell produced

1977

- Sanger and Gilbert
- DNA sequencing techniques are described

1977

- Sharp and Roberts
- Introns discovered

1978

- Botstein
- RFLPs launch the era of molecular mapping of linkage groups

1980

- Sanger Group
- First genome is sequenced, the bacteriophage ΦX174 of *E. coli*

1983

- Mullis
- PCR technique is discovered

1986

- Hood, Smith, Hunkapiller and Hunkapiller
- First automated DNA sequencer

1990

- US Government
- Human Genome Project launched

1990

- Berners-Lee
- HTML language is created

1995

- Celera
- First bacterial genome (H. influenza) is sequenced

1996

- Yeast Genome Consortium
- First eukaryotic genome (yeast) sequenced

2000

- Arabidopsis Genome Initiative
- First flowering plant genome (*Arabidopsis thaliana*) is sequenced

2001

- International Human Genome Sequencing Consortium
- Celera
- The human genome sequence is published