THE SYNTHETIC THEORY OF EVOLUTION

DNA makes RNA, RNA makes protein and proteins make us. —Crick.[1]

f14 Genes, alleles < deoxyribonucleic acid >

> ...using the metaphor of computation to describe the role of genes is bad biology, for it implies an internal self-sufficiency of DNA. Even if we had the complete DNA sequence of an organism and unlimited computational power, we could not compute an organism, because an organism does not compute itself from its genes. (In fact, any computer that did as poor a job of computation as a genetic "program" does of an organism would immediately be thrown into the trash.)
>
> —Richard Lewontin *The Triple Helix: Gene, Organism, and Environment*, 2000.[2]

Mathematical population theorists Ronald Aylmer Fisher (*The Genetical Theory of Natural Selection*, 1930),[3] John Burdon Sanderson Haldane (design disparaging: "[He has] an inordinate fondness for beetles."),[4] and Sewall Wright (conceiver of "adaptive landscapes")[5] were the principle forgers, during the period 1920-1950, of the modern theory of population genetics that unites previously warring concepts (***Footnote f14.1***) of Darwinian natural selection, which is decidedly nonrandom, that guides evolution incrementally (by gradual change) but rapidly by its recursiveness (the output of previous selection is input for subsequent selection), along paths to fitness, and the purely random nature (no inherent directionality, no predictability) of heredity by large sudden changes (saltations, hopeful monsters) that Mendelian rules suggested to some geneticists (principally, typologists William Bateson, Hugo de Vries, and Wilhelm Johannsen). However, a quiet majority of geneticists thought of species' origin as by the gradual accumulation of small mutations and in 1940 this allowed for "The Synthetic Theory of Evolution" found for by geneticist Theodosius Dobzhansky (1900-1975) and zoologists. This congenial consensus or *Synthesis*,[6] which neologism Julian Sorell Huxley (zoologist and Thomas Henry Huxley's grandson) coined (along with *cline*, *clade*, and *ethnology*), was forwarded and developed notably by him, Ernst Mayr (ornithologist), Bernhard Rensch (zoologist), George Gaylord Simpson (vertebrate paleontologist), and **George Ledyard Stebbins** (botanist).

George Ledyard Stebbins (1906-2000)[13]

In 1950, Stebbins made botany a part of the evolutionary synthesis through his monumental book *Variation and Evolution in Plants*[14] which brought plant evolution into line with animal evolution as in Dobzhansky's 1937 *Genetics and the Origin of Species*.[15]

Resolved was the conflict over the object of selection that since the 1920s had been the gene for geneticists, but for most naturalists it was the individual. Elliot Sober showed the way[7] Ernst Mayr recalls:[8] "He pointed out that one must discriminate between selection of an object and selection for an object. The answer to the question of what is being selected for specifies the particular properties for which a given object of selection is favored. However, a particular gene can favor an individual without being the object of selection because it gives properties to the individual that favor its selection. It is a selection for these properties." [9]

Modern paleobiology references are: *Encyclopedia of Paleontology*, Vols. I & II, edited by Ronald Singer [10] and *Palaeobiology: A synthesis*, by Derek Briggs & Peter Crowther.[11] "Evolutionary process" popularizers are: Richard Dawkins, Jared Diamond, Steven Jay Gould, Steve Jones, Steven Pinker, Mark Ridley, Matt Ridley who delights to explain the evolved, relative, testicle size of promiscuous chimpanzees (huge), harem-keeping gorillas (tiny), and humans (in between), George Christopher Williams, and Carl Zimmer.[12]

Gregor Mendel (1823-1884) [21]

"In 1859 I obtained a very fertile descendant with large, tasty seeds from a first generation hybrid. Since in the following year, its progeny retained the desirable characteristics and were uniform, the variety was cultivated in our vegetable garden, and many plants were raised every year up to 1865." (Gregor Mendel to Carl Nägeli, April 1867)

To learn the rules of inheritance, Mendel kept a count of "differentiating characters" (*differierende Merkmale*) (stem length, pod shape, color and texture of seeds, flower color) as these showed up in successive generations of cross-pollinated gardens peas.[22]

A gene is an organic self-replicating entity that along with others makes up the genome (the regulatory and developmental program of an organism). All genes in an individual's genome are not expressed in the phenome (the niche-interacting biochemical, physiological, and morphological parts of an organism) but all, if the program remains successful, have a chance of being reproduced. This holds for asexual or sexual reproduction. Genetic research has shown that the genes are segments of DNA (deoxyribonucleic acid) molecules.

Heredity traits are passed down through generations by genes. The word "gene" was coined in 1909 by the Danish botanist Wilhelm Johanssen for the heredity unit located on a chromosome (a DNA molecule). Trouble is with the word "unit" in this definition for as Nathaniel C. Comfort pointed out in 2001,[16] what DNA expresses is not a discrete chromosome part with fixed boundaries but, by modern definition, are structural genes (each of which can code for an active protein) and regulator genes (each of which can code for a repressor protein which turns on or off a structural gene in response to signals from the cell) that often act in clusters. Also, DNA segments that constitute a single gene are sometimes far apart on a chromosome. The same DNA segments can be combined in different ways to yield a variety of expressions. Examples exist of genes within genes, overlapping genes, and genes that can be read forward or backwards to be differently expressed. Cutting to the chase, "a gene," Jerry A Coyne offers, is a "piece of DNA that is translated into messenger RNA."[17]

Gene mutation (random) can give rise to the appearance of new characteristics that are not already present in a species' population. John Maynard Smith described this in *Shaping Life*, 1999, and provided reductionist and holistic examples of how genome information is used during the development of the individual (the phenotype).[18] Natural *nonrandom* selection explains how information is incorporated in the genome. This is the mechanism that eluded Darwin and completes his theory of evolution. It together with *random* genetic drift due to neutral (non-selective) mechanisms (such as genetic recombination) gives us *neo-Darwinism*.[19] GOD (generation of diversity) is by *unstoppable* genetic mutations.

Gregor Mendel (who in 1843 entered the monastery at Bruno, Czech Republic, and was elected abbot in 1868)[20] in the1860s, by his visual identification and counting study of traits (that disclosed for the first time how the dominant and recessive characters of these ranged through successive generations of say pea plants), discovered one aspect of sexual inheritance which is that parents can fail to always produce offspring with their own dominant features. For example, Mendel's yellow peas were hybrid forms with a recessive green feature latent but able to reappear in later generations. These inferences from tabulations were summarized by Mendel at a local natural history society meeting in 1865 and published in German in 1866 as a long report in the society's journal.[23]

Mendel's First Law of Inheritance: Each organism inherits two copies of a "trait" (allele), one from each parent; and, the one that each parent contributes is selected by chance from the two—designated in his tabulations of "dominant and recessive character (allele) pairs" by capital- and lower-case letters, Aa, Bb, for example—that each has. These traits (alleles) are particulate (they do not blend).

Darwin, who labored under the misapprehension that mated traits blend, could have been enlightened. But, as he made German a foreign language to its native speakers by his thoroughly English pronunciations of its words, perhaps he inclined not to *lesen Sie tschechische Publikationen*.

Mendel's work was appreciated as pioneering, but only in retrospect, after geneticist botanists Hugo de Vries, Carl Correns and Eric von Tschermak at the beginning of the 1900s, and William Bateson soon after, had independently rediscovered his (from gene to function) findings. These had never suffered from not being clear cut. Indeed, statistical analysis of this data by Fisher, who developed probability theory, revealed in 1936 that Mendel evidently did not save data outliers, which he or his field helpers privy to his (preconceived?)[24] theory perceived to be experimental errors. His theory was that both partners in sexual union contributed to the offspring. That they did so in definite amounts disproved the "spermatist" dogma which was the notion that the male contributes all. His refutation of that prevailing wisdom was his life's reward for his dislike of it. That, before his pea counting, had cost him his right to teach, suggests Colin Tudge in *The Impact of the Gene*, 2001.[25] Mutterings at the Abby was that his work was maverick and soon after his death his private notes were purposefully destroyed (nor did it help posterity that his findings remained available in the archived transactions of the Scientific Society of the City of Brunn, Austria—now Brno in the Czech Republic).

Since Mendel himself focused on the formation of hybrids, the iconoclasitic view is, he did not (could not) know that the characters (alleles) he studied were expressions of genes on the chromosomes.[26] The rediscoverers could (and did). They defaulted the laurels to Mendel by seeing in his results the modern concept that complementary pairs of inheritable characters (alleles) received from mates do not blend but can be dominant (*R*) or recessive (*r*) in the sense that offspring will show the characteristic *R* when it inherited the allele pairs *RR*, *Rr*, or *rR*, and will only show the characteristic r when it inherited the allele pair *rr*. (*R* and *r* are alleles, which are genes that occupy the same place on a chromosome.)

Darwin's simple claim was that the individual is favored by selection owing to the individual's overall quality. That is, natural selection acts on the phenotype. The Synthesis (the synthetic theory of evolution) of Darwinian natural selection (1859) and Mendelian inheritance (1865) is at the heart of neo-Darwinism, a term coined by George John Romanes in 1896 to designate "Dawinism *without* inheritance of acquired characters," as advocated by August Weismann who could claim that the individual's "somatic cells" are not perpetuated but what can be perpetuated down the ages are "germ cells" ("genes," in the later terminology of the geneticists). Selection acts on particular alleles in relation to their overall contribution to all individuals that carry copies of them. George C. Williams in *Adaptation and Natural* Selection,1966,[27] updated Darwin's understanding that evolution is not about survival of the species but is all about how natural selection operates on individuals *and* their genes.[28] Neo-Darwinism makes explicit that since alleles cannot be individually exposed to selection, each rides on the quality of the whole genome which itself cannot guarantee the quality of the phenotype (as the phenotype is at the mercy of environmental factors during its growth to fecundity). And more offspring more fit is too simplistic: An analysis of 21,000 records of survival in late 19th-century Utah, reveals that families with fewer children had more surviving grandchildren.[29]

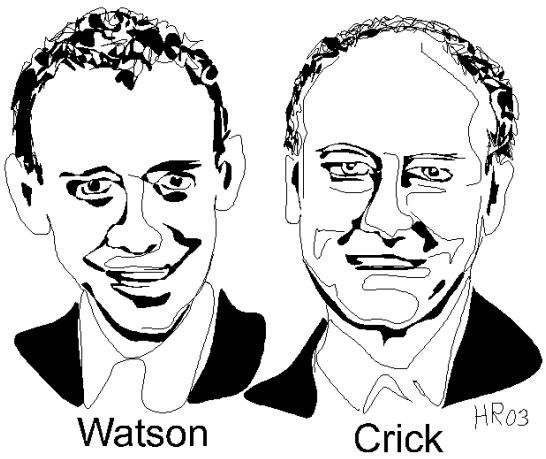
In a review of Robin Marantz Henig's book *The Monk in the Garden*, 2000,[30] Joe Cain finds that whereas "Mendel: focused on simple features that were easy to isolate: pea color and shape, plant height and so forth. The [post-1900] Mendelians went further representing all organisms—including humans—as divisible masses of simple biological atoms called 'traits.'... Eugenicists pressed on, suggesting ways to sort these units in human groups by scientific breeding, sterilization and worse ... we might ask: Why do we package qualities so ambiguous and multifaceted as I.Q. or sexuality into compact unit characters?"

Development, the process of organization of cells (morphogenesis), is achieved by the clumping of like (differentiation) and movement and arrangement of these into distinct tissues (spatial organization). What holds and draws into place is identified to be protruding parts of proteins of the cadherin family in cell surfaces. The clumping is because like cadherins attract like. A surprise is that spatial organization can be due to the drawing power cadherins (of different types) in proportion to their local concentrations rather than simply like attracting like.

Within cells, proteins typically number in the thousands and are: enzymes (which catalyze biochemical reactions in the cell from the formation of complex biological molecules to the breakdown of these), actin (with a structural role that gives a cell its shape, helps form compartments in which different cellular functions are partitioned, and binds with nucleic acids), hormones (for example, insulin), antibodies, and others which are involved in transportation of other molecules (for example, hemoglobin that carries oxygen in red blood cells).

Any change in a gene's DNA molecule is called a *mutation*. Such mutation can occur by chance at any time due to 1) non repair of lesions in DNA that arise from 1a) environmental agents such as the ultraviolet component of sunlight, ionizing radiation and numerous genotoxic chemicals (as are some compounds in cigarette smoke) or 1b) products of normal cellular metabolism such as reactive oxygen species (superoxide anions, hydroxyl radicals and hydrogen peroxide) derived from oxidative respiration and products of lipid peroxidation or 1c) spontaneous disintegration of some chemical bonds in DNA under physiological conditions such as hydrolysis of nucleotide residues that leaves non-instructive abasic sites, or 2) by insertion (retrovirus infections, for example, have left remnants that account of 8% of the human genome),[31] or 3) by misreplication.[32]

Jan H. J. Hoeijmakers in 2001 writes: "In view of the plethora of types of [DNA] lesions, no single repair process can cope with all kinds of damage. Instead, evolution has moulded a tapestry of sophisticated, interwoven DNA repair systems that as a whole cover most (but not all) of the insults inflicted on a cell's vital genetic information. ... Because the problem of DNA damage has existed *ab initio*, DNA repair systems must have arisen early in evolution. This explains why all known repair pathways are highly conserved (usually across the [bacterial]/eukaryotic evolutionary border). At least four main, partly overlapping damage repair pathways operate in mammals—nucleotide-excision repair, base-excision repair, homologous recombination, and end joining." Enzymatic repair of damaged RNA and proteins also takes place.[33] Every day, in every cell of a human body, may occur as many as 300,000 lesions that can stop the regular DNA-replication machinery. However SOS, a class of unusual DNA polymerases, allows DNA replication to proceed. DNA replication is thereby "efficient but relatively imprecise, leaving mistakes to error-correction enzymes which are themselves efficient because their substrates are specific mistakes made by other enzymes," writes Emil Schmidt: "Nature does not exhaust itself for the sake of perfectionism. Errors and infidelity, even wastefulness, provide innovation and robustness, ensuring the perpetuation of life." [34]



2 April 1953 'We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest. ... It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.'[35]

As early as the 1930s, the *Drosophila melanogaster* (fruitfly) pioneer geneticist Thomas Hunt Morgan had proposed that embryonic growth results from genes switching on and off, in the right place, and at the right time. After zoologist-geneticist **James Dewey Watson** (1928-) and ex-physicist **Francis Harry Compton Crick** (1916-2004) collaboratively—given the first good x-ray diffraction images of DNA fibers by chemist-crystallographers Maurice Hugh Frederick Wilkins (1916-2004) and Rosalind Elsie Franklin (1920-1958)[36] —unraveled 1953 thru 1961 the structure of DNA, ways could be devised to trace this activity (of genes turning on and off) as cells mature into liver, muscle, heart, or nerve. For this reason, the complete genome sequence is not a predictor without further biochemical knowledge of what molecules will be produced when the cell "reads" the genome. And currently, little is understood of the relationship of amino-acid sequences and protein-molecule foldings that determine what it does in a cell.

Proteomics is a new study that goes beyond understanding how proteins are made to the identification of their function in a cell.[37] Proteins fold by the formation of covalent bonds (as between two sulfur containing amino acids) and by a variety of non-covalent interactions among the amino acids. Many proteins fold to be globular, some are fibrous or have other distinctive three-dimensional shapes. The physiological environment plays a role in protein folding. Some proteins fold to expose a hydrophobic (water repelling) region which helps them bind with membranes. Most proteins are modified after they are made by the addition of sugars (glycosylation), phosphate (phosphorylation), sulfate and a few other small molecules. Their function is thereby modified and the same can cause some to become active or inactive. Certain proteins work by themselves and need no company. However, others work only when they are bound with other molecules of self or other proteins or cellular constituents. Proteins form such complexes by binding along surface clefts created by folding in a particular fashion, as well as by ionic and other non-covalent interactions.

Chrisriane **Nüsslein-Volhard** "It took me a long time to realize that the issue was gender. ... They expected less of a woman. The attitude was, 'I'll give her a chance, but I'm sure she won't perform.'"[40]

By 1995, developmental biologists such as Chrisriane **Nüsslein-Volhard**, Eric F. Weischaus and Edward B. Lewis (1918-2004), had made the surprising discovery, through their visual examination of thousands of fruit fly mutants, that of the 20,000 or so genes in its chromosomes only a handful (Hox genes, see below) actually generate the earliest blueprint for the insect's body plan. "By looking at the whole embryo from the outside, they could deduce what the genes were doing on the inside." In the case of the fly, morphogen molecules ("makers of structure") that the mother places in the egg, ooze through the embryo and create chemical gradients in it. These gradients in turn activate developmental genes in the egg that oversee its earliest organization.[38] The resulting cascade of biochemical activity ultimately tells cells where they are and what they ought to be. Distinctive behavior patterns can follow the same laws of inheritance as, say, did the stem-heights of Gregor Mendel's peas. Seymour Benzer first established this to be true, beginning in the 1960s, for a variety of behaviors fruit flies that exhibit, and which can be selected for, precise internal clocks verses none, reproductive zest with success versus elaborate courtship dance but no copulation or mere celibacy, and learned avoidance or no of an associated odor or an electric shock. Each *Time, Love, and Memory* behavior, as Jonathan Weiner calls them in his book by that name (published in 1999),[39] have identifiable locations on the fruit-fly chromosomes and the known genes can be experimentally turned on or off and, even when genetically damaged, repaired.

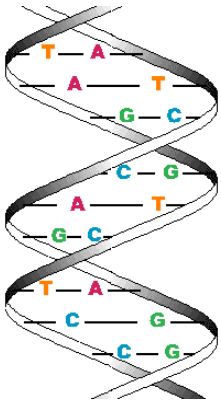


Figure f14.1 The DNA double helix consists of anti-parallel right-coiling helical strands, with complimentary bases (T-A and G-C).

More than a century has passed since Freidrich Miescher in 1869 isolated a novel chemical (DNA) from immune cells left in the pus on bloody bandages. The human genome pieced together by public (Francis Collins) and private (J. Craig Venter) sectors announced June 23, 2000, to be 99.9 percent complete, is DNA, a linear polymer structured as a double helix (specifically, as two *right*-handed anti-parallel coiling helixes) of 3.12 billion pairs of nitrogenous bases (each a cross link from one helix to the other) of T (thymine, a pyrimidine) to A (adenine, a purine), A to T, G (guanine, a purine) to C (cytosine, a pyrimidine), or C to G) (***Figure f14.1***). The significance of this is likened by Norton Zinder to the publication in 1543 of the first book on human anatomy.[41] All the parts of the human anatomy were listed, yet to illuminate the workings and interactions of these, physicians struggle still.

Barbara McClintock (1902-1992)

"That woman is either crazy or a genius!" exclaimed a colleague at Carnegie Institution of Washington, D.C., where she worked for 50 years.[45]

Her study of plant genetics (variation in the coloration of kernels of corn) won her the Nobel Prize for Physiology in 1983.

Less than 2 percent of our genome is genetic code for the proteins that let us live. Much of that is repeated code of transposons, the "jumping genes" first described by Barbara **McClintock** in the 1940s, that insert copies of themselves throughout the genome.[42] Add regulatory sequences, and the total of genes in the chromosomes required to bring you into existence is 5 percent of the genome. In 1995, S. Geo and K. J. Kemphues began to show how cells have evolved ways to interfere with gene activity that is particularly risky for germ line integrity but which, according to John McDonald, could account for some major evolutionary advances.[43] In *Genome, The Autobiography of a Species*, Mark Ridley in 2000 describes battlefield conditions of conflicts "between parental genes and childhood genes, or between male genes and female genes."[44] The remainder, 95 percent! of your genome, is of nonfunctioning (once-functioning) genes, HIV-like gene fossils (about 8 percent of human DNA) that are mostly dormant, and huge stretches of no-function (nonsense) code rife with genetic parasites that, willy nilly, will hitchhike into your children and their children's children.

Cells with different functions (liver cells, brain cells, etc.) produce particular proteins by turning on or turning off appropriate genes along the DNA molecule (***Figure f14.2***). In short, expression of genes is under complex hormonal, developmental, and tissue-specific regulation. At the prompting of its promoter (a region of DNA extending 150-300 base pairs upstream from the transcription start site that contains binding sites for RNA polymerase and a number of proteins that regulate the rate of transcription) a gene is first transcribed into a tightly edited molecule of pure information called *mRNA* [46] (messenger RNA first described by Elliot "Ken" Volkin and Lazarus Astrachan in 1956 as "DNA-like RNA")[47] that is free to travel. Varieties of mRNA are different numbers and orderings of nucleotides U (uracil), A, G, C) in a linear single chain. Each mRNA chain of nucleotides is the genetic code for a protein which is an ordered linear chain of amino acids. (The amino acids that living organisms produce and use are almost exclusively, for evolutionary circumstances yet to be elucidated, the *left*-hand mirror image of the right-hand structural form that is chemically equally possible).[48] The genetic code, discovered by Marshall Nirenberg in 1961, is of nucleotide triplets, called *codens* (***Figure f14.3***).[49]

Note: In RNA, U plays the role of T in DNA code. So the messenger RNA is read out from DNA with the rule of U (in RNA) links to A, and G links to C. The proteins that are coded for are assembled outside of the nucleus. The mRNA travels from where it is made in the nucleus to enter the cytoplasm outside the nuclear membrane where it can meet and bind with an organelle called a *ribosome*. Then a small variety of cytoplasmic RNA, known as *transfer RNA* (tRNA), starts to bring amino acids one at a time as called for to the ribosome and links them into a strand. The ribosome which glides along the mRNA is cued as to what calls to make by reading the mRNA's sequence of letter triplets. Thus, the nucleotide sequence present in the DNA, via the mRNA intermediary, ultimately directs the synthesis of protein with a precise and predetermined sequence. As the words of the code along the mRNA chain are triplets of letters that follow each other end to end, the reading must be told where to start and stop. The initiation sequence includes the start word AUG (which internally to a reading also codes for methionines). Termination is when the reading is of any of three stop words: UAA, UAG, and UGA.[50]

In human cells, the genome size is 3,200,000,000 AT, GC, letter pairs and the chromosome number is 23. By comparison, the respective numbers are for the mouse *Mus musculus*: 3,454,200,000 and 20, for the house fly *Musca domestica*: 900,000,000 and 5, and spoiling any patterns, for the amoeba *Amoeba dubia*: 670,000,000,000 and several hundreds. Bacterium *Mycoplasma genitalium*: 580,000

and 1 (circular) has the smallest genome (with 387 protein- and 43 RNA-coding genes).[51]And what is a recently found virus doing with 1,200,000 (and more than 900 genes)?[52]

In 1995, the journal *Nature* published a 379-page atlas of the human genome in which are listed partial sequences of nearly 30,000 human protein-coding genes. We then differed from mice by 300 genes but the number of human protein-coding genes has since been upped to 40,000 and in 2007 whittled to a firm count of 20,488 (an estimated 100 more may yet be found).[53] (The roundworm has 19,000, the fruitfly 13,000, and yeast 6,000 genes.)[54] J. Craig Venter, a significant contributor to this data base, estimates that of our genes: about 16 percent play a role in our metabolism, 12 percent are used for communication from cell to cell, 4 percent are devoted to helping us and our cells reproduce, and of the remainder, he suspects, most are active only in the brain; he has found 3,000 of these already and planned to scoop the sequencing of the entire human genome.[55] A tie was reached in the joint announcement made at the Whitehouse on June 26, 2000 of the human genome maps of its 3.2 billion AT, CG, letter pairs published in 2001 by Eric Lander in *Nature*, and by J. Craig Venter in *Science*. "The ultimate answer to the commandment 'Know thyself'" commented Walter Gilbert.

Genome sequencing is fundamental to an understanding of how species cope. By comparing genomes, we can reconstruct how one species could evolve from another and say what common ancestor species could have had. This application of the study of genes was first proposed by Paul Linus. Flagrant has been genetic change, reports Michele Cargill in 2003, in olfaction and hearing of humans over that of chimpanzee in which skeletal system genes, she finds, have been most labile.[56]

By contrast, genes controlling development can be anticipated to have been conserved through evolution. Nüsslein-Volhard is now studying zebra fish (*Danio rerio*) to know the mechanisms specific to the shaping of vertebrates structures such as their complex nervous system, blood vessels and kidneys.[57] Transparent embryos of these fish allow developmental abnormalities to be spotted by eye. The female lays 200–300 eggs each week (enough for the identification of candidate genes and positional cloning) and the embryos mature within a couple of months. Millions of zebra fish can also be bred in a small space. The control genes of these fish are now being systematically identified. Their genome has about 1.7 billion base-pairs and, as in other fish, this has long regions of synteny—where the order of the genes is similar to that found in humans and so likely play the same or similar functions in all other vertebrates, including humans.[58]

For example, a steadfast biopacemaker for circadian (pronounced: *sarKAYdian*) (coined 1959: L. *circa dies*, around a day) rhythms found in eukaryotic life as fungi, algae, higher plants, fruit flies, and mammals, has now been described. The surprise is that external environmental changes (such as light and temperature) have only a secondary tuning input. The primary mechanism for each cellular beat of time begins in the nucleus. There, special, always "on" initiator genes make proteins that activate "clock" genes in another region of the cell's DNA. The mRNA output of these activates, outside the nucleus, is the construction there of distinctive clock proteins that accumulate in the cytoplasm. At a certain concentration (that can be different for different organisms) the clock proteins bond in pairs and in this condition they gain entry into the nucleus. There they block the operation of the initiator genes. The clock genes stop production. In time the blocking proteins (by the action of an as yet unidentified enzyme) are eliminated in the nucleus. The cycle, which takes between 22 and 26 hours, resumes. The pacemaker for circadian rhythms in mammals is the pinhead sized suprachiasmatic nucleus (discovered in 1972) of roughly 20,000 neurons that is nestled in the dark, deep within the brain. Light-detecting molecules that track the day and continually adjusts the primary biological clock of animals and plants are proteins called *cryptochromes*. Our ungoverned wake-sleep cycle is 25 hours.[59] Cryptochromes that detect blue light have been found by David E. Somers to govern the daily (circadian) responses of the weed *Arabidopsis thaliana* (which tiny thale cress the British botanist and apothecary William Curtis described in 1777 in his *Flora Londinensis* as a plant of "no particular virtues or uses") genome sequence completed in 2000. Contrary to the prior assumption that DNA contains but one copy of most its genes, of the 25,500 genes of *A. thaliana*, two thirds are sequence-sortable, at one or *more*, into 11,600 families![60] Each *Arabidopsis* plant has two copies of each of it alleles. To shock, it is self fertilizing. *A.*'s cloistered reproduction

should see its each expressed mutant recessive allele pair in descendants, but to doubly shock this is not always so finds Mendelian voyeur Susan J. Lolle.[61] Doth some heritable RNA within the reproductive cell bear forward and insert ancestral instructions for normalcy?

Jeffrey C. Hall and Michael Rosbash have shown that fruit flies with mutations in a cryptochrome gene have altered circadian rhythms. Aziz Sancar has found that in mice lacking one of two mouse cryptochromes, their biological clocks run on a cycle 1 hour longer than those of normal mice. In addition, *A. thalana* genome research is on patterning (floral architecture which is controlled by a small number of genes found throughout the angiosperms), the "dwarf" gene (gibberellin is the response gene now recovered from it that contributed to improved grain yields during the "green revolution"), hormone biosynthesis (of brassinosteroids that regulate giantism or dwarfism in plants and animals), ethylene mutations (which can control many aspects of plant development), seed dispersal (how the several genes responsible can be manipulated to reduce pod-shatter seed loss in oil-seed crops such as canola).

Spawned from *in vivo* genome sequencing projects is Bioinformatics which Victor A. Simossis describes as "*in silico*" research that "uses biological rules and definitions to create methods for pre-processing experimental data and providing guidlines for research carried out at the bench. Methods can be applied to almost every single biological area, spanning from the design of primers for a given DNA sequence for sequencing or PCR [***Footnote f14.2***], to the visualisation of a protein 3D structure for the design of antibodies or drugs, to the determination of the evolutionary history of a single protein sequence, to the full characterization of a proteins properties and structure from sequence alone."[62]

Developmental genes are the molecules that signal by the position they happen to have in the chromosome which parts of the genetic code will be expressed, or suppressed, as a stem cell divides to become (and divides within) tissues and organs. In 1983, DNA sequence patterns recurrent in developmental genes, called *homeobox* of the fruitfly *Drosophila*, were discovered and reasoned to be conserved from a simpler remote ancestor. The homeobox gene tinman is active in heart development in flies *and* vertebrates. *Pax-6*, a developmental gene (discovered by Walter Gehring in 1994), is implicated in eye development across the animal kingdom.[63] Developmental genes are remarkable for their proven need to be conserved (in both structure and function). **Peter W. H. Holland** states that "most animal phyla possess essentially the same genes, and that some (but not all) of these genes change their development roles infrequently in evolution."[64] Evolutionary analyses indicate the ancestor of bilaterally symmetrical creatures had at least seven homeobox (homeotic or Hox) genes, organized into a single complex. Although gene order is endlessly shuffled by chromosomal rearrangements such as inversions and movements of large DNA segments, Hox genes through 600 million years of evolution stayed clustered together along chromosomes in a linear order that corresponds head (anterior) to tail (posterior) to their domains of tissues expression, function along the body axis,[65] and timing of expression. In mammals, the Hox family is a set of 39 developmental control genes, located in four complexes derived from duplications or is a single original cluster.[66] For plants, Michael Purugganan comments that analogous to HOX genes are a class of developmental MADS-box genes.[67]

The millimeter-long soil nematode's *Caenorhabditis elegans*, which consists of just 959 cells, was in 1999 the first complex organism to have its genome sequenced. Presently, it is the leading model organism in developmental biology for: "What is special about *C. elegans* is that we can see which genes are being expressed in which cell," says Paul Sternberg, "and that we know what each of the cells does, and how they relate to each other." [68]

Normal strains of the nematode *C. elegans* pass through four larval stages as they mature into fertile adults. A mutant strain, that, in the early 1990s Victor Ambros began to study, gets stuck at the first

stage (after molting, instead of moving on to the second larval stage, they simply repeated the first stage, the larva growing larger but never becoming a full-fledged adult). In 1993 came the surprise discovery that the gene that turns off other genes that play a role in the worm's development, encodes not a protein but a tiny (22-nucleotide) molecule of RNA.[69] In 1998, double-stranded RNA (dsRNA) was identified by Andrew Fire and Craig C. Mello as a key component of the gene-silencing.[70] Research has been rapid into understanding biological processes called *RNA interference* (RNAi) that involve either blocking the translation of specific messenger RNAs (post-transcriptional gene silencing; PTGS) or preventing transcription of specific regions of DNA into RNA (transcriptional gene silencing; TGS). One RNAi pathway governs distinct aspects of development, tissue-specific differentiation, and maintenance of differentiated functions in both animals and plants.[71]

To explain how so (and long-standing surprise) that the number of protein-coding genes in an organism doesn't seem to reflect its complexity, John Mattick with Michael J. Gagen in 2001 proposed that small RNA genes may account for the diversity and complexity of eukaryotes![72] Rapidly evolving non-coding regions could be what makes humans (non-coding DNA produces RNA and does not code for a protein). Greatly different from our nearest primate relatives is the gene, called *HAR1F*, in a non-coding segment of the genome. In the phenome, HAR1F is produced by brain cells called *Cajal–Retzius* cells, which regulate how the six layers of the cortex are laid down during development. The gene may interact with a protein called *reelin*, which plays a vital role in this layering. "But it's wild speculation," says geneticist David Haussler.[73]

Free radicals are atoms or molecular units such as hydroxyl (•OH) that have an unpaired electron. Electrons seek to travel in pairs. Free radicals are formed during aerobic respiration in cells. Each will find numerous opportunities to cut an electron from another molecule. That creates a free radical of it and so a damaging chain reaction is set off. In 1954, Denham Harman began to consider that taken-for-granted aging is due to such damage that the body fails to repair.[74] Evolutionary theory would predict that cells will have evolved to protect themselves. In 1967, biochemists discovered in cells the ubiquitous enzyme superoxide dismutase that sops up free-radicals.[75]

Fast forward to the present genomic revolution. This, Nicholas Wade in *Life Script*, 2001, posits will result in medicine's most radical shift in perspective since the germ theory of disease.[76] Addressable are: Allergies and infectious diseases that, at their root, have a vulnerability based in the proteins and RNA products of genes; Genetic diseases, through rational drug design (like *Gleevec*, that shows promise in treating a form of leukemia), in contrast to past trial-and-error discovery, to target known sites on molecules implicated in disease; Replacement of worn-out body parts, with new grown from stem cells, and guaranteed not to be rejected by the patient; Extend life spans, by staving off senescence (if aging is not a necessary outcome of life but is merely a byproduct); and, Pharmacogenomics (the study of the ways different genetic constitutions respond to drugs) for personalized medicine, making cross-reactions, allergic responses to drugs and lethal underdoses, a thing of the past. ❑

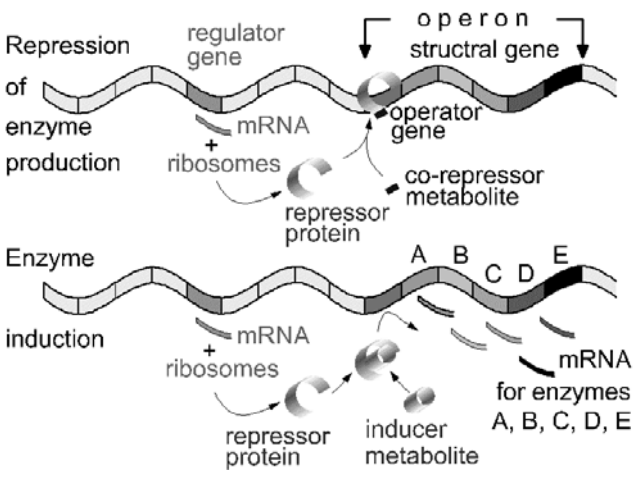


Figure f14.2 [77] A regulator gene occupying one site on the DNA strand controls the production of a repressor protein which, in combination with a co-repressor metabolite, inhibits the activity of an operator gene on another site of the DNA strand. In the presence of an inducer metabolite (that is the nutrient to be digested), the repressor protein is unable to block the operator site. The operator gene controls the activity of adjacent structural genes on which there is an assembly of messenger RNA (mRNA) molecules involved in manufacture of those enzymes used to digest the nutrient. An example is the lac operon which is a cluster of genes encoding proteins that metabolize the milk sugar lactose. In *E. coli* it allows the bacterium to have established its current home in the colons of mammals.

Figure f14.3 [50] The genetic code is the algorithm that translates information (nucleotide triplets in genes) contained in DNA into specific amino acids in proteins. The four bases (nucleotide), U, C, A, G, that serve as letters of the genetic alphabet are arranged in 64 three-letter words called *codons*. The genetic code, for as yet to be explained reasons, is highly redundant as the 64 codons call for the production of only 20 amino acids.[78] Arg (arginine) is called for by six of the codons, and likewise for leucine, and serine. Others are called for by three or two codons. Only Met (methionine)and Trp (tryptophan) are read for their production from one codon each. The reading is done by ribosomes which, as they do so, add the called for amino acid to the protein sequence that they are writing. Not so used are three Stop codons that signal the end of a protein chain. The genetic code is the same in bacteria, archaea and eucarya and so antedates life's split into these domains. In certain archaea, selenocysteine[79] and pyrrolysine are additional natural amino acids encoded (S. K. Blight in 2004 has shown) in DNA using the general mechanism employed for the common set of 20 amino acids.[80]

Footnote f14.1 "You shall never be Fellow here as long as I live" Pearson hissed at Hurst, after a particularly contentious meeting about the genetics of horse coat color at the Royal Society in London. On the one side, the Mendelians, including Francis Galton, William Bateson and Charles Hurst, who accepted Mendelism but considered natural selection as ineffective, seeing evolution as occurring by 'macromutations', or single genetic changes of very large effect. On the other side stood the biometricians, most notably Karl Pearson and Raphael Weldon, who accepted the ubiquity of Darwinian selection but rejected Mendelian genetics.[81]

Footnote f14.2 " ...the easiest way to ready DNA for duplication is to boil it, separating it into its two constituent strands. Repeated boiling and cooling of a solution containing both DNA and Taq polymerase duplicates — or 'amplifies' — even the most minute quantities of original DNA. The method is called the 'polymerase chain reaction', or PCR, and it is brilliantly clever." —Richard Dawkins. [82]

Comment f15.1 **The ways of flesh**

So why in 2001 did stem-cell pioneer Roger Pedersen leave the University of California, San Francisco? For "very favorable circumstances and tremendous support" in Britain rather than "sitting on my hands for the next few years" in the United States he told *The Wall Street Journal*. He took a faculty position at Cambridge University where human-embryo research, including work on stem cells, is legal and publicly supported in Britain.[1] How so, the latter?

In 1990, the British parliament was persuaded in favor (actually they were in favour) of allowing a limited range of research projects on pre-gastrulation embryos by arguments that such embryos have not progressed to the point that they can be considered as individuals (of which, in natural reproduction, 80% will perish before birth). The distinction of organic matter being not a human (for example an embryo before gastrulation) and a human (an individual) is lost on 'Pro-life' groups who define the beginning of human life as the union of sperm and egg, and equate the harvesting of human embryonic stem cells to homicide. One might shed tears when a thorn has nicked out a part of ones body but the tears are not for the loss of the hundreds of potential clones that the nuclei of the removed cells could become given the circumstances that produced the sheep-clone Dolly from the teat-cell of its, well, clone.

The natural phenomenon of identical twins—around 1 in 250 human births—mocks the simple view that at fertilization a single life (soul) starts. Before gastrulation, the cell, or cloned mass of them, like a nicked out hair plug awaiting relocation into a bald pate, is at all times alive but is not an individual or a group of individuals. But under other circumstances, as insertion of the material into the right environment, any one of the cells has the potential to split (that is to be an embryo) to form two or more viable cells which then split and again up to an embryo comprising a hollow ball of cells called a *blastocyst*. Thereafter, to 'gastrulation', when its cells begin to migrate into distinct layers that form the basis of the adult body plan. Because of twinning, and of adult nuclear-transfer cloning, many bioethicists conclude that it is only after gastrulation that an embryo can begin to be considered to be an individual human being. [2]