



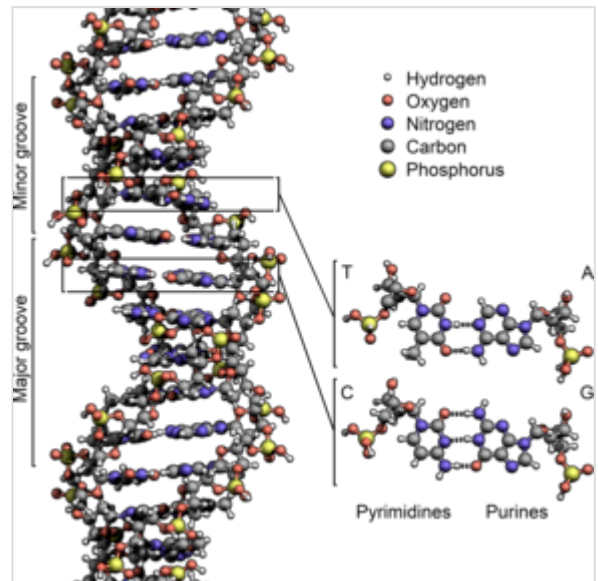
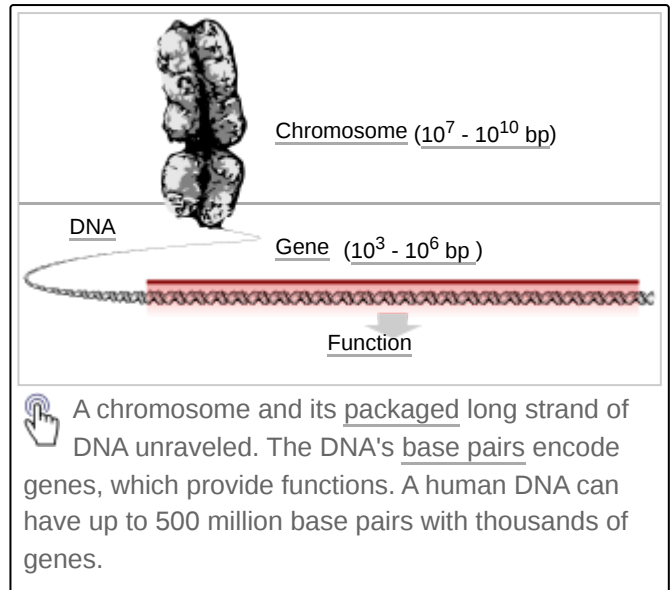
DNA

Deoxyribonucleic acid

(/diːˈɒksɪˌraɪboʊnjuːkliːɪk, -kleɪ/ [ⓘ] ^[1] **DNA**) is a polymer composed of two polynucleotide chains that coil around each other to form a double helix. The polymer carries genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses. DNA and ribonucleic acid (RNA) are nucleic acids. Alongside proteins, lipids and complex carbohydrates (polysaccharides), nucleic acids are one of the four major types of macromolecules that are essential for all known forms of life.

The two DNA strands are known as polynucleotides as they are composed of simpler monomeric units called nucleotides.^{[2][3]} Each nucleotide is composed of one of four nitrogen-containing nucleobases (cytosine [C], guanine [G], adenine [A] or thymine [T]), a sugar called deoxyribose, and a phosphate group. The nucleotides are joined to one another in a chain by covalent bonds (known as the phosphodiester linkage) between the sugar of one nucleotide and the phosphate of the next, resulting in an alternating sugar-phosphate backbone. The nitrogenous bases of the two separate polynucleotide strands are bound together, according to base pairing rules (A with T and C with G), with hydrogen bonds to make double-stranded DNA. The complementary nitrogenous bases are divided into two groups, the single-ringed pyrimidines and the double-ringed purines. In DNA, the pyrimidines are thymine and cytosine; the purines are adenine and guanine.

Both strands of double-stranded DNA store the same biological information. This information is replicated when the two strands separate. A large part of DNA (more than 98% for humans) is non-coding, meaning that these sections do not serve as patterns for protein sequences. The two strands of DNA run in opposite directions to each other and are thus antiparallel. Attached to each sugar is one of four types of nucleobases (or *bases*). It is the sequence of these four nucleobases along the backbone that encodes genetic information. RNA strands are created



The structure of the DNA double helix (type B-DNA). The atoms in the structure are colour-coded by element and the detailed structures of two base pairs are shown in the bottom right.

using DNA strands as a template in a process called transcription, where DNA bases are exchanged for their corresponding bases except in the case of thymine (T), for which RNA substitutes uracil (U).^[4] Under the genetic code, these RNA strands specify the sequence of amino acids within proteins in a process called translation.

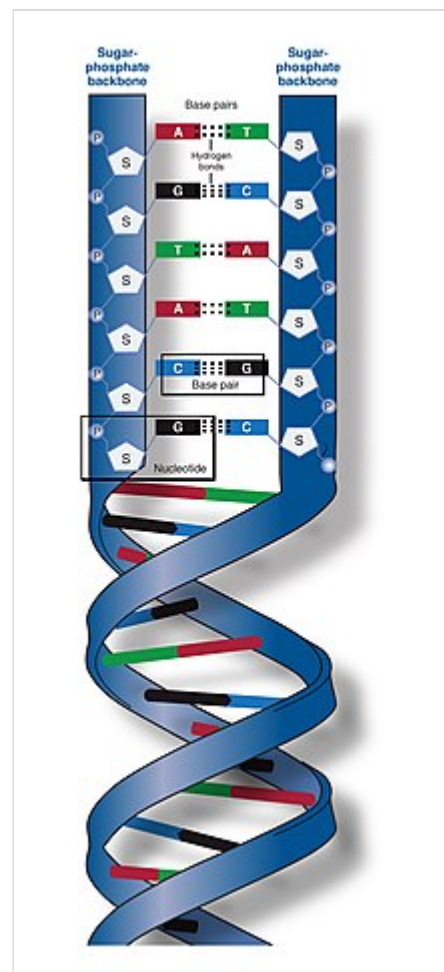
Within eukaryotic cells, DNA is organized into long structures called chromosomes. Before typical cell division, these chromosomes are duplicated in the process of DNA replication, providing a complete set of chromosomes for each daughter cell. Eukaryotic organisms (animals, plants, fungi and protists) store most of their DNA inside the cell nucleus as nuclear DNA, and some in the mitochondria as mitochondrial DNA or in chloroplasts as chloroplast DNA.^[5] In contrast, prokaryotes (bacteria and archaea) store their DNA only in the cytoplasm, in circular chromosomes. Within eukaryotic chromosomes, chromatin proteins, such as histones, compact and organize DNA. These compacting structures guide the interactions between DNA and other proteins, helping control which parts of the DNA are transcribed.

Properties

DNA is a long polymer made from repeating units called nucleotides.^{[6][7]} The structure of DNA is dynamic along its length, being capable of coiling into tight loops and other shapes.^[8] In all species it is composed of two helical chains, bound to each other by hydrogen bonds. Both chains are coiled around the same axis, and have the same pitch of 34 ångströms (3.4 nm). The pair of chains have a radius of 10 Å (1.0 nm).^[9] According to another study, when measured in a different solution, the DNA chain measured 22–26 Å (2.2–2.6 nm) wide, and one nucleotide unit measured 3.3 Å (0.33 nm) long.^[10] The buoyant density of most DNA is 1.7g/cm³.^[11]

DNA does not usually exist as a single strand, but instead as a pair of strands that are held tightly together.^{[9][12]} These two long strands coil around each other, in the shape of a double helix. The nucleotide contains both a segment of the backbone of the molecule (which holds the chain together) and a nucleobase (which interacts with the other DNA strand in the helix). A nucleobase linked to a sugar is called a nucleoside, and a base linked to a sugar and to one or more phosphate groups is called a nucleotide. A biopolymer comprising multiple linked nucleotides (as in DNA) is called a polynucleotide.^[13]

The backbone of the DNA strand is made from alternating phosphate and sugar groups.^[14] The sugar in DNA is 2-deoxyribose, which is a pentose (five-carbon) sugar. The sugars are joined by phosphate groups that form phosphodiester bonds between the third and fifth carbon atoms of adjacent sugar rings. These are known as the 3'-end (three prime end), and 5'-end (five prime end) carbons, the prime symbol being used to distinguish these carbon atoms from those of the base to which the deoxyribose forms a glycosidic bond.^[12]



Simplified diagram

Therefore, any DNA strand normally has one end at which there is a phosphate group attached to the 5' carbon of a ribose (the 5' phosphoryl) and another end at which there is a free hydroxyl group attached to the 3' carbon of a ribose (the 3' hydroxyl). The orientation of the 3' and 5' carbons along the sugar-phosphate backbone confers directionality (sometimes called polarity) to each DNA strand. In a nucleic acid double helix, the direction of the nucleotides in one strand is opposite to their direction in the other strand: the strands are antiparallel. The asymmetric ends of DNA strands are said to have a directionality of five prime end (5'), and three prime end (3'), with the 5' end having a terminal phosphate group and the 3' end a terminal hydroxyl group. One major difference between DNA and RNA is the sugar, with the 2-deoxyribose in DNA being replaced by the related pentose sugar ribose in RNA.^[12]

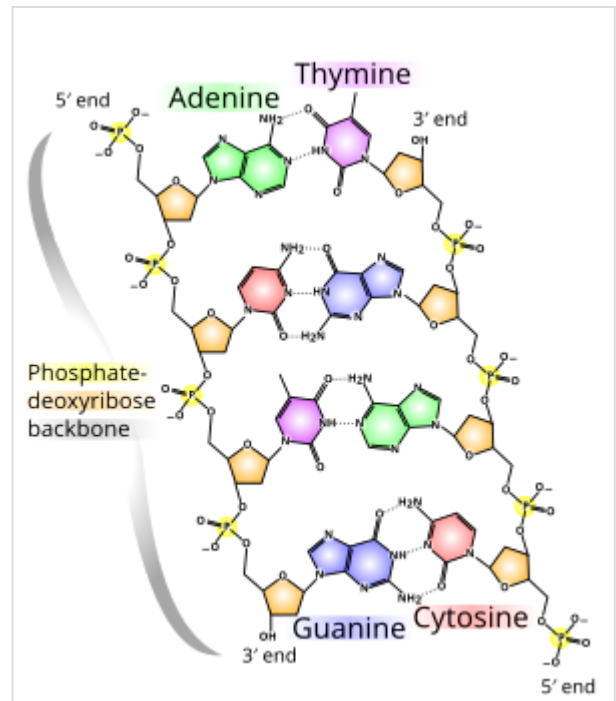
The DNA double helix is stabilized primarily by two forces: hydrogen bonds between nucleotides and base-stacking interactions among aromatic nucleobases.^[16] The four bases found in DNA are adenine (A), cytosine (C), guanine (G) and thymine (T). These four bases are attached to the sugar-phosphate to form the complete nucleotide, as shown for adenosine monophosphate. Adenine pairs with thymine and guanine pairs with cytosine, forming A-T and G-C base pairs.^{[17][18]}

Nucleobase classification

The nucleobases are classified into two types: the purines, A and G, which are fused five- and six-membered heterocyclic compounds, and the pyrimidines, the six-membered rings C and T.^[12] A fifth pyrimidine nucleobase, uracil (U), usually takes the place of thymine in RNA and differs from thymine by lacking a methyl group on its ring. In addition to RNA and DNA, many artificial nucleic acid analogues have been created to study the properties of nucleic acids, or for use in biotechnology.^[19]

Non-canonical bases

Modified bases occur in DNA. The first of these recognized was 5-methylcytosine, which was found in the genome of *Mycobacterium tuberculosis* in 1925.^[20] The reason for the presence of these noncanonical bases in bacterial viruses (bacteriophages) is to avoid the restriction enzymes present in



Chemical structure of DNA; hydrogen bonds shown as dotted lines. Each end of the double helix has an exposed 5' phosphate on one strand and an exposed 3' hydroxyl group (—OH) on the other.



A section of DNA. The bases lie horizontally between the two spiraling strands^[15] (animated version).

bacteria. This enzyme system acts at least in part as a molecular immune system protecting bacteria from infection by viruses.^[21] Modifications of the bases cytosine and adenine, the more common and modified DNA bases, play vital roles in the epigenetic control of gene expression in plants and animals.^[22]

A number of noncanonical bases are known to occur in DNA.^[23] Most of these are modifications of the canonical bases plus uracil.

- **Modified Adenine**
 - N6-carbamoyl-methyladenine
 - N6-methyladenine
- **Modified Guanine**
 - 7-Deazaguanine
 - 7-Methylguanine
- **Modified Cytosine**
 - N4-Methylcytosine
 - 5-Carboxylcytosine
 - 5-Formylcytosine
 - 5-Glycosylhydroxymethylcytosine
 - 5-Hydroxycytosine
 - 5-Methylcytosine
- **Modified Thymidine**
 - α -Glutamythymidine
 - α -Putrescinythymine
- **Uracil and modifications**
 - Base J
 - Uracil
 - 5-Dihydroxypentauracil
 - 5-Hydroxymethyldeoxyuracil
- **Others**
 - Deoxyarchaeosine
 - 2,6-Diaminopurine (2-Aminoadenine)

Grooves

Twin helical strands form the DNA backbone. Another double helix may be found tracing the spaces, or grooves, between the strands. These voids are adjacent to the base pairs and may provide a binding site. As the strands are not symmetrically located with respect to each other, the grooves are unequally sized. The major groove is 22 ångströms (2.2 nm) wide, while the minor groove is 12 Å (1.2 nm) in width.^[24] Due to the larger width of the major groove, the edges of the bases are more accessible in the major groove than in the minor groove. As a result, proteins such as transcription factors that can bind to specific sequences in double-stranded DNA usually make contact with the sides of the bases exposed in

the major groove.^[25] This situation varies in unusual conformations of DNA within the cell (*see below*), but the major and minor grooves are always named to reflect the differences in width that would be seen if the DNA was twisted back into the ordinary B form.

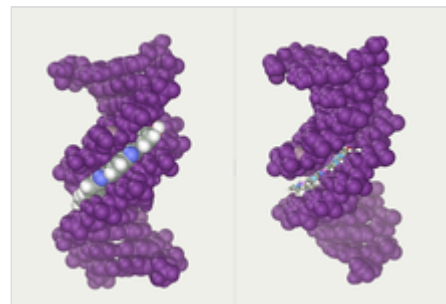
Base pairing

In a DNA double helix, each type of nucleobase on one strand bonds with just one type of nucleobase on the other strand. This is called complementary base pairing. Purines form hydrogen bonds to pyrimidines, with adenine bonding only to thymine in two hydrogen bonds, and cytosine bonding only to guanine in three hydrogen bonds. This arrangement of two nucleotides binding together across the double helix (from six-carbon ring to six-carbon ring) is called a Watson-Crick base pair. DNA with high GC-content is more stable than DNA with low GC-content. A Hoogsteen base pair (hydrogen bonding the 6-carbon ring to the 5-carbon ring) is a rare variation of base-pairing.^[26] As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily. The two strands of DNA in a double helix can thus be pulled apart like a zipper, either by a mechanical force or high temperature.^[27] As a result of this base pair complementarity, all the information in the double-stranded sequence of a DNA helix is duplicated on each strand, which is vital in DNA replication. This reversible and specific interaction between complementary base pairs is critical for all the functions of DNA in organisms.^[7]

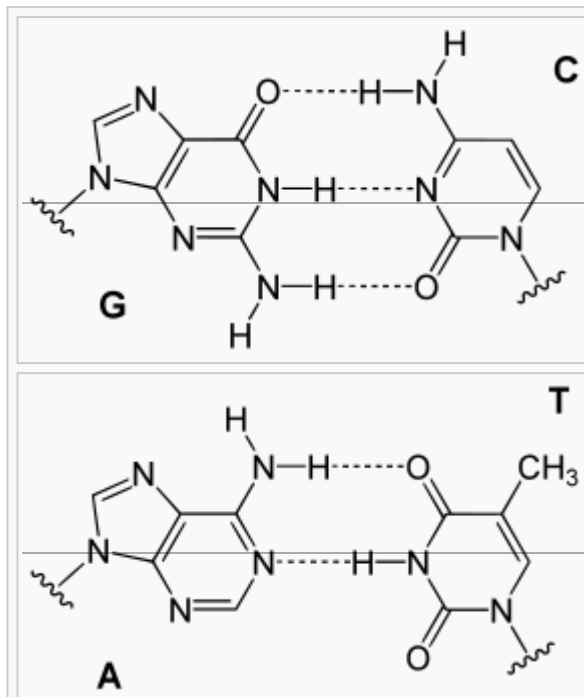
ssDNA vs. dsDNA

Most DNA molecules are actually two polymer strands, bound together in a helical fashion by noncovalent bonds; this double-stranded (dsDNA) structure is maintained largely by the intrastrand base stacking interactions, which are strongest for G, C stacks. The two strands can come apart—a process known as melting—to form two single-stranded DNA (ssDNA) molecules. Melting occurs at high temperatures, low salt and high pH (low pH also melts DNA, but since DNA is unstable due to acid depurination, low pH is rarely used).

The stability of the dsDNA form depends not only on the GC-content (% G, C basepairs) but also on sequence (since stacking is sequence specific) and also length (longer molecules are more stable). The stability can be measured in various ways; a common way is the melting temperature (also called T_m value), which is the temperature at which 50% of the double-strand molecules are converted to single-strand molecules; melting temperature is dependent on ionic strength and the concentration of DNA. As a



DNA major and minor grooves. The latter is a binding site for the Hoechst stain dye 33258.



Top, a **GC** base pair with three hydrogen bonds. Bottom, an **AT** base pair with two hydrogen bonds. Non-covalent hydrogen bonds between the pairs are shown as dashed lines.

result, it is both the percentage of GC base pairs and the overall length of a DNA double helix that determines the strength of the association between the two strands of DNA. Long DNA helices with a high GC-content have more strongly interacting strands, while short helices with high AT content have more weakly interacting strands.^[28] In biology, parts of the DNA double helix that need to separate easily, such as the TATAAT Pribnow box in some promoters, tend to have a high AT content, making the strands easier to pull apart.^[29]

In the laboratory, the strength of this interaction can be measured by finding the melting temperature T_m necessary to break half of the hydrogen bonds. When all the base pairs in a DNA double helix melt, the strands separate and exist in solution as two entirely independent molecules. These single-stranded DNA molecules have no single common shape, but some conformations are more stable than others.^[30]

Amount

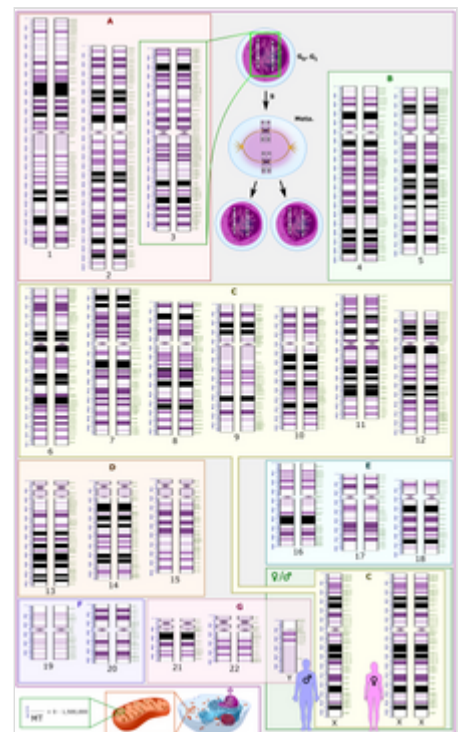
In humans, the total female diploid nuclear genome per cell extends for 6.37 Gigabase pairs (Gbp), is 208.23 cm long and weighs 6.51 picograms (pg).^[31] Male values are 6.27 Gbp, 205.00 cm, 6.41 pg.^[31] Each DNA polymer can contain hundreds of millions of nucleotides, such as in chromosome 1. Chromosome 1 is the largest human chromosome with approximately 220 million base pairs, and would be 85 mm long if straightened.^[32]

In eukaryotes, in addition to nuclear DNA, there is also mitochondrial DNA (mtDNA) which encodes certain proteins used by the mitochondria. The mtDNA is usually relatively small in comparison to the nuclear DNA. For example, the human mitochondrial DNA forms closed circular molecules, each of which contains 16,569^{[33][34]} DNA base pairs,^[35] with each such molecule normally containing a full set of the mitochondrial genes. Each human mitochondrion contains, on average, approximately 5 such mtDNA molecules.^[35] Each human cell contains approximately 100 mitochondria, giving a total number of mtDNA molecules per human cell of approximately 500.^[35] However, the amount of mitochondria per cell also varies by cell type, and an egg cell can contain 100,000 mitochondria, corresponding to up to 1,500,000 copies of the mitochondrial genome (constituting up to 90% of the DNA of the cell).^[36]

Sense and antisense

A DNA sequence is called a "sense" sequence if it is the same as that of a messenger RNA copy that is translated into protein.^[37]

The sequence on the opposite strand is called the "antisense" sequence. Both sense and antisense sequences can exist on different parts of the same strand of DNA (i.e. both strands can contain both sense and antisense sequences). In both prokaryotes and eukaryotes,



Schematic karyogram of a human. It shows 22 homologous chromosomes, both the female (XX) and male (XY) versions of the sex chromosome (bottom right), as well as the mitochondrial genome (to scale at bottom left). The blue scale to the left of each chromosome pair (and the mitochondrial genome) shows its length in terms of millions of DNA base pairs.

antisense RNA sequences are produced, but the functions of these RNAs are not entirely clear.^[38] One proposal is that antisense RNAs are involved in regulating gene expression through RNA-RNA base pairing.^[39]

A few DNA sequences in prokaryotes and eukaryotes, and more in plasmids and viruses, blur the distinction between sense and antisense strands by having overlapping genes.^[40] In these cases, some DNA sequences do double duty, encoding one protein when read along one strand, and a second protein when read in the opposite direction along the other strand. In bacteria, this overlap may be involved in the regulation of gene transcription,^[41] while in viruses, overlapping genes increase the amount of information that can be encoded within the small viral genome.^[42]

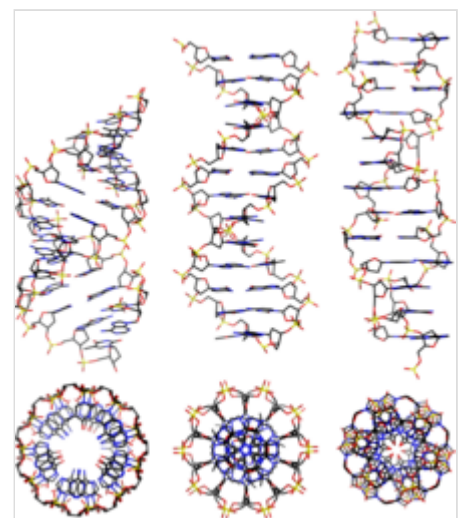
Supercoiling

DNA can be twisted like a rope in a process called DNA supercoiling. With DNA in its "relaxed" state, a strand usually circles the axis of the double helix once every 10.4 base pairs, but if the DNA is twisted the strands become more tightly or more loosely wound.^[43] If the DNA is twisted in the direction of the helix, this is positive supercoiling, and the bases are held more tightly together. If they are twisted in the opposite direction, this is negative supercoiling, and the bases come apart more easily. In nature, most DNA has slight negative supercoiling that is introduced by enzymes called topoisomerases.^[44] These enzymes are also needed to relieve the twisting stresses introduced into DNA strands during processes such as transcription and DNA replication.^[45]

Alternative DNA structures

DNA exists in many possible conformations that include A-DNA, B-DNA, and Z-DNA forms, although only B-DNA and Z-DNA have been directly observed in functional organisms.^[14] The conformation that DNA adopts depends on the hydration level, DNA sequence, the amount and direction of supercoiling, chemical modifications of the bases, the type and concentration of metal ions, and the presence of polyamines in solution.^[46]

The first published reports of A-DNA X-ray diffraction patterns—and also B-DNA—used analyses based on Patterson functions that provided only a limited amount of structural information for oriented fibers of DNA.^{[47][48]} An alternative analysis was proposed by Wilkins *et al.* in 1953 for the *in vivo* B-DNA X-ray diffraction-scattering patterns of highly hydrated DNA fibers in terms of squares of Bessel functions.^[49] In the same journal, James Watson and Francis Crick presented their molecular modeling analysis of the DNA X-ray diffraction patterns to suggest that the structure was a double helix.^[9]



From left to right, the structures of A, B and Z-DNA

Although the *B-DNA form* is most common under the conditions found in cells,^[50] it is not a well-defined conformation but a family of related DNA conformations^[51] that occur at the high hydration levels present in cells. Their corresponding X-ray diffraction and scattering patterns are characteristic of molecular paracrystals with a significant degree of disorder.^{[52][53]}

Compared to B-DNA, the A-DNA form is a wider right-handed spiral, with a shallow, wide minor groove and a narrower, deeper major groove. The A form occurs under non-physiological conditions in partly dehydrated samples of DNA, while in the cell it may be produced in hybrid pairings of DNA and RNA strands, and in enzyme-DNA complexes.^{[54][55]} Segments of DNA where the bases have been chemically modified by methylation may undergo a larger change in conformation and adopt the Z form. Here, the strands turn about the helical axis in a left-handed spiral, the opposite of the more common B form.^[56] These unusual structures can be recognized by specific Z-DNA binding proteins and may be involved in the regulation of transcription.^[57]

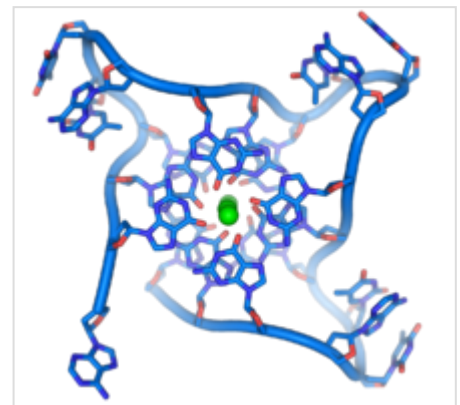
Alternative DNA chemistry

For many years, exobiologists have proposed the existence of a shadow biosphere, a postulated microbial biosphere of Earth that uses radically different biochemical and molecular processes than currently known life. One of the proposals was the existence of lifeforms that use arsenic instead of phosphorus in DNA. A report in 2010 of the possibility in the bacterium GFAJ-1 was announced,^{[58][59]} though the research was disputed,^{[59][60]} and evidence suggests the bacterium actively prevents the incorporation of arsenic into the DNA backbone and other biomolecules.^[61]

Quadruplex structures

At the ends of the linear chromosomes are specialized regions of DNA called telomeres. The main function of these regions is to allow the cell to replicate chromosome ends using the enzyme telomerase, as the enzymes that normally replicate DNA cannot copy the extreme 3' ends of chromosomes.^[63] These specialized chromosome caps also help protect the DNA ends, and stop the DNA repair systems in the cell from treating them as damage to be corrected.^[64] In human cells, telomeres are usually lengths of single-stranded DNA containing several thousand repeats of a simple TTAGGG sequence.^[65]

These guanine-rich sequences may stabilize chromosome ends by forming structures of stacked sets of four-base units, rather than the usual base pairs found in other DNA molecules. Here, four guanine bases, known as a guanine tetrad, form a flat plate. These flat four-base units then stack on top of each other to form a stable G-quadruplex structure.^[66] These structures are stabilized by hydrogen bonding between the edges of the bases and chelation of a metal ion in the centre of each four-base unit.^[67] Other structures can also be formed, with the central set of four bases coming from either a single strand folded around the bases, or several different parallel strands, each contributing one base to the central structure.

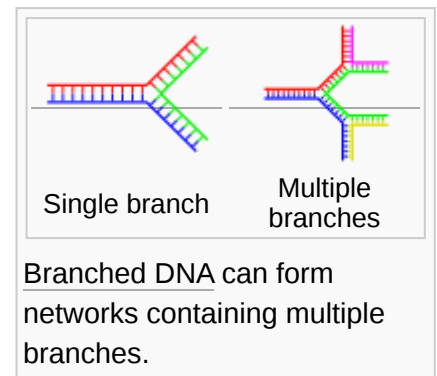


DNA quadruplex formed by telomere repeats. The looped conformation of the DNA backbone is very different from the typical DNA helix. The green spheres in the center represent potassium ions.^[62]

In addition to these stacked structures, telomeres also form large loop structures called telomere loops, or T-loops. Here, the single-stranded DNA curls around in a long circle stabilized by telomere-binding proteins.^[68] At the very end of the T-loop, the single-stranded telomere DNA is held onto a region of double-stranded DNA by the telomere strand disrupting the double-helical DNA and base pairing to one of the two strands. This triple-stranded structure is called a displacement loop or D-loop.^[66]

Branched DNA

In DNA, fraying occurs when non-complementary regions exist at the end of an otherwise complementary double-strand of DNA. However, branched DNA can occur if a third strand of DNA is introduced and contains adjoining regions able to hybridize with the frayed regions of the pre-existing double-strand. Although the simplest example of branched DNA involves only three strands of DNA, complexes involving additional strands and multiple branches are also possible.^[69] Branched DNA can be used in nanotechnology to construct geometric shapes, see the section on uses in technology below.



Artificial bases

Several artificial nucleobases have been synthesized, and successfully incorporated in the eight-base DNA analogue named Hachimoji DNA. Dubbed S, B, P, and Z, these artificial bases are capable of bonding with each other in a predictable way (S–B and P–Z), maintain the double helix structure of DNA, and be transcribed to RNA. Their existence could be seen as an indication that there is nothing special about the four natural nucleobases that evolved on Earth.^{[70][71]} On the other hand, DNA is tightly related to RNA which does not only act as a transcript of DNA but also performs as molecular machines many tasks in cells. For this purpose it has to fold into a structure. It has been shown that to allow to create all possible structures at least four bases are required for the corresponding RNA,^[72] while a higher number is also possible but this would be against the natural principle of least effort.

Acidity

The phosphate groups of DNA give it similar acidic properties to phosphoric acid and it can be considered as a strong acid. It will be fully ionized at a normal cellular pH, releasing protons which leave behind negative charges on the phosphate groups. These negative charges protect DNA from breakdown by hydrolysis by repelling nucleophiles which could hydrolyze it.^[73]

Macroscopic appearance

Pure DNA extracted from cells forms white, stringy clumps.^[74]

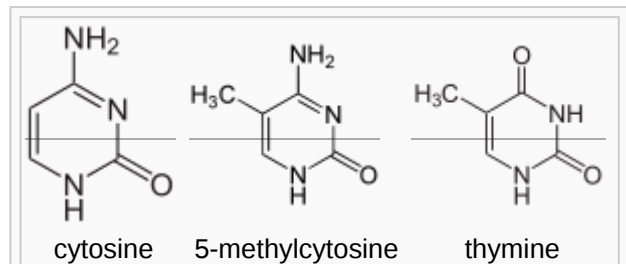
Chemical modifications and altered DNA packaging



Impure DNA extracted from an orange

Base modifications and DNA packaging

The expression of genes is influenced by how the DNA is packaged in chromosomes, in a structure called chromatin. Base modifications can be involved in packaging, with regions that have low or no gene expression usually containing high levels of methylation of cytosine bases. DNA packaging and its influence on gene expression can also occur by covalent modifications of the histone protein core around which DNA is wrapped in the chromatin structure or else by remodeling carried out by chromatin remodeling complexes (see Chromatin remodeling). There is, further, crosstalk between DNA methylation and histone modification, so they can coordinately affect chromatin and gene expression.^[75]



Structure of cytosine with and without the 5-methyl group. Deamination converts 5-methylcytosine into thymine.

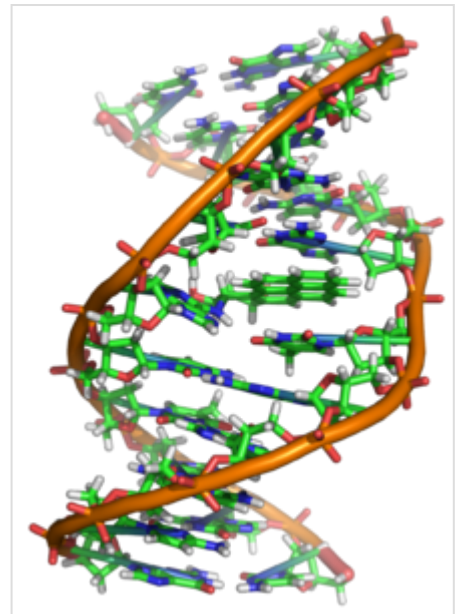
For one example, cytosine methylation produces 5-methylcytosine, which is important for X-inactivation of chromosomes.^[76] The average level of methylation varies between organisms—the worm *Caenorhabditis elegans* lacks cytosine methylation, while vertebrates have higher levels, with up to 1% of their DNA containing 5-methylcytosine.^[77] Despite the importance of 5-methylcytosine, it can deaminate to leave a thymine base, so methylated cytosines are particularly prone to mutations.^[78] Other base modifications include adenine methylation in bacteria, the presence of 5-hydroxymethylcytosine in the brain,^[79] and the glycosylation of uracil to produce the "J-base" in kinetoplastids.^{[80][81]}

Damage

DNA can be damaged by many sorts of mutagens, which change the DNA sequence. Mutagens include oxidizing agents, alkylating agents and also high-energy electromagnetic radiation such as ultraviolet light and X-rays. The type of DNA damage produced depends on the type of mutagen. For example, UV light can damage DNA by producing thymine dimers, which are cross-links between pyrimidine bases.^[83] On the other hand, oxidants such as free radicals or hydrogen peroxide produce multiple forms of damage, including base modifications, particularly of guanine, and double-strand breaks.^[84] A typical human cell contains about 150,000 bases that have suffered oxidative damage.^[85] Of these oxidative lesions, the most dangerous are double-strand breaks, as these are difficult to repair and can produce point mutations, insertions, deletions from the DNA sequence, and chromosomal translocations.^[86] These mutations can cause cancer. Because of inherent limits in the DNA repair mechanisms, if humans lived long enough, they would all eventually develop cancer.^{[87][88]} DNA damages that are naturally occurring, due to normal cellular processes that produce reactive oxygen species, the hydrolytic activities of cellular water, etc., also occur frequently. Although most of these

damages are repaired, in any cell some DNA damage may remain despite the action of repair processes. These remaining DNA damages accumulate with age in mammalian postmitotic tissues. This accumulation appears to be an important underlying cause of aging.^{[89][90][91]}

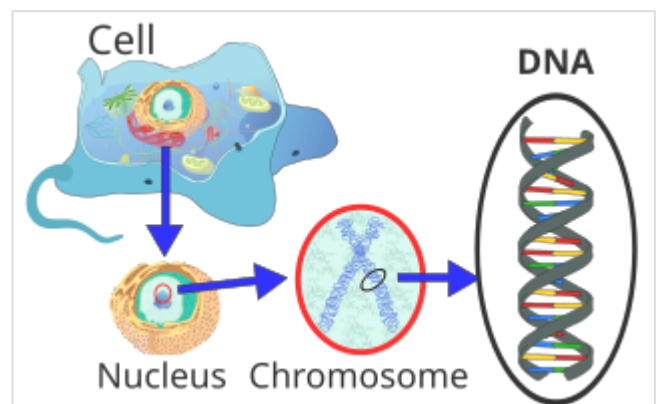
Many mutagens fit into the space between two adjacent base pairs, this is called intercalation. Most intercalators are aromatic and planar molecules; examples include ethidium bromide, acridines, daunomycin, and doxorubicin. For an intercalator to fit between base pairs, the bases must separate, distorting the DNA strands by unwinding of the double helix. This inhibits both transcription and DNA replication, causing toxicity and mutations.^[92] As a result, DNA intercalators may be carcinogens, and in the case of thalidomide, a teratogen.^[93] Others such as benzo[a]pyrene diol epoxide and aflatoxin form DNA adducts that induce errors in replication.^[94] Nevertheless, due to their ability to inhibit DNA transcription and replication, other similar toxins are also used in chemotherapy to inhibit rapidly growing cancer cells.^[95]



A covalent adduct between a metabolically activated form of benzo[a]pyrene, the major mutagen in tobacco smoke, and DNA^[82]

Biological functions

DNA usually occurs as linear chromosomes in eukaryotes, and circular chromosomes in prokaryotes. The set of chromosomes in a cell makes up its genome; the human genome has approximately 3 billion base pairs of DNA arranged into 46 chromosomes.^[96] The information carried by DNA is held in the sequence of pieces of DNA called genes. Transmission of genetic information in genes is achieved via complementary base pairing. For example, in transcription, when a cell uses the information in a gene, the DNA sequence is copied into a complementary RNA sequence through the attraction between the DNA and the correct RNA nucleotides. Usually, this RNA copy is then used to make a matching protein sequence in a process called translation, which depends on the same interaction between RNA nucleotides. In an alternative fashion, a cell may copy its genetic information in a process called DNA replication. The details of these functions are covered in other articles; here the focus is on the interactions between DNA and other molecules that mediate the function of the genome.



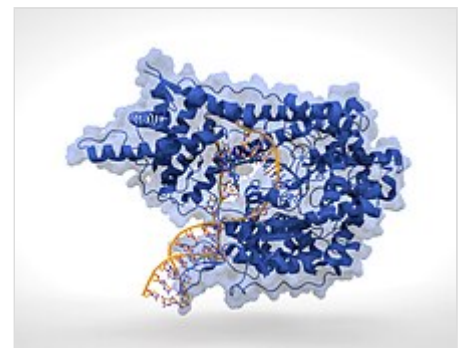
Location of eukaryote nuclear DNA within the chromosomes

Genes and genomes

Genomic DNA is tightly and orderly packed in the process called DNA condensation, to fit the small available volumes of the cell. In eukaryotes, DNA is located in the cell nucleus, with small amounts in mitochondria and chloroplasts. In prokaryotes, the DNA is held within an irregularly shaped body in the cytoplasm called the nucleoid.^[97] The genetic information in a genome is held within genes, and the complete set of this information in an organism is called its genotype. A gene is a unit of heredity and is a region of DNA that influences a particular characteristic in an organism. Genes contain an open reading frame that can be transcribed, and regulatory sequences such as promoters and enhancers, which control transcription of the open reading frame.

In many species, only a small fraction of the total sequence of the genome encodes protein. For example, only about 1.5% of the human genome consists of protein-coding exons, with over 50% of human DNA consisting of non-coding repetitive sequences.^[98] The reasons for the presence of so much noncoding DNA in eukaryotic genomes and the extraordinary differences in genome size, or C-value, among species, represent a long-standing puzzle known as the "C-value enigma".^[99] However, some DNA sequences that do not code protein may still encode functional non-coding RNA molecules, which are involved in the regulation of gene expression.^[100]

Some noncoding DNA sequences play structural roles in chromosomes. Telomeres and centromeres typically contain few genes but are important for the function and stability of chromosomes.^{[64][102]} An abundant form of noncoding DNA in humans are pseudogenes, which are copies of genes that have been disabled by mutation.^[103] These sequences are usually just molecular fossils, although they can occasionally serve as raw genetic material for the creation of new genes through the process of gene duplication and divergence.^[104]



T7 RNA polymerase (blue)
producing an mRNA (green) from a
DNA template (orange)^[101]

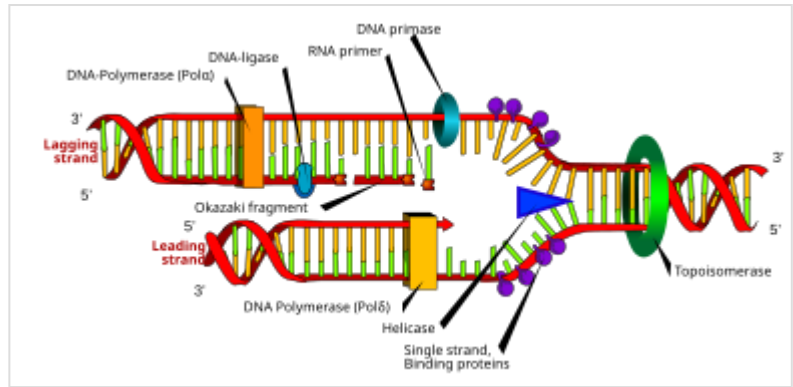
Transcription and translation

A gene is a sequence of DNA that contains genetic information and can influence the phenotype of an organism. Within a gene, the sequence of bases along a DNA strand defines a messenger RNA sequence, which then defines one or more protein sequences. The relationship between the nucleotide sequences of genes and the amino-acid sequences of proteins is determined by the rules of translation, known collectively as the genetic code. The genetic code consists of three-letter 'words' called codons formed from a sequence of three nucleotides (e.g. ACT, CAG, TTT).

In transcription, the codons of a gene are copied into messenger RNA by RNA polymerase. This RNA copy is then decoded by a ribosome that reads the RNA sequence by base-pairing the messenger RNA to transfer RNA, which carries amino acids. Since there are 4 bases in 3-letter combinations, there are 64 possible codons (4^3 combinations). These encode the twenty standard amino acids, giving most amino acids more than one possible codon. There are also three 'stop' or 'nonsense' codons signifying the end of the coding region; these are the TAG, TAA, and TGA codons, (UAG, UAA, and UGA on the mRNA).

Replication

Cell division is essential for an organism to grow, but, when a cell divides, it must replicate the DNA in its genome so that the two daughter cells have the same genetic information as their parent. The double-stranded structure of DNA provides a simple mechanism for DNA replication. Here, the two strands are separated and then each strand's complementary DNA sequence is recreated by an enzyme called DNA polymerase. This enzyme makes the complementary strand by finding the correct base through complementary base pairing and bonding it onto the original strand. As DNA polymerases can only extend a DNA strand in a 5' to 3' direction, different mechanisms are used to copy the antiparallel strands of the double helix.^[105] In this way, the base on the old strand dictates which base appears on the new strand, and the cell ends up with a perfect copy of its DNA.



DNA replication: The double helix is unwound by a helicase and topoisomerase. Next, one DNA polymerase produces the leading strand copy. Another DNA polymerase binds to the lagging strand. This enzyme makes discontinuous segments (called Okazaki fragments) before DNA ligase joins them together.

Extracellular nucleic acids

Naked extracellular DNA (eDNA), most of it released by cell death, is nearly ubiquitous in the environment. Its concentration in soil may be as high as 2 µg/L, and its concentration in natural aquatic environments may be as high at 88 µg/L.^[106] Various possible functions have been proposed for eDNA: it may be involved in horizontal gene transfer,^[107] it may provide nutrients,^[108] and it may act as a buffer to recruit or titrate ions or antibiotics.^[109] Extracellular DNA acts as a functional extracellular matrix component in the biofilms of several bacterial species. It may act as a recognition factor to regulate the attachment and dispersal of specific cell types in the biofilm;^[110] it may contribute to biofilm formation;^[111] and it may contribute to the biofilm's physical strength and resistance to biological stress.^[112]

Cell-free fetal DNA is found in the blood of the mother, and can be sequenced to determine a great deal of information about the developing fetus.^[113]

Under the name of environmental DNA eDNA has seen increased use in the natural sciences as a survey tool for ecology, monitoring the movements and presence of species in water, air, or on land, and assessing an area's biodiversity.^{[114][115]}

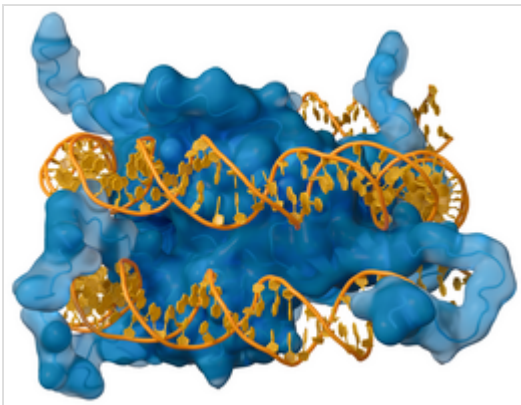
Neutrophil extracellular traps

Neutrophil extracellular traps (NETs) are networks of extracellular fibers, primarily composed of DNA, which allow neutrophils, a type of white blood cell, to kill extracellular pathogens while minimizing damage to the host cells.

Interactions with proteins

All the functions of DNA depend on interactions with proteins. These protein interactions can be non-specific, or the protein can bind specifically to a single DNA sequence. Enzymes can also bind to DNA and of these, the polymerases that copy the DNA base sequence in transcription and DNA replication are particularly important.

DNA-binding proteins



Interaction of DNA (in orange) with histones (in blue). These proteins' basic amino acids bind to the acidic phosphate groups on DNA.

Structural proteins that bind DNA are well-understood examples of non-specific DNA-protein interactions. Within chromosomes, DNA is held in complexes with structural proteins. These proteins organize the DNA into a compact structure called chromatin. In eukaryotes, this structure involves DNA binding to a complex of small basic proteins called histones, while in prokaryotes multiple types of proteins are involved.^{[116][117]} The histones form a disk-shaped complex called a nucleosome, which contains two complete turns of double-stranded DNA wrapped around its surface. These non-specific interactions are formed through basic residues in the histones, making ionic bonds to the acidic sugar-phosphate backbone of the DNA, and are thus largely independent of the base sequence.^[118] Chemical modifications of these basic amino acid residues include

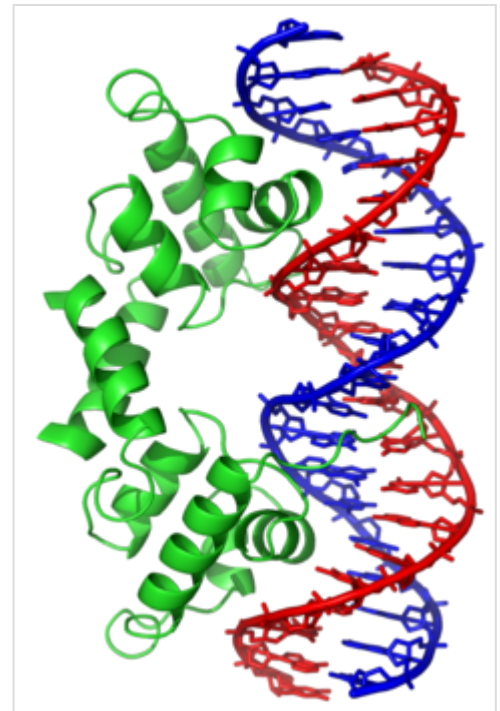
methylation, phosphorylation, and acetylation.^[119] These chemical changes alter the strength of the interaction between the DNA and the histones, making the DNA more or less accessible to transcription factors and changing the rate of transcription.^[120] Other non-specific DNA-binding proteins in chromatin include the high-mobility group proteins, which bind to bent or distorted DNA.^[121] These proteins are important in bending arrays of nucleosomes and arranging them into the larger structures that make up chromosomes.^[122]

A distinct group of DNA-binding proteins is the DNA-binding proteins that specifically bind single-stranded DNA. In humans, replication protein A is the best-understood member of this family and is used in processes where the double helix is separated, including DNA replication, recombination, and DNA repair.^[123] These binding proteins seem to stabilize single-stranded DNA and protect it from forming stem-loops or being degraded by nucleases.

In contrast, other proteins have evolved to bind to particular DNA sequences. The most intensively studied of these are the various transcription factors, which are proteins that regulate transcription. Each transcription factor binds to one particular set of DNA sequences and activates or inhibits the transcription of genes that have these sequences close to their promoters. The transcription factors do this in two ways. Firstly, they can bind the RNA polymerase responsible for transcription, either directly or

through other mediator proteins; this locates the polymerase at the promoter and allows it to begin transcription.^[125] Alternatively, transcription factors can bind enzymes that modify the histones at the promoter. This changes the accessibility of the DNA template to the polymerase.^[126]

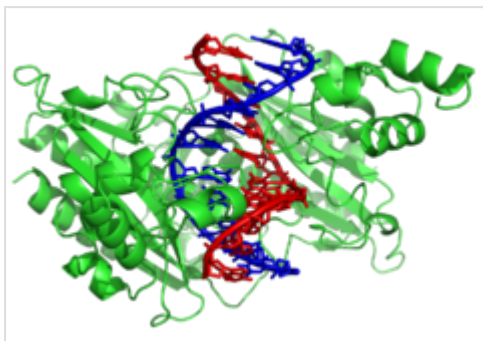
As these DNA targets can occur throughout an organism's genome, changes in the activity of one type of transcription factor can affect thousands of genes.^[127] Consequently, these proteins are often the targets of the signal transduction processes that control responses to environmental changes or cellular differentiation and development. The specificity of these transcription factors' interactions with DNA come from the proteins making multiple contacts to the edges of the DNA bases, allowing them to "read" the DNA sequence. Most of these base-interactions are made in the major groove, where the bases are most accessible.^[25]



The lambda repressor helix-turn-helix transcription factor bound to its DNA target^[124]

DNA-modifying enzymes

Nucleases and ligases



The restriction enzyme EcoRV (green) in a complex with its substrate DNA^[128]

Nucleases are enzymes that cut DNA strands by catalyzing the hydrolysis of the phosphodiester bonds. Nucleases that hydrolyse nucleotides from the ends of DNA strands are called exonucleases, while endonucleases cut within strands. The most frequently used nucleases in molecular biology are the restriction endonucleases, which cut DNA at specific sequences. For instance, the EcoRV enzyme shown to the left recognizes the 6-base sequence 5'-GATATC-3' and makes a cut at the horizontal line. In nature, these enzymes protect bacteria against phage infection by digesting the phage DNA when it enters the bacterial cell, acting as part of the restriction modification system.^[129] In technology, these sequence-

specific nucleases are used in molecular cloning and DNA fingerprinting.

Enzymes called DNA ligases can rejoin cut or broken DNA strands.^[130] Ligases are particularly important in lagging strand DNA replication, as they join the short segments of DNA produced at the replication fork into a complete copy of the DNA template. They are also used in DNA repair and genetic recombination.^[130]

Topoisomerases and helicases

Topoisomerases are enzymes with both nuclease and ligase activity. These proteins change the amount of supercoiling in DNA. Some of these enzymes work by cutting the DNA helix and allowing one section to rotate, thereby reducing its level of supercoiling; the enzyme then seals the DNA break.^[44] Other types of

these enzymes are capable of cutting one DNA helix and then passing a second strand of DNA through this break, before rejoining the helix.^[131] Topoisomerases are required for many processes involving DNA, such as DNA replication and transcription.^[45]

Helicases are proteins that are a type of molecular motor. They use the chemical energy in nucleoside triphosphates, predominantly adenosine triphosphate (ATP), to break hydrogen bonds between bases and unwind the DNA double helix into single strands.^[132] These enzymes are essential for most processes where enzymes need to access the DNA bases.

Polymerases

Polymerases are enzymes that synthesize polynucleotide chains from nucleoside triphosphates. The sequence of their products is created based on existing polynucleotide chains—which are called *templates*. These enzymes function by repeatedly adding a nucleotide to the 3' hydroxyl group at the end of the growing polynucleotide chain. As a consequence, all polymerases work in a 5' to 3' direction.^[133] In the active site of these enzymes, the incoming nucleoside triphosphate base-pairs to the template: this allows polymerases to accurately synthesize the complementary strand of their template. Polymerases are classified according to the type of template that they use.

In DNA replication, DNA-dependent DNA polymerases make copies of DNA polynucleotide chains. To preserve biological information, it is essential that the sequence of bases in each copy are precisely complementary to the sequence of bases in the template strand. Many DNA polymerases have a proofreading activity. Here, the polymerase recognizes the occasional mistakes in the synthesis reaction by the lack of base pairing between the mismatched nucleotides. If a mismatch is detected, a 3' to 5' exonuclease activity is activated and the incorrect base removed.^[134] In most organisms, DNA polymerases function in a large complex called the replisome that contains multiple accessory subunits, such as the DNA clamp or helicases.^[135]

RNA-dependent DNA polymerases are a specialized class of polymerases that copy the sequence of an RNA strand into DNA. They include reverse transcriptase, which is a viral enzyme involved in the infection of cells by retroviruses, and telomerase, which is required for the replication of telomeres.^{[63][136]} For example, HIV reverse transcriptase is an enzyme for AIDS virus replication.^[136] Telomerase is an unusual polymerase because it contains its own RNA template as part of its structure. It synthesizes telomeres at the ends of chromosomes. Telomeres prevent fusion of the ends of neighboring chromosomes and protect chromosome ends from damage.^[64]

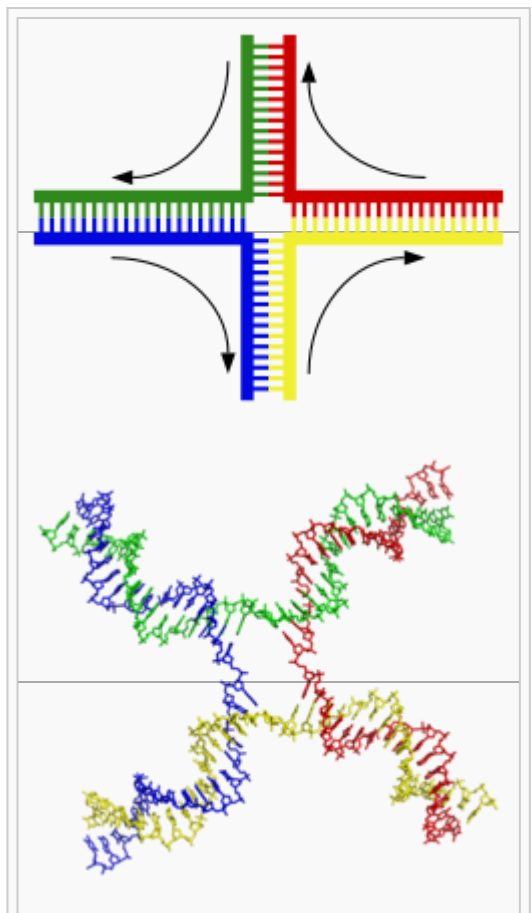
Transcription is carried out by a DNA-dependent RNA polymerase that copies the sequence of a DNA strand into RNA. To begin transcribing a gene, the RNA polymerase binds to a sequence of DNA called a promoter and separates the DNA strands. It then copies the gene sequence into a messenger RNA transcript until it reaches a region of DNA called the terminator, where it halts and detaches from the DNA. As with human DNA-dependent DNA polymerases, RNA polymerase II, the enzyme that transcribes most of the genes in the human genome, operates as part of a large protein complex with multiple regulatory and accessory subunits.^[137]

Genetic recombination

A DNA helix usually does not interact with other segments of DNA, and in human cells, the different chromosomes even occupy separate areas in the nucleus called "chromosome territories".^[139] This physical separation of different chromosomes is important for the ability of DNA to function as a stable repository for information, as one of the few times chromosomes interact is in chromosomal crossover which occurs during sexual reproduction, when genetic recombination occurs. Chromosomal crossover is when two DNA helices break, swap a section and then rejoin.

Recombination allows chromosomes to exchange genetic information and produces new combinations of genes, which increases the efficiency of natural selection and can be important in the rapid evolution of new proteins.^[140] Genetic recombination can also be involved in DNA repair, particularly in the cell's response to double-strand breaks.^[141]

The most common form of chromosomal crossover is homologous recombination, where the two chromosomes involved share very similar sequences. Non-homologous recombination can be damaging to cells, as it can produce chromosomal translocations and genetic abnormalities. The recombination reaction is catalyzed by enzymes known as recombinases, such as RAD51.^[142] The first step in recombination is a double-stranded break caused by either an endonuclease or damage to the DNA.^[143] A series of steps catalyzed in part by the recombinase then leads to joining of the two helices by at least one Holliday junction, in which a segment of a single strand in each helix is annealed to the complementary strand in the other helix. The Holliday junction is a tetrahedral junction structure that can be moved along the pair of chromosomes, swapping one strand for another. The recombination reaction is then halted by cleavage of the junction and re-ligation of the released DNA.^[144] Only strands of like polarity exchange DNA during recombination. There are two types of cleavage: east-west cleavage and north-south cleavage. The north-south cleavage nicks both strands of DNA, while the east-west cleavage has one strand of DNA intact. The formation of a Holliday junction during recombination makes it possible for genetic diversity, genes to exchange on chromosomes, and expression of wild-type viral genomes.



Structure of the Holliday junction intermediate in genetic recombination. The four separate DNA strands are coloured red, blue, green and yellow.^[138]

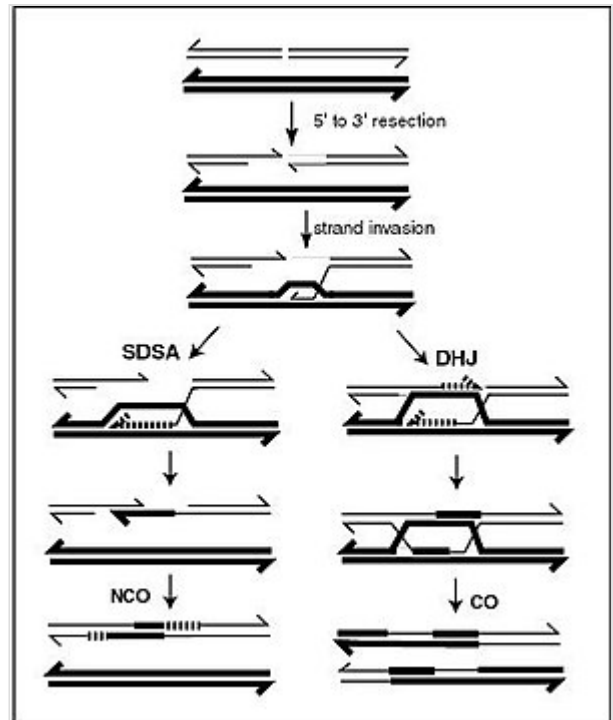
Evolution

DNA contains the genetic information that allows all forms of life to function, grow and reproduce. However, it is unclear how long in the 4-billion-year history of life DNA has performed this function, as it has been proposed that the earliest forms of life may have used RNA as their genetic material.^{[145][146]} RNA may have acted as the central part of early cell metabolism as it can both transmit genetic information and carry out catalysis as part of ribozymes.^[147] This ancient RNA world where nucleic acid would have been used for both catalysis and genetics may have influenced the evolution of the current

genetic code based on four nucleotide bases. This would occur, since the number of different bases in such an organism is a trade-off between a small number of bases increasing replication accuracy and a large number of bases increasing the catalytic efficiency of ribozymes.^[148] However, there is no direct evidence of ancient genetic systems, as recovery of DNA from most fossils is impossible because DNA survives in the environment for less than one million years, and slowly degrades into short fragments in solution.^[149] Claims for older DNA have been made, most notably a report of the isolation of a viable bacterium from a salt crystal 250 million years old,^[150] but these claims are controversial.^{[151][152]}

Building blocks of DNA (adenine, guanine, and related organic molecules) may have been formed extraterrestrially in outer space.^{[153][154][155]} Complex DNA and RNA organic compounds of life, including uracil, cytosine, and thymine, have also been formed in the laboratory under conditions mimicking those found in outer space, using starting chemicals, such as pyrimidine, found in meteorites. Pyrimidine, like polycyclic aromatic hydrocarbons (PAHs), the most carbon-rich chemical found in the universe, may have been formed in red giants or in interstellar cosmic dust and gas clouds.^[156]

Ancient DNA has been recovered from ancient organisms at a timescale where genome evolution can be directly observed, including from extinct organisms up to millions of years old, such as the woolly mammoth.^{[157][158]}



A current model of meiotic recombination, initiated by a double-strand break or gap, followed by pairing with an homologous chromosome and strand invasion to initiate the recombinational repair process. Repair of the gap can lead to crossover (CO) or non-crossover (NCO) of the flanking regions. CO recombination is thought to occur by the Double Holliday Junction (DHJ) model, illustrated on the right, above. NCO recombinants are thought to occur primarily by the Synthesis Dependent Strand Annealing (SDSA) model, illustrated on the left, above. Most recombination events appear to be the SDSA type.

Uses in technology

Genetic engineering

Methods have been developed to purify DNA from organisms, such as phenol-chloroform extraction, and to manipulate it in the laboratory, such as restriction digests and the polymerase chain reaction. Modern biology and biochemistry make intensive use of these techniques in recombinant DNA technology. Recombinant DNA is a man-made DNA sequence that has been assembled from other DNA sequences.

They can be transformed into organisms in the form of plasmids or in the appropriate format, by using a viral vector.^[159] The genetically modified organisms produced can be used to produce products such as recombinant proteins, used in medical research,^[160] or be grown in agriculture.^{[161][162]}

DNA profiling

Forensic scientists can use DNA in blood, semen, skin, saliva or hair found at a crime scene to identify a matching DNA of an individual, such as a perpetrator.^[163] This process is formally termed DNA profiling, also called *DNA fingerprinting*. In DNA profiling, the lengths of variable sections of repetitive DNA, such as short tandem repeats and minisatellites, are compared between people. This method is usually an extremely reliable technique for identifying a matching DNA.^[164] However, identification can be complicated if the scene is contaminated with DNA from several people.^[165] DNA profiling was developed in 1984 by British geneticist Sir Alec Jeffreys,^[166] and first used in forensic science to convict Colin Pitchfork in the 1988 Enderby murders case.^[167]

The development of forensic science and the ability to now obtain genetic matching on minute samples of blood, skin, saliva, or hair has led to re-examining many cases. Evidence can now be uncovered that was scientifically impossible at the time of the original examination. Combined with the removal of the double jeopardy law in some places, this can allow cases to be reopened where prior trials have failed to produce sufficient evidence to convince a jury. People charged with serious crimes may be required to provide a sample of DNA for matching purposes. The most obvious defense to DNA matches obtained forensically is to claim that cross-contamination of evidence has occurred. This has resulted in meticulous strict handling procedures with new cases of serious crime.

DNA profiling is also used successfully to positively identify victims of mass casualty incidents,^[168] bodies or body parts in serious accidents, and individual victims in mass war graves, via matching to family members.

DNA profiling is also used in DNA paternity testing to determine if someone is the biological parent or grandparent of a child with the probability of parentage is typically 99.99% when the alleged parent is biologically related to the child. Normal DNA sequencing methods happen after birth, but there are new methods to test paternity while a mother is still pregnant.^[169]

DNA enzymes or catalytic DNA

Deoxyribozymes, also called DNazymes or catalytic DNA, were first discovered in 1994.^[170] They are mostly single stranded DNA sequences isolated from a large pool of random DNA sequences through a combinatorial approach called in vitro selection or systematic evolution of ligands by exponential enrichment (SELEX). DNazymes catalyze variety of chemical reactions including RNA-DNA cleavage, RNA-DNA ligation, amino acids phosphorylation-dephosphorylation, carbon-carbon bond formation, etc. DNazymes can enhance catalytic rate of chemical reactions up to 100,000,000,000-fold over the uncatalyzed reaction.^[171] The most extensively studied class of DNazymes is RNA-cleaving types which have been used to detect different metal ions and designing therapeutic agents. Several metal-specific DNazymes have been reported including the GR-5 DNzyme (lead-specific),^[170] the CA1-3

DNAzymes (copper-specific),^[172] the 39E DNAzyme (uranyl-specific) and the NaA43 DNAzyme (sodium-specific).^[173] The NaA43 DNAzyme, which is reported to be more than 10,000-fold selective for sodium over other metal ions, was used to make a real-time sodium sensor in cells.

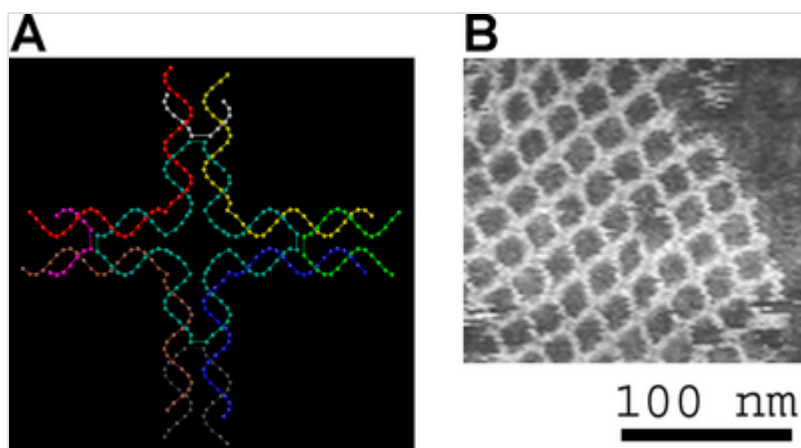
Bioinformatics

Bioinformatics involves the development of techniques to store, data mine, search and manipulate biological data, including DNA nucleic acid sequence data. These have led to widely applied advances in computer science, especially string searching algorithms, machine learning, and database theory.^[174] String searching or matching algorithms, which find an occurrence of a sequence of letters inside a larger sequence of letters, were developed to search for specific sequences of nucleotides.^[175] The DNA sequence may be aligned with other DNA sequences to identify homologous sequences and locate the specific mutations that make them distinct. These techniques, especially multiple sequence alignment, are used in studying phylogenetic relationships and protein function.^[176] Data sets representing entire genomes' worth of DNA sequences, such as those produced by the Human Genome Project, are difficult to use without the annotations that identify the locations of genes and regulatory elements on each chromosome. Regions of DNA sequence that have the characteristic patterns associated with protein- or RNA-coding genes can be identified by gene finding algorithms, which allow researchers to predict the presence of particular gene products and their possible functions in an organism even before they have been isolated experimentally.^[177] Entire genomes may also be compared, which can shed light on the evolutionary history of particular organism and permit the examination of complex evolutionary events.

DNA nanotechnology

DNA nanotechnology uses the unique molecular recognition properties of DNA and other nucleic acids to create self-assembling branched DNA complexes with useful properties.^[179] DNA is thus used as a structural material rather than as a carrier of biological information. This has led to the creation of two-dimensional periodic lattices (both tile-based and using the DNA origami method) and three-dimensional structures in the shapes of polyhedra.^[180] Nanomechanical devices and algorithmic self-assembly have also been demonstrated,^[181] and these DNA structures have been used to

template the arrangement of other molecules such as gold nanoparticles and streptavidin proteins.^[182] DNA and other nucleic acids are the basis of aptamers, synthetic oligonucleotide ligands for specific target molecules used in a range of biotechnology and biomedical applications.^[183]



The DNA structure at left (schematic shown) will self-assemble into the structure visualized by atomic force microscopy at right. DNA nanotechnology is the field that seeks to design nanoscale structures using the molecular recognition properties of DNA molecules.^[178]

History and anthropology

Because DNA collects mutations over time, which are then inherited, it contains historical information, and, by comparing DNA sequences, geneticists can infer the evolutionary history of organisms, their phylogeny.^[184] This field of phylogenetics is a powerful tool in evolutionary biology. If DNA sequences within a species are compared, population geneticists can learn the history of particular populations. This can be used in studies ranging from ecological genetics to anthropology.

Information storage

DNA as a storage device for information has enormous potential since it has much higher storage density compared to electronic devices. However, high costs, slow read and write times (memory latency), and insufficient reliability has prevented its practical use.^{[185][186]}

History

DNA was first isolated by the Swiss physician Friedrich Miescher who, in 1869, discovered a microscopic substance in the pus of discarded surgical bandages. As it resided in the nuclei of cells, he called it "nuclein".^{[187][188]} In 1878, Albrecht Kossel isolated the non-protein component of "nuclein", nucleic acid, and later isolated its five primary nucleobases.^{[189][190]}

In 1909, Phoebus Levene identified the base, sugar, and phosphate nucleotide unit of RNA (then named "yeast nucleic acid").^{[191][192][193]} In 1929, Levene identified deoxyribose sugar in "thymus nucleic acid" (DNA).^[194] Levene suggested that DNA consisted of a string of four nucleotide units linked together through the phosphate groups ("tetranucleotide hypothesis"). Levene thought the chain was short and the bases repeated in a fixed order. In 1927, Nikolai Koltsov proposed that inherited traits would be inherited via a "giant hereditary molecule" made up of "two mirror strands that would replicate in a semi-conservative fashion using each strand as a template".^{[195][196]} In 1928, Frederick Griffith in his experiment discovered that traits of the "smooth" form of *Pneumococcus* could be transferred to the "rough" form of the same bacteria by mixing killed "smooth" bacteria with the live "rough" form.^{[197][198]} This system provided the first clear suggestion that DNA carries genetic information.

In 1933, while studying virgin sea urchin eggs, Jean Brachet suggested that DNA is found in the cell nucleus and that RNA is present exclusively in the cytoplasm. At the time, "yeast nucleic acid" (RNA) was thought to occur only in plants, while "thymus nucleic acid" (DNA) only in animals. The latter was thought to be a tetramer, with the function of buffering cellular pH.^{[199][200]}

In 1937, William Astbury produced the first X-ray diffraction patterns that showed that DNA had a regular structure.^[201]



Maclyn McCarty (left) shakes hands with Francis Crick and James Watson, co-originators of the double-helix model based on the X-ray diffraction data and insights of Rosalind Franklin and Raymond Gosling.

In 1943, Oswald Avery, along with co-workers Colin MacLeod and Maclyn McCarty, identified DNA as the transforming principle, supporting Griffith's suggestion (Avery–MacLeod–McCarty experiment).^[202] Erwin Chargaff developed and published observations now known as Chargaff's rules, stating that in DNA from any species of any organism, the amount of guanine should be equal to cytosine and the amount of adenine should be equal to thymine.^{[203][204]}



A blue plaque outside The Eagle pub in Cambridge, England commemorating Crick and Watson

Late in 1951, Francis Crick started working with James Watson at the Cavendish Laboratory within the University of Cambridge. DNA's role in heredity was confirmed in 1952 when Alfred Hershey and Martha Chase in the Hershey–Chase experiment showed that DNA is the genetic material of the enterobacteria phage T2.^[205]

In May 1952, Raymond Gosling, a graduate student working under the supervision of Rosalind Franklin, took an X-ray diffraction image, labeled as "Photo 51",^[206] at high hydration levels of DNA.

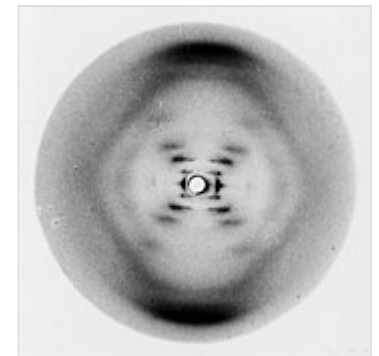
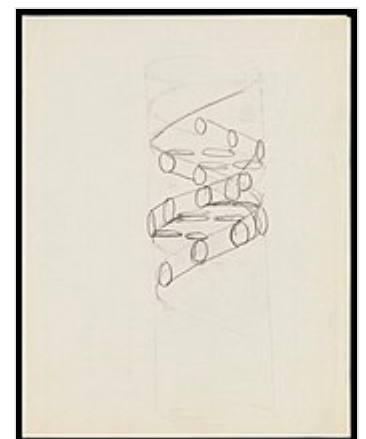


Photo 51, showing X-ray diffraction pattern of DNA

This photo was given to Watson and Crick by Maurice Wilkins and was critical to their obtaining the correct structure of DNA. Franklin told Crick and Watson that the backbones had to be on the outside. Before then, Linus Pauling, and Watson and Crick, had erroneous models with the chains inside and the bases pointing outwards. Franklin's identification of the space group for DNA crystals revealed to Crick that the two DNA strands were antiparallel.^[207] In February 1953, Linus Pauling and Robert Corey proposed a model for nucleic acids containing three intertwined chains, with the phosphates near the axis, and the bases on the outside.^[208] Watson and Crick completed their model, which is now accepted as the first correct model of the double helix of DNA. On 28 February 1953 Crick interrupted patrons' lunchtime at The Eagle pub in Cambridge, England to announce that he and Watson had "discovered the secret of life".^[209]

The 25 April 1953 issue of the journal Nature published a series of five articles giving the Watson and Crick double-helix structure DNA and evidence supporting it.^[210] The structure was reported in a letter titled "MOLECULAR STRUCTURE OF NUCLEIC ACIDS A Structure for Deoxyribose Nucleic Acid", in which they said, "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."^[9] This letter was followed by a letter from Franklin and Gosling, which was the first publication of their own X-ray diffraction data and of their original analysis method.^{[48][211]} Then followed a letter by Wilkins and two of his colleagues, which contained an analysis of *in vivo* B-DNA X-ray patterns, and which supported the presence *in vivo* of the Watson and Crick structure.^[49]



Pencil sketch of the DNA double helix by Francis Crick in 1953

In April 2023, scientists, based on new evidence, concluded that Rosalind Franklin was a contributor and "equal player" in the discovery process of DNA, rather than otherwise, as may have been presented subsequently after the time of the discovery.^{[212][213][214]}

In 1962, after Franklin's death, Watson, Crick, and Wilkins jointly received the Nobel Prize in Physiology or Medicine.^[215] Nobel Prizes are awarded only to living recipients. A debate continues about who should receive credit for the discovery.^[216]

In an influential presentation in 1957, Crick laid out the central dogma of molecular biology, which foretold the relationship between DNA, RNA, and proteins, and articulated the "adaptor hypothesis".^[217] Final confirmation of the replication mechanism that was implied by the double-helical structure followed in 1958 through the Meselson–Stahl experiment.^[218] Further work by Crick and co-workers showed that the genetic code was based on non-overlapping triplets of bases, called codons, allowing Har Gobind Khorana, Robert W. Holley, and Marshall Warren Nirenberg to decipher the genetic code.^[219] These findings represent the birth of molecular biology.^[220]

In 1986, DNA analysis was first used for criminal investigative purposes when police in the UK requested Alec Jeffreys of the University of Leicester to verify or disprove a suspect's rape-murder "confession". In this particular case, the suspect had confessed to two rape-murders, but had later retracted his confession. DNA testing at the university labs soon disproved the veracity of the suspect's original "confession", and the suspect was exonerated from the murder-rape charges.^[221]

See also

- Autosome – Any chromosome other than a sex chromosome
- Crystallography – Scientific study of crystal structures
- DNA Day – Holiday celebrated on April 25
- DNA microarray – Collection of microscopic DNA spots attached to a solid surface
- DNA sequencing – Process of determining the nucleic acid sequence
- Genetic disorder – Health problem caused by one or more abnormalities in the genome
- Genetic genealogy – DNA testing to infer relationships
- Haplotype – Group of genes from one parent
- Meiosis – Cell division producing haploid gametes
- Nucleic acid notation – Universal notation using the Roman characters A, C, G, and T to call the four DNA nucleotides
- Nucleic acid sequence – Succession of nucleotides in a nucleic acid
- *Proteopedia Forms of DNA* (http://www.proteopedia.org/wiki/index.php/Forms_of_DNA)
- Ribosomal DNA – Genes coding for ribosomal RNA
- Southern blot – DNA analysis technique
- X-ray scattering techniques – family of non-destructive analytical techniques
- Xeno nucleic acid – Synthetic nucleic acid analogues

References

1. "deoxyribonucleic acid" (<https://www.merriam-webster.com/dictionary/deoxyribonucleic%20acid>). *Merriam-Webster.com Dictionary*. Merriam-Webster.

2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2014). *Molecular Biology of the Cell* (<http://www.garlandscience.com/product/isbn/9780815344322>) (6th ed.). Garland. p. Chapter 4: DNA, Chromosomes and Genomes. ISBN 978-0-8153-4432-2. Archived (<http://web.archive.org/web/20140714210549/http://www.garlandscience.com/product/isbn/9780815344322>) from the original on 14 July 2014.
3. Purcell A. "DNA" (<http://basicbiology.net/micro/genetics/dna>). *Basic Biology*. Archived (<http://web.archive.org/web/20170105045651/http://basicbiology.net/micro/genetics/dna/>) from the original on 5 January 2017.
4. "Uracil" (<https://www.genome.gov/genetics-glossary/Uracil>). *Genome.gov*. Retrieved 21 November 2019.
5. Russell P (2001). *iGenetics* (https://archive.org/details/igenetics0000russ_v6o1). New York: Benjamin Cummings. ISBN 0-8053-4553-1.
6. Saenger W (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag. ISBN 0-387-90762-9.
7. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Peter W (2002). *Molecular Biology of the Cell* (<https://www.ncbi.nlm.nih.gov/books/NBK21054/>) (Fourth ed.). New York and London: Garland Science. ISBN 0-8153-3218-1. OCLC 145080076 (<https://search.worldcat.org/oclc/145080076>). Archived (<https://web.archive.org/web/20161101022040/https://www.ncbi.nlm.nih.gov/books/NBK21054/>) from the original on 1 November 2016.
8. Irobalieva RN, Fogg JM, Catanese DJ, Catanese DJ, Sutthibutpong T, Chen M, Barker AK, Ludtke SJ, Harris SA, Schmid MF, Chiu W, Zechiedrich L (October 2015). "Structural diversity of supercoiled DNA" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4608029>). *Nature Communications*. **6**: 8440. Bibcode:2015NatCo...6.8440I (<https://ui.adsabs.harvard.edu/abs/2015NatCo...6.8440I>). doi:10.1038/ncomms9440 (<https://doi.org/10.1038%2Fncomm9440>). ISSN 2041-1723 (<https://search.worldcat.org/issn/2041-1723>). PMC 4608029 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4608029>). PMID 26455586 (<https://pubmed.ncbi.nlm.nih.gov/26455586>).
9. Watson JD, Crick FH (April 1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (<http://www.nature.com/nature/dna50/watsoncrick.pdf>) (PDF). *Nature*. **171** (4356): 737–38. Bibcode:1953Natur.171..737W (<https://ui.adsabs.harvard.edu/abs/1953Natur.171..737W>). doi:10.1038/171737a0 (<https://doi.org/10.1038%2F171737a0>). ISSN 0028-0836 (<https://search.worldcat.org/issn/0028-0836>). PMID 13054692 (<https://pubmed.ncbi.nlm.nih.gov/13054692>). S2CID 4253007 (<https://api.semanticscholar.org/CorpusID:4253007>). Archived (<https://web.archive.org/web/20070204110320/http://www.nature.com/nature/dna50/watsoncrick.pdf>) (PDF) from the original on 4 February 2007.
10. Mandelkern M, Elias JG, Eden D, Crothers DM (October 1981). "The dimensions of DNA in solution". *Journal of Molecular Biology*. **152** (1): 153–61. doi:10.1016/0022-2836(81)90099-1 (<https://doi.org/10.1016%2F0022-2836%2881%2990099-1>). ISSN 0022-2836 (<https://search.worldcat.org/issn/0022-2836>). PMID 7338906 (<https://pubmed.ncbi.nlm.nih.gov/7338906>).
11. Arrighi, Frances E.; Mandel, Manley; Bergendahl, Janet; Hsu, T. C. (June 1970). "Buoyant densities of DNA of mammals". *Biochemical Genetics*. **4** (3): 367–376. doi:10.1007/BF00485753 (<https://doi.org/10.1007%2FBF00485753>). ISSN 0006-2928 (<https://search.worldcat.org/issn/0006-2928>). PMID 4991030 (<https://pubmed.ncbi.nlm.nih.gov/4991030>). S2CID 27950750 (<https://api.semanticscholar.org/CorpusID:27950750>).
12. Berg J, Tymoczko J, Stryer L (2002). *Biochemistry*. W.H. Freeman and Company. ISBN 0-7167-4955-6.

13. IUPAC-IUB Commission on Biochemical Nomenclature (CBN) (December 1970). "Abbreviations and Symbols for Nucleic Acids, Polynucleotides and their Constituents. Recommendations 1970" (<https://web.archive.org/web/20070205191106/http://www.chem.qmul.ac.uk/iupac/misc/naabb.html>). *The Biochemical Journal*. **120** (3): 449–54. doi:10.1042/bj1200449 (<https://doi.org/10.1042%2Fbj1200449>). ISSN 0306-3283 (<https://search.worldcat.org/issn/0306-3283>). PMC 1179624 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1179624>). PMID 5499957 (<https://pubmed.ncbi.nlm.nih.gov/5499957>). Archived from the original (<http://www.chem.qmul.ac.uk/iupac/misc/naabb.html>) on 5 February 2007.
14. Ghosh A, Bansal M (April 2003). "A glossary of DNA structures from A to Z". *Acta Crystallographica Section D*. **59** (Pt 4): 620–26. Bibcode:2003AcCrD..59..620G (<https://ui.adsabs.harvard.edu/abs/2003AcCrD..59..620G>). doi:10.1107/S0907444903003251 (<https://doi.org/10.1107%2FS0907444903003251>). ISSN 0907-4449 (<https://search.worldcat.org/issn/0907-4449>). PMID 12657780 (<https://pubmed.ncbi.nlm.nih.gov/12657780>).
15. Edwards KJ, Brown DG, Spink N, Skelly JV, Neidle S. "RCSB PDB – 1D65: Molecular structure of the B-DNA dodecamer d(CGCAAATTTGCG)2. An examination of propeller twist and minor-groove water structure at 2.2 Å resolution" (<https://www.rcsb.org/structure/1D65>). *www.rcsb.org*. Retrieved 27 March 2023.
16. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360284>). *Nucleic Acids Research*. **34** (2): 564–74. doi:10.1093/nar/gkj454 (<https://doi.org/10.1093%2Fnar%2Fgkj454>). ISSN 0305-1048 (<https://search.worldcat.org/issn/0305-1048>). PMC 1360284 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360284>). PMID 16449200 (<https://pubmed.ncbi.nlm.nih.gov/16449200>).
17. Tropp BE (2012). *Molecular Biology* (4th ed.). Sudbury, Mass.: Jones and Barlett Learning. ISBN 978-0-7637-8663-2.
18. Carr S (1953). "Watson-Crick Structure of DNA" (https://www.mun.ca/biology/scarr/Watson-Crick_Model.html). Memorial University of Newfoundland. Archived (https://web.archive.org/web/20160719095721/http://www.mun.ca/biology/scarr/Watson-Crick_Model.html) from the original on 19 July 2016. Retrieved 13 July 2016.
19. Verma S, Eckstein F (1998). "Modified oligonucleotides: synthesis and strategy for users" (<https://doi.org/10.1146%2Fannurev.biochem.67.1.99>). *Annual Review of Biochemistry*. **67**: 99–134. doi:10.1146/annurev.biochem.67.1.99 (<https://doi.org/10.1146%2Fannurev.biochem.67.1.99>). ISSN 0066-4154 (<https://search.worldcat.org/issn/0066-4154>). PMID 9759484 (<https://pubmed.ncbi.nlm.nih.gov/9759484>).
20. Johnson TB, Coghill RD (1925). "Pyrimidines. CIII. The discovery of 5-methylcytosine in tuberculinic acid, the nucleic acid of the tubercle bacillus". *Journal of the American Chemical Society*. **47**: 2838–44. doi:10.1021/ja01688a030 (<https://doi.org/10.1021%2Fja01688a030>). ISSN 0002-7863 (<https://search.worldcat.org/issn/0002-7863>).
21. Weigele P, Raleigh EA (October 2016). "Biosynthesis and Function of Modified Bases in Bacteria and Their Viruses" (<https://doi.org/10.1021%2Facs.chemrev.6b00114>). *Chemical Reviews*. **116** (20): 12655–12687. doi:10.1021/acs.chemrev.6b00114 (<https://doi.org/10.1021%2Facs.chemrev.6b00114>). ISSN 0009-2665 (<https://search.worldcat.org/issn/0009-2665>). PMID 27319741 (<https://pubmed.ncbi.nlm.nih.gov/27319741>).
22. Kumar S, Chinnusamy V, Mohapatra T (2018). "Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6305559>). *Frontiers in Genetics*. **9**: 640. doi:10.3389/fgene.2018.00640 (<https://doi.org/10.3389%2Ffgene.2018.00640>). ISSN 1664-8021 (<https://search.worldcat.org/issn/1664-8021>). PMC 6305559 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6305559>). PMID 30619465 (<https://pubmed.ncbi.nlm.nih.gov/30619465>).
23. Carell T, Kurz MQ, Müller M, Rossa M, Spada F (April 2018). "Non-canonical Bases in the Genome: The Regulatory Information Layer in DNA". *Angewandte Chemie*. **57** (16): 4296–4312. doi:10.1002/anie.201708228 (<https://doi.org/10.1002%2Fanie.201708228>). PMID 28941008 (<https://pubmed.ncbi.nlm.nih.gov/28941008>).

24. Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE (October 1980). "Crystal structure analysis of a complete turn of B-DNA". *Nature*. **287** (5784): 755–58. Bibcode:1980Natur.287..755W (<https://ui.adsabs.harvard.edu/abs/1980Natur.287..755W>). doi:10.1038/287755a0 (<https://doi.org/10.1038%2F287755a0>). PMID 7432492 (<https://pubmed.ncbi.nlm.nih.gov/7432492>). S2CID 4315465 (<https://api.semanticscholar.org/CorpusID:4315465>).
25. Pabo CO, Sauer RT (1984). "Protein-DNA recognition". *Annual Review of Biochemistry*. **53**: 293–321. doi:10.1146/annurev.bi.53.070184.001453 (<https://doi.org/10.1146%2Fannurev.bi.53.070184.001453>). PMID 6236744 (<https://pubmed.ncbi.nlm.nih.gov/6236744>).
26. Nikolova EN, Zhou H, Gottardo FL, Alvey HS, Kimsey IJ, Al-Hashimi HM (2013). "A historical account of Hoogsteen base-pairs in duplex DNA" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3844552>). *Biopolymers*. **99** (12): 955–68. doi:10.1002/bip.22334 (<https://doi.org/10.1002%2Fbip.22334>). PMC 3844552 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3844552>). PMID 23818176 (<https://pubmed.ncbi.nlm.nih.gov/23818176>).
27. Clausen-Schaumann H, Rief M, Tolkdorf C, Gaub HE (April 2000). "Mechanical stability of single DNA molecules" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1300792>). *Biophysical Journal*. **78** (4): 1997–2007. Bibcode:2000BpJ....78.1997C (<https://ui.adsabs.harvard.edu/abs/2000BpJ....78.1997C>). doi:10.1016/S0006-3495(00)76747-6 (<https://doi.org/10.1016%2FS0006-3495%2800%2976747-6>). PMC 1300792 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1300792>). PMID 10733978 (<https://pubmed.ncbi.nlm.nih.gov/10733978>).
28. Chalikian TV, Völker J, Plum GE, Breslauer KJ (July 1999). "A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC22151>). *Proceedings of the National Academy of Sciences of the United States of America*. **96** (14): 7853–58. Bibcode:1999PNAS...96.7853C (<https://ui.adsabs.harvard.edu/abs/1999PNAS...96.7853C>). doi:10.1073/pnas.96.14.7853 (<https://doi.org/10.1073%2Fpnas.96.14.7853>). PMC 22151 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC22151>). PMID 10393911 (<https://pubmed.ncbi.nlm.nih.gov/10393911>).
29. deHaseth PL, Helmann JD (June 1995). "Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA". *Molecular Microbiology*. **16** (5): 817–24. doi:10.1111/j.1365-2958.1995.tb02309.x (<https://doi.org/10.1111%2Fj.1365-2958.1995.tb02309.x>). PMID 7476180 (<https://pubmed.ncbi.nlm.nih.gov/7476180>). S2CID 24479358 (<https://api.semanticscholar.org/CorpusID:24479358>).
30. Isaksson J, Acharya S, Barman J, Cheruku P, Chattopadhyaya J (December 2004). "Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern" (<http://www.boc.uu.se/boc14www/thesis/johan2005/Paper%20V/Paper%20V.pdf>) (PDF). *Biochemistry*. **43** (51): 15996–6010. doi:10.1021/bi048221v (<https://doi.org/10.1021%2Fbi048221v>). PMID 15609994 (<https://pubmed.ncbi.nlm.nih.gov/15609994>). Archived (<https://web.archive.org/web/20070610205112/http://www.boc.uu.se/boc14www/thesis/johan2005/Paper%20V/Paper%20V.pdf>) (PDF) from the original on 10 June 2007.
31. Piovesan A, Pelleri MC, Antonaros F, Strippoli P, Caracausi M, Vitale L (2019). "On the length, weight and GC content of the human genome" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6391780>). *BMC Res Notes*. **12** (1): 106. doi:10.1186/s13104-019-4137-z (<https://doi.org/10.1186%2Fs13104-019-4137-z>). PMC 6391780 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6391780>). PMID 30813969 (<https://pubmed.ncbi.nlm.nih.gov/30813969>).
32. Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. (May 2006). "The DNA sequence and biological annotation of human chromosome 1" (<https://doi.org/10.1038%2Fnature04727>). *Nature*. **441** (7091): 315–21. Bibcode:2006Natur.441..315G (<https://ui.adsabs.harvard.edu/abs/2006Natur.441..315G>). doi:10.1038/nature04727 (<https://doi.org/10.1038%2Fnature04727>). PMID 16710414 (<https://pubmed.ncbi.nlm.nih.gov/16710414>).

33. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. (April 1981). "Sequence and organization of the human mitochondrial genome". *Nature*. **290** (5806): 457–465. Bibcode:1981Natur.290..457A (<https://ui.adsabs.harvard.edu/abs/1981Natur.290..457A>). doi:10.1038/290457a0 (<https://doi.org/10.1038%2F290457a0>). PMID 7219534 (<https://pubmed.ncbi.nlm.nih.gov/7219534>). S2CID 4355527 (<https://api.semanticscholar.org/CorpusID:4355527>).
34. "Untitled" (<https://web.archive.org/web/20110813123936/http://chemistry.umeche.maine.edu/CHY431/MitoDNA.html>). Archived from the original (<http://chemistry.umeche.maine.edu/CHY431/MitoDNA.html>) on 13 August 2011. Retrieved 13 June 2012.
35. Satoh M, Kuroiwa T (September 1991). "Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell". *Experimental Cell Research*. **196** (1): 137–140. doi:10.1016/0014-4827(91)90467-9 (<https://doi.org/10.1016%2F0014-4827%2891%2990467-9>). PMID 1715276 (<https://pubmed.ncbi.nlm.nih.gov/1715276>).
36. Zhang D, Keilty D, Zhang ZF, Chian RC (March 2017). "Mitochondria in oocyte aging: current understanding" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5506767>). *Facts, Views & Vision in ObGyn*. **9** (1): 29–38. PMC 5506767 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5506767>). PMID 28721182 (<https://pubmed.ncbi.nlm.nih.gov/28721182>).
37. Designation of the two strands of DNA (<http://www.chem.qmul.ac.uk/iubmb/newsletter/misc/DNA.html>) Archived (<https://web.archive.org/web/20080424015915/http://www.chem.qmul.ac.uk/iubmb/newsletter/misc/DNA.html>) 24 April 2008 at the *Wayback Machine* JCBN/NC-IUB Newsletter 1989. Retrieved 7 May 2008
38. Hüttenhofer A, Schattner P, Polacek N (May 2005). "Non-coding RNAs: hope or hype?". *Trends in Genetics*. **21** (5): 289–97. doi:10.1016/j.tig.2005.03.007 (<https://doi.org/10.1016%2Fj.tig.2005.03.007>). PMID 15851066 (<https://pubmed.ncbi.nlm.nih.gov/15851066>).
39. Munroe SH (November 2004). "Diversity of antisense regulation in eukaryotes: multiple mechanisms, emerging patterns". *Journal of Cellular Biochemistry*. **93** (4): 664–71. doi:10.1002/jcb.20252 (<https://doi.org/10.1002%2Fjcb.20252>). PMID 15389973 (<https://pubmed.ncbi.nlm.nih.gov/15389973>). S2CID 23748148 (<https://api.semanticscholar.org/CorpusID:23748148>).
40. Makalowska I, Lin CF, Makalowski W (February 2005). "Overlapping genes in vertebrate genomes". *Computational Biology and Chemistry*. **29** (1): 1–12. doi:10.1016/j.compbiolchem.2004.12.006 (<https://doi.org/10.1016%2Fj.compbiolchem.2004.12.006>). PMID 15680581 (<https://pubmed.ncbi.nlm.nih.gov/15680581>).
41. Johnson ZI, Chisholm SW (November 2004). "Properties of overlapping genes are conserved across microbial genomes" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC525685>). *Genome Research*. **14** (11): 2268–72. doi:10.1101/gr.2433104 (<https://doi.org/10.1101%2Fgr.2433104>). PMC 525685 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC525685>). PMID 15520290 (<https://pubmed.ncbi.nlm.nih.gov/15520290>).
42. Lamb RA, Horvath CM (August 1991). "Diversity of coding strategies in influenza viruses" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7173306>). *Trends in Genetics*. **7** (8): 261–66. doi:10.1016/0168-9525(91)90326-L (<https://doi.org/10.1016%2F0168-9525%2891%2990326-L>). PMC 7173306 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7173306>). PMID 1771674 (<https://pubmed.ncbi.nlm.nih.gov/1771674>).
43. Benham CJ, Mielke SP (2005). "DNA mechanics" (<https://web.archive.org/web/20190301225243/http://pdfs.semanticscholar.org/ab63/d57290ebf9bc3536fd3f2257a2b509076fc1.pdf>) (PDF). *Annual Review of Biomedical Engineering*. **7**: 21–53. doi:10.1146/annurev.bioeng.6.062403.132016 (<https://doi.org/10.1146%2Fannurev.bioeng.6.062403.132016>). PMID 16004565 (<https://pubmed.ncbi.nlm.nih.gov/16004565>). S2CID 1427671 (<https://api.semanticscholar.org/CorpusID:1427671>). Archived from the original (<http://pdfs.semanticscholar.org/ab63/d57290ebf9bc3536fd3f2257a2b509076fc1.pdf>) (PDF) on 1 March 2019.

44. Champoux JJ (2001). "DNA topoisomerases: structure, function, and mechanism". *Annual Review of Biochemistry*. **70**: 369–413. doi:10.1146/annurev.biochem.70.1.369 (<https://doi.org/10.1146/annurev.biochem.70.1.369>). PMID 11395412 (<https://pubmed.ncbi.nlm.nih.gov/11395412>). S2CID 18144189 (<https://api.semanticscholar.org/CorpusID:18144189>).
45. Wang JC (June 2002). "Cellular roles of DNA topoisomerases: a molecular perspective". *Nature Reviews Molecular Cell Biology*. **3** (6): 430–40. doi:10.1038/nrm831 (<https://doi.org/10.1038/nrm831>). PMID 12042765 (<https://pubmed.ncbi.nlm.nih.gov/12042765>). S2CID 205496065 (<https://api.semanticscholar.org/CorpusID:205496065>).
46. Basu HS, Feuerstein BG, Zarling DA, Shafer RH, Marton LJ (October 1988). "Recognition of Z-RNA and Z-DNA determinants by polyamines in solution: experimental and theoretical studies". *Journal of Biomolecular Structure & Dynamics*. **6** (2): 299–309. doi:10.1080/07391102.1988.10507714 (<https://doi.org/10.1080/07391102.1988.10507714>). PMID 2482766 (<https://pubmed.ncbi.nlm.nih.gov/2482766>).
47.
 - Franklin RE, Gosling RG (6 March 1953). "The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content" (<http://journals.iucr.org/q/issues/1953/08-09/00/a00979/a00979.pdf>) (PDF). *Acta Crystallogr.* **6** (8–9): 673–77. Bibcode:1953AcCry...6..673F (<https://ui.adsabs.harvard.edu/abs/1953AcCry...6..673F>). doi:10.1107/S0365110X53001939 (<https://doi.org/10.1107/S0365110X53001939>). Archived (<https://web.archive.org/web/20160109043915/http://journals.iucr.org/q/issues/1953/08-09/00/a00979/a00979.pdf>) (PDF) from the original on 9 January 2016.
 - Franklin RE, Gosling RG (1953). "The structure of sodium thymonucleate fibres. II. The cylindrically symmetrical Patterson function" (<http://journals.iucr.org/q/issues/1953/08-09/00/a00980/a00980.pdf>) (PDF). *Acta Crystallogr.* **6** (8–9): 678–85. Bibcode:1953AcCry...6..678F (<https://ui.adsabs.harvard.edu/abs/1953AcCry...6..678F>). doi:10.1107/S0365110X53001940 (<https://doi.org/10.1107/S0365110X53001940>). Archived (<https://web.archive.org/web/20170629084321/http://journals.iucr.org/q/issues/1953/08-09/00/a00980/a00980.pdf>) (PDF) from the original on 29 June 2017.
48. Franklin RE, Gosling RG (April 1953). "Molecular configuration in sodium thymonucleate" (<http://www.nature.com/nature/dna50/franklingosling.pdf>) (PDF). *Nature*. **171** (4356): 740–41. Bibcode:1953Natur.171..740F (<https://ui.adsabs.harvard.edu/abs/1953Natur.171..740F>). doi:10.1038/171740a0 (<https://doi.org/10.1038/171740a0>). PMID 13054694 (<https://pubmed.ncbi.nlm.nih.gov/13054694>). S2CID 4268222 (<https://api.semanticscholar.org/CorpusID:4268222>). Archived (<https://web.archive.org/web/20110103160712/http://www.nature.com/nature/dna50/franklingosling.pdf>) (PDF) from the original on 3 January 2011.
49. Wilkins MH, Stokes AR, Wilson HR (April 1953). "Molecular structure of deoxypentose nucleic acids" (<http://www.nature.com/nature/dna50/wilkins.pdf>) (PDF). *Nature*. **171** (4356): 738–40. Bibcode:1953Natur.171..738W (<https://ui.adsabs.harvard.edu/abs/1953Natur.171..738W>). doi:10.1038/171738a0 (<https://doi.org/10.1038/171738a0>). PMID 13054693 (<https://pubmed.ncbi.nlm.nih.gov/13054693>). S2CID 4280080 (<https://api.semanticscholar.org/CorpusID:4280080>). Archived (<https://web.archive.org/web/20110513234223/http://www.nature.com/nature/dna50/wilkins.pdf>) (PDF) from the original on 13 May 2011.
50. Leslie AG, Arnott S, Chandrasekaran R, Ratliff RL (October 1980). "Polymorphism of DNA double helices". *Journal of Molecular Biology*. **143** (1): 49–72. doi:10.1016/0022-2836(80)90124-2 ([https://doi.org/10.1016/0022-2836\(80\)90124-2](https://doi.org/10.1016/0022-2836(80)90124-2)). PMID 7441761 (<https://pubmed.ncbi.nlm.nih.gov/7441761>).
51. Baianu IC (1980). "Structural Order and Partial Disorder in Biological systems" (<http://cogprints.org/3822/>). *Bull. Math. Biol.* **42** (4): 137–41. doi:10.1007/BF02462372 (<https://doi.org/10.1007/BF02462372>). S2CID 189888972 (<https://api.semanticscholar.org/CorpusID:189888972>).
52. Hosemann R, Bagchi RN (1962). *Direct analysis of diffraction by matter*. Amsterdam – New York: North-Holland Publishers.

53. Baianu IC (1978). "X-ray scattering by partially disordered membrane systems" (<https://web.archive.org/web/20200314050140/http://journals.iucr.org/a/issues/1978/05/00/a15615/a15615.pdf>) (PDF). *Acta Crystallogr A*. **34** (5): 751–53. Bibcode:1978AcCrA..34..751B (<https://ui.adsabs.harvard.edu/abs/1978AcCrA..34..751B>). doi:10.1107/S0567739478001540 (<https://doi.org/10.1107%2FS0567739478001540>). Archived from the original (<http://journals.iucr.org/a/issues/1978/05/00/a15615/a15615.pdf>) (PDF) on 14 March 2020. Retrieved 29 August 2019.
54. Wahl MC, Sundaralingam M (1997). "Crystal structures of A-DNA duplexes". *Biopolymers*. **44** (1): 45–63. doi:10.1002/(SICI)1097-0282(1997)44:1<45::AID-BIP4>3.0.CO;2-# (<https://doi.org/10.1002%2F%28SICI%291097-0282%281997%2944%3A1%3C45%3A%3AAID-BIP4%3E3.0.CO%3B2-%23>). PMID 9097733 (<https://pubmed.ncbi.nlm.nih.gov/9097733>).
55. Lu XJ, Shakked Z, Olson WK (July 2000). "A-form conformational motifs in ligand-bound DNA structures". *Journal of Molecular Biology*. **300** (4): 819–40. doi:10.1006/jmbi.2000.3690 (<https://doi.org/10.1006%2Fjmbi.2000.3690>). PMID 10891271 (<https://pubmed.ncbi.nlm.nih.gov/10891271>).
56. Rothenburg S, Koch-Nolte F, Haag F (December 2001). "DNA methylation and Z-DNA formation as mediators of quantitative differences in the expression of alleles". *Immunological Reviews*. **184**: 286–98. doi:10.1034/j.1600-065x.2001.1840125.x (<https://doi.org/10.1034%2Fj.1600-065x.2001.1840125.x>). PMID 12086319 (<https://pubmed.ncbi.nlm.nih.gov/12086319>). S2CID 20589136 (<https://api.semanticscholar.org/CorpusID:20589136>).
57. Oh DB, Kim YG, Rich A (December 2002). "Z-DNA-binding proteins can act as potent effectors of gene expression in vivo" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC139201>). *Proceedings of the National Academy of Sciences of the United States of America*. **99** (26): 16666–71. Bibcode:2002PNAS...9916666O (<https://ui.adsabs.harvard.edu/abs/2002PNAS...9916666O>). doi:10.1073/pnas.262672699 (<https://doi.org/10.1073%2Fpnas.262672699>). PMC 139201 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC139201>). PMID 12486233 (<https://pubmed.ncbi.nlm.nih.gov/12486233>).
58. Palmer J (2 December 2010). "Arsenic-loving bacteria may help in hunt for alien life" (<http://www.bbc.co.uk/news/science-environment-11886943>). *BBC News*. Archived (<https://web.archive.org/web/20101203045804/http://www.bbc.co.uk/news/science-environment-11886943>) from the original on 3 December 2010. Retrieved 2 December 2010.
59. Bortman H (2 December 2010). "Arsenic-Eating Bacteria Opens New Possibilities for Alien Life" (<http://www.space.com/scienceastronomy/arsenic-bacteria-alien-life-101202.html>). *Space.com*. Archived (<https://web.archive.org/web/20101204235915/http://www.space.com/scienceastronomy/arsenic-bacteria-alien-life-101202.html>) from the original on 4 December 2010. Retrieved 2 December 2010.
60. Katsnelson A (2 December 2010). "Arsenic-eating microbe may redefine chemistry of life" (<http://www.nature.com/news/2010/101202/full/news.2010.645.html>). *Nature News*. doi:10.1038/news.2010.645 (<https://doi.org/10.1038%2Fnews.2010.645>). Archived (<https://web.archive.org/web/20120212155007/http://www.nature.com/news/2010/101202/full/news.2010.645.html>) from the original on 12 February 2012.
61. Cressey D (3 October 2012). "'Arsenic-life' Bacterium Prefers Phosphorus after all". *Nature News*. doi:10.1038/nature.2012.11520 (<https://doi.org/10.1038%2Fnature.2012.11520>). S2CID 87341731 (<https://api.semanticscholar.org/CorpusID:87341731>).
62. "Structure and packing of human telomeric DNA" (<http://ndbserver.rutgers.edu/service/ndb/atlas/summary?searchTarget=UD0017>). *ndbserver.rutgers.edu*. Retrieved 18 May 2023.
63. Greider CW, Blackburn EH (December 1985). "Identification of a specific telomere terminal transferase activity in Tetrahymena extracts" (<https://doi.org/10.1016%2F0092-8674%2885%2990170-9>). *Cell*. **43** (2 Pt 1): 405–13. doi:10.1016/0092-8674(85)90170-9 (<https://doi.org/10.1016%2F0092-8674%2885%2990170-9>). PMID 3907856 (<https://pubmed.ncbi.nlm.nih.gov/3907856>).

64. Nugent CI, Lundblad V (April 1998). "The telomerase reverse transcriptase: components and regulation" (<https://doi.org/10.1101%2Fgad.12.8.1073>). *Genes & Development*. **12** (8): 1073–85. doi:10.1101/gad.12.8.1073 (<https://doi.org/10.1101%2Fgad.12.8.1073>). PMID 9553037 (<https://pubmed.ncbi.nlm.nih.gov/9553037>).
65. Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW (November 1997). "Normal human chromosomes have long G-rich telomeric overhangs at one end" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC316649>). *Genes & Development*. **11** (21): 2801–09. doi:10.1101/gad.11.21.2801 (<https://doi.org/10.1101%2Fgad.11.21.2801>). PMC 316649 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC316649>). PMID 9353250 (<https://pubmed.ncbi.nlm.nih.gov/9353250>).
66. Burge S, Parkinson GN, Hazel P, Todd AK, Neidle S (2006). "Quadruplex DNA: sequence, topology and structure" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1636468>). *Nucleic Acids Research*. **34** (19): 5402–15. doi:10.1093/nar/gkl655 (<https://doi.org/10.1093%2Fnar%2Fgkl655>). PMC 1636468 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1636468>). PMID 17012276 (<https://pubmed.ncbi.nlm.nih.gov/17012276>).
67. Parkinson GN, Lee MP, Neidle S (June 2002). "Crystal structure of parallel quadruplexes from human telomeric DNA". *Nature*. **417** (6891): 876–80. Bibcode:2002Natur.417..876P (<https://ui.adsabs.harvard.edu/abs/2002Natur.417..876P>). doi:10.1038/nature755 (<https://doi.org/10.1038%2Fnature755>). PMID 12050675 (<https://pubmed.ncbi.nlm.nih.gov/12050675>). S2CID 4422211 (<https://api.semanticscholar.org/CorpusID:4422211>).
68. Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T (May 1999). "Mammalian telomeres end in a large duplex loop". *Cell*. **97** (4): 503–14. CiteSeerX 10.1.1.335.2649 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.335.2649>). doi:10.1016/S0092-8674(00)80760-6 (<https://doi.org/10.1016%2FS0092-8674%2800%2980760-6>). PMID 10338214 (<https://pubmed.ncbi.nlm.nih.gov/10338214>). S2CID 721901 (<https://api.semanticscholar.org/CorpusID:721901>).
69. Seeman NC (November 2005). "DNA enables nanoscale control of the structure of matter" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3478329>). *Quarterly Reviews of Biophysics*. **38** (4): 363–71. doi:10.1017/S0033583505004087 (<https://doi.org/10.1017%2FS0033583505004087>). PMC 3478329 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3478329>). PMID 16515737 (<https://pubmed.ncbi.nlm.nih.gov/16515737>).
70. Warren M (21 February 2019). "Four new DNA letters double life's alphabet" (<https://doi.org/10.1038%2Fd41586-019-00650-8>). *Nature*. **566** (7745): 436. Bibcode:2019Natur.566..436W (<https://ui.adsabs.harvard.edu/abs/2019Natur.566..436W>). doi:10.1038/d41586-019-00650-8 (<https://doi.org/10.1038%2Fd41586-019-00650-8>). PMID 30809059 (<https://pubmed.ncbi.nlm.nih.gov/30809059>).
71. Hoshika S, Leal NA, Kim MJ, Kim MS, Karalkar NB, Kim HJ, et al. (22 February 2019). "Hachimoji DNA and RNA: A genetic system with eight building blocks (paywall)" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6413494>). *Science*. **363** (6429): 884–887. Bibcode:2019Sci...363..884H (<https://ui.adsabs.harvard.edu/abs/2019Sci...363..884H>). doi:10.1126/science.aat0971 (<https://doi.org/10.1126%2Fscience.aat0971>). PMC 6413494 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6413494>). PMID 30792304 (<https://pubmed.ncbi.nlm.nih.gov/30792304>).
72. Burghardt B, Hartmann AK (February 2007). "RNA secondary structure design" (<https://link.aps.org/doi/10.1103/PhysRevE.75.021920>). *Physical Review E*. **75** (2): 021920. arXiv:physics/0609135 (<https://arxiv.org/abs/physics/0609135>). Bibcode:2007PhRvE..75b1920B (<https://ui.adsabs.harvard.edu/abs/2007PhRvE..75b1920B>). doi:10.1103/PhysRevE.75.021920 (<https://doi.org/10.1103%2FPhysRevE.75.021920>). PMID 17358380 (<https://pubmed.ncbi.nlm.nih.gov/17358380>). S2CID 17574854 (<https://api.semanticscholar.org/CorpusID:17574854>).
73. Reusch W. "Nucleic Acids" (<https://www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/nucacids.htm>). Michigan State University. Retrieved 30 June 2022.

74. "How To Extract DNA From Anything Living" (<https://learn.genetics.utah.edu/content/labs/extraction/howto/>). University of Utah. Retrieved 30 June 2022.
75. Hu Q, Rosenfeld MG (2012). "Epigenetic regulation of human embryonic stem cells" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3488762>). *Frontiers in Genetics*. **3**: 238. doi:10.3389/fgene.2012.00238 (<https://doi.org/10.3389%2Ffgene.2012.00238>). PMC 3488762 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3488762>). PMID 23133442 (<https://pubmed.ncbi.nlm.nih.gov/23133442>).
76. Klose RJ, Bird AP (February 2006). "Genomic DNA methylation: the mark and its mediators". *Trends in Biochemical Sciences*. **31** (2): 89–97. doi:10.1016/j.tibs.2005.12.008 (<https://doi.org/10.1016%2Fj.tibs.2005.12.008>). PMID 16403636 (<https://pubmed.ncbi.nlm.nih.gov/16403636>).
77. Bird A (January 2002). "DNA methylation patterns and epigenetic memory" (<https://doi.org/10.1101%2Fgad.947102>). *Genes & Development*. **16** (1): 6–21. doi:10.1101/gad.947102 (<https://doi.org/10.1101%2Fgad.947102>). PMID 11782440 (<https://pubmed.ncbi.nlm.nih.gov/11782440>).
78. Walsh CP, Xu GL (2006). "Cytosine methylation and DNA repair". *DNA Methylation: Basic Mechanisms*. Current Topics in Microbiology and Immunology. Vol. 301. pp. 283–315. doi:10.1007/3-540-31390-7_11 (https://doi.org/10.1007%2F3-540-31390-7_11). ISBN 3-540-29114-8. PMID 16570853 (<https://pubmed.ncbi.nlm.nih.gov/16570853>).
79. Kriaucionis S, Heintz N (May 2009). "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3263819>). *Science*. **324** (5929): 929–30. Bibcode:2009Sci...324..929K (<https://ui.adsabs.harvard.edu/abs/2009Sci...324..929K>). doi:10.1126/science.1169786 (<https://doi.org/10.1126%2Fscience.1169786>). PMC 3263819 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3263819>). PMID 19372393 (<https://pubmed.ncbi.nlm.nih.gov/19372393>).
80. Ratel D, Ravanat JL, Berger F, Wion D (March 2006). "N6-methyladenine: the other methylated base of DNA" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754416>). *BioEssays*. **28** (3): 309–15. doi:10.1002/bies.20342 (<https://doi.org/10.1002%2Fbies.20342>). PMC 2754416 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2754416>). PMID 16479578 (<https://pubmed.ncbi.nlm.nih.gov/16479578>).
81. Gommers-Ampt JH, Van Leeuwen F, de Beer AL, Vliegenthart JF, Dizdaroglu M, Kowalak JA, Crain PF, Borst P (December 1993). "beta-D-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan *T. brucei*". *Cell*. **75** (6): 1129–36. doi:10.1016/0092-8674(93)90322-H (<https://doi.org/10.1016%2F0092-8674%2893%2990322-H>). hdl:1874/5219 (<https://hdl.handle.net/1874%2F5219>). PMID 8261512 (<https://pubmed.ncbi.nlm.nih.gov/8261512>). S2CID 24801094 (<https://api.semanticscholar.org/CorpusID:24801094>).
82. Created from PDB 1JDG (<http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=1JDG>) Archived (<https://web.archive.org/web/20080922150848/http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=1JDG>) 22 September 2008 at the Wayback Machine
83. Douki T, Reynaud-Angelin A, Cadet J, Sage E (August 2003). "Bipyrimidine photoproducts rather than oxidative lesions are the main type of DNA damage involved in the genotoxic effect of solar UVA radiation". *Biochemistry*. **42** (30): 9221–26. doi:10.1021/bi034593c (<https://doi.org/10.1021%2Fbi034593c>). PMID 12885257 (<https://pubmed.ncbi.nlm.nih.gov/12885257>).
84. Cadet J, Delatour T, Douki T, Gasparutto D, Pouget JP, Ravanat JL, Sauvaigo S (March 1999). "Hydroxyl radicals and DNA base damage". *Mutation Research*. **424** (1–2): 9–21. Bibcode:1999MRFMM.424....9C (<https://ui.adsabs.harvard.edu/abs/1999MRFMM.424....9C>). doi:10.1016/S0027-5107(99)00004-4 (<https://doi.org/10.1016%2FS0027-5107%2899%2900004-4>). PMID 10064846 (<https://pubmed.ncbi.nlm.nih.gov/10064846>).

85. Beckman KB, Ames BN (August 1997). "Oxidative decay of DNA" (<https://doi.org/10.1074%2Fjbc.272.32.19633>). *The Journal of Biological Chemistry*. **272** (32): 19633–36. doi:10.1074/jbc.272.32.19633 (<https://doi.org/10.1074%2Fjbc.272.32.19633>). PMID 9289489 (<https://pubmed.ncbi.nlm.nih.gov/9289489>).
86. Valerie K, Povirk LF (September 2003). "Regulation and mechanisms of mammalian double-strand break repair" (<https://doi.org/10.1038%2Fsj.onc.1206679>). *Oncogene*. **22** (37): 5792–812. doi:10.1038/sj.onc.1206679 (<https://doi.org/10.1038%2Fsj.onc.1206679>). PMID 12947387 (<https://pubmed.ncbi.nlm.nih.gov/12947387>).
87. Johnson G (28 December 2010). "Unearthing Prehistoric Tumors, and Debate" (<https://www.nytimes.com/2010/12/28/health/28cancer.html>). *The New York Times*. Archived (<https://web.archive.org/web/20170624233156/http://www.nytimes.com/2010/12/28/health/28cancer.html>) from the original on 24 June 2017. "If we lived long enough, sooner or later we all would get cancer."
88. Alberts B, Johnson A, Lewis J, et al. (2002). "The Preventable Causes of Cancer" (<https://www.ncbi.nlm.nih.gov/books/NBK26897/>). *Molecular biology of the cell* (4th ed.). New York: Garland Science. ISBN 0-8153-4072-9. Archived (<https://web.archive.org/web/20160102193148/http://www.ncbi.nlm.nih.gov/books/NBK26897/>) from the original on 2 January 2016. "A certain irreducible background incidence of cancer is to be expected regardless of circumstances: mutations can never be absolutely avoided, because they are an inescapable consequence of fundamental limitations on the accuracy of DNA replication, as discussed in Chapter 5. If a human could live long enough, it is inevitable that at least one of his or her cells would eventually accumulate a set of mutations sufficient for cancer to develop."
89. Bernstein H, Payne CM, Bernstein C, Garewal H, Dvorak K (2008). "Cancer and aging as consequences of un-repaired DNA damage" (https://www.novapublishers.com/catalog/product_info.php?products_id=43247). In Kimura H, Suzuki A (eds.). *New Research on DNA Damage*. New York: Nova Science Publishers. pp. 1–47. ISBN 978-1-60456-581-2. Archived (https://web.archive.org/web/20141025091740/https://www.novapublishers.com/catalog/product_info.php?products_id=43247) from the original on 25 October 2014.
90. Hoeijmakers JH (October 2009). "DNA damage, aging, and cancer". *The New England Journal of Medicine*. **361** (15): 1475–85. doi:10.1056/NEJMr0804615 (<https://doi.org/10.1056%2FNEJMr0804615>). PMID 19812404 (<https://pubmed.ncbi.nlm.nih.gov/19812404>).
91. Freitas AA, de Magalhães JP (2011). "A review and appraisal of the DNA damage theory of ageing". *Mutation Research*. **728** (1–2): 12–22. Bibcode:2011MRRMR.728...12F (<https://ui.adsabs.harvard.edu/abs/2011MRRMR.728...12F>). doi:10.1016/j.mrrev.2011.05.001 (<https://doi.org/10.1016%2Fj.mrrev.2011.05.001>). PMID 21600302 (<https://pubmed.ncbi.nlm.nih.gov/21600302>).
92. Ferguson LR, Denny WA (September 1991). "The genetic toxicology of acridines". *Mutation Research*. **258** (2): 123–60. doi:10.1016/0165-1110(91)90006-H (<https://doi.org/10.1016%2F0165-1110%2891%2990006-H>). PMID 1881402 (<https://pubmed.ncbi.nlm.nih.gov/1881402>).
93. Stephens TD, Bunde CJ, Fillmore BJ (June 2000). "Mechanism of action in thalidomide teratogenesis". *Biochemical Pharmacology*. **59** (12): 1489–99. doi:10.1016/S0006-2952(99)00388-3 (<https://doi.org/10.1016%2FS0006-2952%2899%2900388-3>). PMID 10799645 (<https://pubmed.ncbi.nlm.nih.gov/10799645>).
94. Jeffrey AM (1985). "DNA modification by chemical carcinogens". *Pharmacology & Therapeutics*. **28** (2): 237–72. doi:10.1016/0163-7258(85)90013-0 (<https://doi.org/10.1016%2F0163-7258%2885%2990013-0>). PMID 3936066 (<https://pubmed.ncbi.nlm.nih.gov/3936066>).
95. Braña MF, Cacho M, Gradillas A, de Pascual-Teresa B, Ramos A (November 2001). "Intercalators as anticancer drugs". *Current Pharmaceutical Design*. **7** (17): 1745–80. doi:10.2174/1381612013397113 (<https://doi.org/10.2174%2F1381612013397113>). PMID 11562309 (<https://pubmed.ncbi.nlm.nih.gov/11562309>).

96. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. (February 2001). "The sequence of the human genome" (<https://doi.org/10.1126%2Fscience.1058040>). *Science*. **291** (5507): 1304–51. Bibcode:2001Sci...291.1304V (<https://ui.adsabs.harvard.edu/abs/2001Sci...291.1304V>). doi:10.1126/science.1058040 (<https://doi.org/10.1126%2Fscience.1058040>). PMID 11181995 (<https://pubmed.ncbi.nlm.nih.gov/11181995>).
97. Thanbichler M, Wang SC, Shapiro L (October 2005). "The bacterial nucleoid: a highly organized and dynamic structure" (<https://doi.org/10.1002%2Fjcb.20519>). *Journal of Cellular Biochemistry*. **96** (3): 506–21. doi:10.1002/jcb.20519 (<https://doi.org/10.1002%2Fjcb.20519>). PMID 15988757 (<https://pubmed.ncbi.nlm.nih.gov/15988757>).
98. Wolfsberg TG, McEntyre J, Schuler GD (February 2001). "Guide to the draft human genome" (<https://zenodo.org/record/1233093>). *Nature*. **409** (6822): 824–26. Bibcode:2001Natur.409..824W (<https://ui.adsabs.harvard.edu/abs/2001Natur.409..824W>). doi:10.1038/35057000 (<https://doi.org/10.1038%2F35057000>). PMID 11236998 (<https://pubmed.ncbi.nlm.nih.gov/11236998>).
99. Gregory TR (January 2005). "The C-value enigma in plants and animals: a review of parallels and an appeal for partnership" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4246714>). *Annals of Botany*. **95** (1): 133–46. doi:10.1093/aob/mci009 (<https://doi.org/10.1093%2Faob%2Fmci009>). PMC 4246714 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4246714>). PMID 15596463 (<https://pubmed.ncbi.nlm.nih.gov/15596463>).
100. Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, et al. (June 2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2212820>). *Nature*. **447** (7146): 799–816. Bibcode:2007Natur.447..799B (<https://ui.adsabs.harvard.edu/abs/2007Natur.447..799B>). doi:10.1038/nature05874 (<https://doi.org/10.1038%2Fnature05874>). PMC 2212820 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2212820>). PMID 17571346 (<https://pubmed.ncbi.nlm.nih.gov/17571346>).
101. Yin YW, Steitz TA. "RCSB PDB – 1MSW: Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase" (<https://www.rcsb.org/structure/1MSW>). *www.rcsb.org*. Retrieved 27 March 2023.
102. Pidoux AL, Allshire RC (March 2005). "The role of heterochromatin in centromere function" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1569473>). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. **360** (1455): 569–79. doi:10.1098/rstb.2004.1611 (<https://doi.org/10.1098%2Frstb.2004.1611>). PMC 1569473 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1569473>). PMID 15905142 (<https://pubmed.ncbi.nlm.nih.gov/15905142>).
103. Harrison PM, Hegyi H, Balasubramanian S, Luscombe NM, Bertone P, Echols N, Johnson T, Gerstein M (February 2002). "Molecular fossils in the human genome: identification and analysis of the pseudogenes in chromosomes 21 and 22" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC155275>). *Genome Research*. **12** (2): 272–80. doi:10.1101/gr.207102 (<https://doi.org/10.1101%2Fgr.207102>). PMC 155275 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC155275>). PMID 11827946 (<https://pubmed.ncbi.nlm.nih.gov/11827946>).
104. Harrison PM, Gerstein M (May 2002). "Studying genomes through the aeons: protein families, pseudogenes and proteome evolution". *Journal of Molecular Biology*. **318** (5): 1155–74. doi:10.1016/S0022-2836(02)00109-2 (<https://doi.org/10.1016%2FS0022-2836%2802%2900109-2>). PMID 12083509 (<https://pubmed.ncbi.nlm.nih.gov/12083509>).
105. Albà M (2001). "Replicative DNA polymerases" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC150442>). *Genome Biology*. **2** (1): REVIEWS3002. doi:10.1186/gb-2001-2-1-reviews3002 (<https://doi.org/10.1186%2Fgb-2001-2-1-reviews3002>). PMC 150442 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC150442>). PMID 11178285 (<https://pubmed.ncbi.nlm.nih.gov/11178285>).

106. Tani K, Nasu M (2010). "Roles of Extracellular DNA in Bacterial Ecosystems". In Kikuchi Y, Rykova EY (eds.). *Extracellular Nucleic Acids* (<https://archive.org/details/extracellularnuc00kiku>). Springer. pp. 25 (<https://archive.org/details/extracellularnuc00kiku/page/n35>)–38. ISBN 978-3-642-12616-1.
107. Vlassov VV, Laktionov PP, Rykova EY (July 2007). "Extracellular nucleic acids". *BioEssays*. **29** (7): 654–67. doi:10.1002/bies.20604 (<https://doi.org/10.1002%2Fbies.20604>). PMID 17563084 (<https://pubmed.ncbi.nlm.nih.gov/17563084>). S2CID 32463239 (<https://api.semanticscholar.org/CorpusID:32463239>).
108. Finkel SE, Kolter R (November 2001). "DNA as a nutrient: novel role for bacterial competence gene homologs" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC100116>). *Journal of Bacteriology*. **183** (21): 6288–93. doi:10.1128/JB.183.21.6288-6293.2001 (<https://doi.org/10.1128%2FJB.183.21.6288-6293.2001>). PMC 100116 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC100116>). PMID 11591672 (<https://pubmed.ncbi.nlm.nih.gov/11591672>).
109. Mulcahy H, Charron-Mazenod L, Lewenza S (November 2008). "Extracellular DNA chelates cations and induces antibiotic resistance in *Pseudomonas aeruginosa* biofilms" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2581603>). *PLOS Pathogens*. **4** (11): e1000213. doi:10.1371/journal.ppat.1000213 (<https://doi.org/10.1371%2Fjournal.ppat.1000213>). PMC 2581603 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2581603>). PMID 19023416 (<https://pubmed.ncbi.nlm.nih.gov/19023416>).
110. Berne C, Kysela DT, Brun YV (August 2010). "A bacterial extracellular DNA inhibits settling of motile progeny cells within a biofilm" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2962764>). *Molecular Microbiology*. **77** (4): 815–29. doi:10.1111/j.1365-2958.2010.07267.x (<https://doi.org/10.1111%2Fj.1365-2958.2010.07267.x>). PMC 2962764 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2962764>). PMID 20598083 (<https://pubmed.ncbi.nlm.nih.gov/20598083>).
111. Whitchurch CB, Tolker-Nielsen T, Ragas PC, Mattick JS (February 2002). "Extracellular DNA required for bacterial biofilm formation". *Science*. **295** (5559): 1487. doi:10.1126/science.295.5559.1487 (<https://doi.org/10.1126%2Fscience.295.5559.1487>). PMID 11859186 (<https://pubmed.ncbi.nlm.nih.gov/11859186>).
112. Hu W, Li L, Sharma S, Wang J, McHardy I, Lux R, Yang Z, He X, Gimzewski JK, Li Y, Shi W (2012). "DNA builds and strengthens the extracellular matrix in *Myxococcus xanthus* biofilms by interacting with exopolysaccharides" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530553>). *PLOS ONE*. **7** (12): e51905. Bibcode:2012PLoSO...751905H (<https://ui.adsabs.harvard.edu/abs/2012PLoSO...751905H>). doi:10.1371/journal.pone.0051905 (<https://doi.org/10.1371%2Fjournal.pone.0051905>). PMC 3530553 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530553>). PMID 23300576 (<https://pubmed.ncbi.nlm.nih.gov/23300576>).
113. Hui L, Bianchi DW (February 2013). "Recent advances in the prenatal interrogation of the human fetal genome" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4378900>). *Trends in Genetics*. **29** (2): 84–91. doi:10.1016/j.tig.2012.10.013 (<https://doi.org/10.1016%2Fj.tig.2012.10.013>). PMC 4378900 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4378900>). PMID 23158400 (<https://pubmed.ncbi.nlm.nih.gov/23158400>).
114. Foote AD, Thomsen PF, Sveegaard S, Wahlberg M, Kielgast J, Kyhn LA, et al. (2012). "Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430683>). *PLOS ONE*. **7** (8): e41781. Bibcode:2012PLoSO...741781F (<https://ui.adsabs.harvard.edu/abs/2012PLoSO...741781F>). doi:10.1371/journal.pone.0041781 (<https://doi.org/10.1371%2Fjournal.pone.0041781>). PMC 3430683 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430683>). PMID 22952587 (<https://pubmed.ncbi.nlm.nih.gov/22952587>).
115. "Researchers Detect Land Animals Using DNA in Nearby Water Bodies" (<https://www.the-scientist.com/news-opinion/researchers-detect-land-animals-using-dna-in-nearby-water-bodies-67481>).

116. Sandman K, Pereira SL, Reeve JN (December 1998). "Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11147202>). *Cellular and Molecular Life Sciences*. **54** (12): 1350–64. doi:10.1007/s000180050259 (<https://doi.org/10.1007%2Fs000180050259>). PMC 11147202 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11147202>). PMID 9893710 (<https://pubmed.ncbi.nlm.nih.gov/9893710>). S2CID 21101836 (<https://api.semanticscholar.org/CorpusID:21101836>).
117. Dame RT (May 2005). "The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin" (<https://doi.org/10.1111%2Fj.1365-2958.2005.04598.x>). *Molecular Microbiology*. **56** (4): 858–70. doi:10.1111/j.1365-2958.2005.04598.x (<https://doi.org/10.1111%2Fj.1365-2958.2005.04598.x>). PMID 15853876 (<https://pubmed.ncbi.nlm.nih.gov/15853876>). S2CID 26965112 (<https://api.semanticscholar.org/CorpusID:26965112>).
118. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (September 1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution". *Nature*. **389** (6648): 251–60. Bibcode:1997Natur.389..251L (<https://ui.adsabs.harvard.edu/abs/1997Natur.389..251L>). doi:10.1038/38444 (<https://doi.org/10.1038%2F38444>). PMID 9305837 (<https://pubmed.ncbi.nlm.nih.gov/9305837>). S2CID 4328827 (<https://api.semanticscholar.org/CorpusID:4328827>).
119. Jenuwein T, Allis CD (August 2001). "Translating the histone code" (<http://www.gs.washington.edu/academics/courses/braun/55104/readings/jenuwein.pdf>) (PDF). *Science*. **293** (5532): 1074–80. doi:10.1126/science.1063127 (<https://doi.org/10.1126%2Fscience.1063127>). PMID 11498575 (<https://pubmed.ncbi.nlm.nih.gov/11498575>). S2CID 1883924 (<https://api.semanticscholar.org/CorpusID:1883924>). Archived (<https://web.archive.org/web/20170808142426/http://www.gs.washington.edu/academics/courses/braun/55104/readings/jenuwein.pdf>) (PDF) from the original on 8 August 2017.
120. Ito T (2003). "Nucleosome Assembly and Remodeling". *Protein Complexes that Modify Chromatin*. Current Topics in Microbiology and Immunology. Vol. 274. pp. 1–22. doi:10.1007/978-3-642-55747-7_1 (https://doi.org/10.1007%2F978-3-642-55747-7_1). ISBN 978-3-540-44208-0. PMID 12596902 (<https://pubmed.ncbi.nlm.nih.gov/12596902>).
121. Thomas JO (August 2001). "HMG1 and 2: architectural DNA-binding proteins". *Biochemical Society Transactions*. **29** (Pt 4): 395–401. doi:10.1042/BST0290395 (<https://doi.org/10.1042%2FBST0290395>). PMID 11497996 (<https://pubmed.ncbi.nlm.nih.gov/11497996>).
122. Grosschedl R, Giese K, Pagel J (March 1994). "HMG domain proteins: architectural elements in the assembly of nucleoprotein structures". *Trends in Genetics*. **10** (3): 94–100. doi:10.1016/0168-9525(94)90232-1 (<https://doi.org/10.1016%2F0168-9525%2894%2990232-1>). PMID 8178371 (<https://pubmed.ncbi.nlm.nih.gov/8178371>).
123. Iftode C, Daniely Y, Borowiec JA (1999). "Replication protein A (RPA): the eukaryotic SSB". *Critical Reviews in Biochemistry and Molecular Biology*. **34** (3): 141–80. doi:10.1080/10409239991209255 (<https://doi.org/10.1080%2F10409239991209255>). PMID 10473346 (<https://pubmed.ncbi.nlm.nih.gov/10473346>).
124. Beamer LJ, Pabo CO. "RCSB PDB – 1LMB: Refined 1.8 Å crystal structure of the lambda repressor-operator complex" (<https://www.rcsb.org/structure/1LMB>). *www.rcsb.org*. Retrieved 27 March 2023.
125. Myers LC, Kornberg RD (2000). "Mediator of transcriptional regulation". *Annual Review of Biochemistry*. **69**: 729–49. doi:10.1146/annurev.biochem.69.1.729 (<https://doi.org/10.1146%2Fannurev.biochem.69.1.729>). PMID 10966474 (<https://pubmed.ncbi.nlm.nih.gov/10966474>).
126. Spiegelman BM, Heinrich R (October 2004). "Biological control through regulated transcriptional coactivators" (<https://doi.org/10.1016%2Fj.cell.2004.09.037>). *Cell*. **119** (2): 157–67. doi:10.1016/j.cell.2004.09.037 (<https://doi.org/10.1016%2Fj.cell.2004.09.037>). PMID 15479634 (<https://pubmed.ncbi.nlm.nih.gov/15479634>).

127. Li Z, Van Calcar S, Qu C, Cavenee WK, Zhang MQ, Ren B (July 2003). "A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC166200>). *Proceedings of the National Academy of Sciences of the United States of America*. **100** (14): 8164–69. Bibcode:2003PNAS..100.8164L (<https://ui.adsabs.harvard.edu/abs/2003PNAS..100.8164L>). doi:10.1073/pnas.1332764100 (<https://doi.org/10.1073/pnas.1332764100>). PMC 166200 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC166200>). PMID 12808131 (<https://pubmed.ncbi.nlm.nih.gov/12808131>).
128. Kostrewa D, Winkler FK. "RCSB PDB – 1RVA: Mg²⁺ binding to the active site of EcoRV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2 Å resolution" (<https://www.rcsb.org/structure/1RVA>). *www.rcsb.org*. Retrieved 27 March 2023.
129. Bickle TA, Krüger DH (June 1993). "Biology of DNA restriction" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372918>). *Microbiological Reviews*. **57** (2): 434–50. doi:10.1128/MMBR.57.2.434-450.1993 (<https://doi.org/10.1128/MMBR.57.2.434-450.1993>). PMC 372918 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372918>). PMID 8336674 (<https://pubmed.ncbi.nlm.nih.gov/8336674>).
130. Doherty AJ, Suh SW (November 2000). "Structural and mechanistic conservation in DNA ligases" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC113121>). *Nucleic Acids Research*. **28** (21): 4051–58. doi:10.1093/nar/28.21.4051 (<https://doi.org/10.1093/nar/28.21.4051>). PMC 113121 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC113121>). PMID 11058099 (<https://pubmed.ncbi.nlm.nih.gov/11058099>).
131. Schoeffler AJ, Berger JM (December 2005). "Recent advances in understanding structure-function relationships in the type II topoisomerase mechanism". *Biochemical Society Transactions*. **33** (Pt 6): 1465–70. doi:10.1042/BST20051465 (<https://doi.org/10.1042/BST20051465>). PMID 16246147 (<https://pubmed.ncbi.nlm.nih.gov/16246147>).
132. Tuteja N, Tuteja R (May 2004). "Unraveling DNA helicases. Motif, structure, mechanism and function" (<http://repository.ias.ac.in/52775/1/40-pub.pdf>) (PDF). *European Journal of Biochemistry*. **271** (10): 1849–63. doi:10.1111/j.1432-1033.2004.04094.x (<https://doi.org/10.1111/j.1432-1033.2004.04094.x>). PMID 15128295 (<https://pubmed.ncbi.nlm.nih.gov/15128295>).
133. Joyce CM, Steitz TA (November 1995). "Polymerase structures and function: variations on a theme?" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC177480>). *Journal of Bacteriology*. **177** (22): 6321–29. doi:10.1128/jb.177.22.6321-6329.1995 (<https://doi.org/10.1128/jb.177.22.6321-6329.1995>). PMC 177480 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC177480>). PMID 7592405 (<https://pubmed.ncbi.nlm.nih.gov/7592405>).
134. Hubscher U, Maga G, Spadari S (2002). "Eukaryotic DNA polymerases" (<https://web.archive.org/web/20210126170051/http://pdfs.semanticscholar.org/e941/98efed7eb8fa606b87d9a44c118c235a62e9.pdf>) (PDF). *Annual Review of Biochemistry*. **71**: 133–63. doi:10.1146/annurev.biochem.71.090501.150041 (<https://doi.org/10.1146/annurev.biochem.71.090501.150041>). PMID 12045093 (<https://pubmed.ncbi.nlm.nih.gov/12045093>). S2CID 26171993 (<https://api.semanticscholar.org/CorpusID:26171993>). Archived from the original (<http://pdfs.semanticscholar.org/e941/98efed7eb8fa606b87d9a44c118c235a62e9.pdf>) (PDF) on 26 January 2021.
135. Johnson A, O'Donnell M (2005). "Cellular DNA replicases: components and dynamics at the replication fork". *Annual Review of Biochemistry*. **74**: 283–315. doi:10.1146/annurev.biochem.73.011303.073859 (<https://doi.org/10.1146/annurev.biochem.73.011303.073859>). PMID 15952889 (<https://pubmed.ncbi.nlm.nih.gov/15952889>).
136. Tarrago-Litvak L, Andréola ML, Nevinsky GA, Sarih-Cottin L, Litvak S (May 1994). "The reverse transcriptase of HIV-1: from enzymology to therapeutic intervention" (<https://doi.org/10.1096/fasebj.8.8.7514143>). *FASEB Journal*. **8** (8): 497–503. doi:10.1096/fasebj.8.8.7514143 (<https://doi.org/10.1096/fasebj.8.8.7514143>). PMID 7514143 (<https://pubmed.ncbi.nlm.nih.gov/7514143>). S2CID 39614573 (<https://api.semanticscholar.org/CorpusID:39614573>).

137. Martinez E (December 2002). "Multi-protein complexes in eukaryotic gene transcription". *Plant Molecular Biology*. **50** (6): 925–47. doi:10.1023/A:1021258713850 (<https://doi.org/10.1023%2FA%3A1021258713850>). PMID 12516863 (<https://pubmed.ncbi.nlm.nih.gov/12516863/>). S2CID 24946189 (<https://api.semanticscholar.org/CorpusID:24946189>).
138. Thorpe JH, Gale BC, Teixeira SC, Cardin CJ. "RCSB PDB – 1M6G: Structural Characterisation of the Holliday Junction TCGGTACCGA" (<https://www.rcsb.org/structure/1M6G>). *www.rcsb.org*. Retrieved 27 March 2023.
139. Cremer T, Cremer C (April 2001). "Chromosome territories, nuclear architecture and gene regulation in mammalian cells". *Nature Reviews Genetics*. **2** (4): 292–301. doi:10.1038/35066075 (<https://doi.org/10.1038%2F35066075>). PMID 11283701 (<https://pubmed.ncbi.nlm.nih.gov/11283701/>). S2CID 8547149 (<https://api.semanticscholar.org/CorpusID:8547149>).
140. Pál C, Papp B, Lercher MJ (May 2006). "An integrated view of protein evolution". *Nature Reviews Genetics*. **7** (5): 337–48. doi:10.1038/nrg1838 (<https://doi.org/10.1038%2Fnrg1838>). PMID 16619049 (<https://pubmed.ncbi.nlm.nih.gov/16619049/>). S2CID 23225873 (<https://api.semanticscholar.org/CorpusID:23225873>).
141. O'Driscoll M, Jeggo PA (January 2006). "The role of double-strand break repair – insights from human genetics". *Nature Reviews Genetics*. **7** (1): 45–54. doi:10.1038/nrg1746 (<https://doi.org/10.1038%2Fnrg1746>). PMID 16369571 (<https://pubmed.ncbi.nlm.nih.gov/16369571/>). S2CID 7779574 (<https://api.semanticscholar.org/CorpusID:7779574>).
142. Vispé S, Defais M (October 1997). "Mammalian Rad51 protein: a RecA homologue with pleiotropic functions". *Biochimie*. **79** (9–10): 587–92. doi:10.1016/S0300-9084(97)82007-X (<https://doi.org/10.1016%2FS0300-9084%2897%2982007-X>). PMID 9466696 (<https://pubmed.ncbi.nlm.nih.gov/9466696/>).
143. Neale MJ, Keeney S (July 2006). "Clarifying the mechanics of DNA strand exchange in meiotic recombination" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5607947>). *Nature*. **442** (7099): 153–58. Bibcode:2006Natur.442..153N (<https://ui.adsabs.harvard.edu/abs/2006Natur.442..153N>). doi:10.1038/nature04885 (<https://doi.org/10.1038%2Fnature04885>). PMC 5607947 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5607947>). PMID 16838012 (<https://pubmed.ncbi.nlm.nih.gov/16838012/>).
144. Dickman MJ, Ingleston SM, Sedelnikova SE, Rafferty JB, Lloyd RG, Grasby JA, Hornby DP (November 2002). "The RuvABC resolvasome" (<https://doi.org/10.1046%2Fj.1432-1033.2002.03250.x>). *European Journal of Biochemistry*. **269** (22): 5492–501. doi:10.1046/j.1432-1033.2002.03250.x (<https://doi.org/10.1046%2Fj.1432-1033.2002.03250.x>). PMID 12423347 (<https://pubmed.ncbi.nlm.nih.gov/12423347/>). S2CID 39505263 (<https://api.semanticscholar.org/CorpusID:39505263>).
145. Joyce GF (July 2002). "The antiquity of RNA-based evolution". *Nature*. **418** (6894): 214–21. Bibcode:2002Natur.418..214J (<https://ui.adsabs.harvard.edu/abs/2002Natur.418..214J>). doi:10.1038/418214a (<https://doi.org/10.1038%2F418214a>). PMID 12110897 (<https://pubmed.ncbi.nlm.nih.gov/12110897/>). S2CID 4331004 (<https://api.semanticscholar.org/CorpusID:4331004>).
146. Orgel LE (2004). "Prebiotic chemistry and the origin of the RNA world". *Critical Reviews in Biochemistry and Molecular Biology*. **39** (2): 99–123. CiteSeerX 10.1.1.537.7679 (<https://cite-seerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.537.7679>). doi:10.1080/10409230490460765 (<https://doi.org/10.1080%2F10409230490460765>). PMID 15217990 (<https://pubmed.ncbi.nlm.nih.gov/15217990/>). S2CID 4939632 (<https://api.semanticscholar.org/CorpusID:4939632>).
147. Davenport RJ (May 2001). "Ribozymes. Making copies in the RNA world". *Science*. **292** (5520): 1278a–1278. doi:10.1126/science.292.5520.1278a (<https://doi.org/10.1126%2Fscience.292.5520.1278a>). PMID 11360970 (<https://pubmed.ncbi.nlm.nih.gov/11360970/>). S2CID 85976762 (<https://api.semanticscholar.org/CorpusID:85976762>).

148. Szathmáry E (April 1992). "What is the optimum size for the genetic alphabet?" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC48712>). *Proceedings of the National Academy of Sciences of the United States of America*. **89** (7): 2614–18. Bibcode:1992PNAS...89.2614S (<https://ui.adsabs.harvard.edu/abs/1992PNAS...89.2614S>). doi:10.1073/pnas.89.7.2614 (<https://doi.org/10.1073%2Fpnas.89.7.2614>). PMC 48712 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC48712>). PMID 1372984 (<https://pubmed.ncbi.nlm.nih.gov/1372984>).
149. Lindahl T (April 1993). "Instability and decay of the primary structure of DNA". *Nature*. **362** (6422): 709–15. Bibcode:1993Natur.362..709L (<https://ui.adsabs.harvard.edu/abs/1993Natur.362..709L>). doi:10.1038/362709a0 (<https://doi.org/10.1038%2F362709a0>). PMID 8469282 (<https://pubmed.ncbi.nlm.nih.gov/8469282>). S2CID 4283694 (<https://api.semanticscholar.org/CorpusID:4283694>).
150. Vreeland RH, Rosenzweig WD, Powers DW (October 2000). "Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal". *Nature*. **407** (6806): 897–900. Bibcode:2000Natur.407..897V (<https://ui.adsabs.harvard.edu/abs/2000Natur.407..897V>). doi:10.1038/35038060 (<https://doi.org/10.1038%2F35038060>). PMID 11057666 (<https://pubmed.ncbi.nlm.nih.gov/11057666>). S2CID 9879073 (<https://api.semanticscholar.org/CorpusID:9879073>).
151. Hebsgaard MB, Phillips MJ, Willerslev E (May 2005). "Geologically ancient DNA: fact or artefact?". *Trends in Microbiology*. **13** (5): 212–20. doi:10.1016/j.tim.2005.03.010 (<https://doi.org/10.1016%2Fj.tim.2005.03.010>). PMID 15866038 (<https://pubmed.ncbi.nlm.nih.gov/15866038>).
152. Nickle DC, Learn GH, Rain MW, Mullins JI, Mittler JE (January 2002). "Curiously modern DNA for a "250 million-year-old" bacterium". *Journal of Molecular Evolution*. **54** (1): 134–37. Bibcode:2002JMolE..54..134N (<https://ui.adsabs.harvard.edu/abs/2002JMolE..54..134N>). doi:10.1007/s00239-001-0025-x (<https://doi.org/10.1007%2Fs00239-001-0025-x>). PMID 11734907 (<https://pubmed.ncbi.nlm.nih.gov/11734907>). S2CID 24740859 (<https://api.semanticscholar.org/CorpusID:24740859>).
153. Callahan MP, Smith KE, Cleaves HJ, Ruzicka J, Stern JC, Glavin DP, House CH, Dworkin JP (August 2011). "Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3161613>). *Proceedings of the National Academy of Sciences of the United States of America*. **108** (34): 13995–98. Bibcode:2011PNAS..10813995C (<https://ui.adsabs.harvard.edu/abs/2011PNAS..10813995C>). doi:10.1073/pnas.1106493108 (<https://doi.org/10.1073%2Fpnas.1106493108>). PMC 3161613 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3161613>). PMID 21836052 (<https://pubmed.ncbi.nlm.nih.gov/21836052>).
154. Steigerwald J (8 August 2011). "NASA Researchers: DNA Building Blocks Can Be Made in Space" (<https://www.nasa.gov/topics/solarsystem/features/dna-meteorites.html>). NASA. Archived (<https://web.archive.org/web/20150623004556/http://www.nasa.gov/topics/solarsystem/features/dna-meteorites.html>) from the original on 23 June 2015. Retrieved 10 August 2011.
155. ScienceDaily Staff (9 August 2011). "DNA Building Blocks Can Be Made in Space, NASA Evidence Suggests" (<https://www.sciencedaily.com/releases/2011/08/110808220659.htm>). *ScienceDaily*. Archived (<https://web.archive.org/web/20110905105043/https://www.sciencedaily.com/releases/2011/08/110808220659.htm>) from the original on 5 September 2011. Retrieved 9 August 2011.
156. Marlaire R (3 March 2015). "NASA Ames Reproduces the Building Blocks of Life in Laboratory" (<https://www.nasa.gov/content/nasa-ames-reproduces-the-building-blocks-of-life-in-laboratory>). NASA. Archived (<https://web.archive.org/web/20150305083306/http://www.nasa.gov/content/nasa-ames-reproduces-the-building-blocks-of-life-in-laboratory/>) from the original on 5 March 2015. Retrieved 5 March 2015.
157. Hunt K (17 February 2021). "World's oldest DNA sequenced from a mammoth that lived more than a million years ago" (<https://www.cnn.com/2021/02/17/world/mammoth-oldest-dna-million-years-ago-scn/index.html>). *CNN News*. Retrieved 17 February 2021.

158. Callaway E (17 February 2021). "Million-year-old mammoth genomes shatter record for oldest ancient DNA – Permafrost-preserved teeth, up to 1.6 million years old, identify a new kind of mammoth in Siberia" (<https://doi.org/10.1038%2Fd41586-021-00436-x>). *Nature*. **590** (7847): 537–538. Bibcode:2021Natur.590..537C (<https://ui.adsabs.harvard.edu/abs/2021Natur.590..537C>). doi:10.1038/d41586-021-00436-x (<https://doi.org/10.1038%2Fd41586-021-00436-x>). ISSN 0028-0836 (<https://search.worldcat.org/issn/0028-0836>). PMID 33597786 (<https://pubmed.ncbi.nlm.nih.gov/33597786>).
159. Goff SP, Berg P (December 1976). "Construction of hybrid viruses containing SV40 and lambda phage DNA segments and their propagation in cultured monkey cells". *Cell*. **9** (4 PT 2): 695–705. doi:10.1016/0092-8674(76)90133-1 (<https://doi.org/10.1016%2F0092-8674%2876%2990133-1>). PMID 189942 (<https://pubmed.ncbi.nlm.nih.gov/189942>). S2CID 41788896 (<https://api.semanticscholar.org/CorpusID:41788896>).
160. Houdebine LM (2007). "Transgenic animal models in biomedical research". *Target Discovery and Validation Reviews and Protocols*. Methods in Molecular Biology. Vol. 360. pp. 163–202. doi:10.1385/1-59745-165-7:163 (<https://doi.org/10.1385%2F1-59745-165-7%3A163>). ISBN 978-1-59745-165-9. PMID 17172731 (<https://pubmed.ncbi.nlm.nih.gov/17172731>).
161. Daniell H, Dhingra A (April 2002). "Multigene engineering: dawn of an exciting new era in biotechnology" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3481857>). *Current Opinion in Biotechnology*. **13** (2): 136–41. doi:10.1016/S0958-1669(02)00297-5 (<https://doi.org/10.1016%2FS0958-1669%2802%2900297-5>). PMC 3481857 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3481857>). PMID 11950565 (<https://pubmed.ncbi.nlm.nih.gov/11950565>).
162. Job D (November 2002). "Plant biotechnology in agriculture". *Biochimie*. **84** (11): 1105–10. doi:10.1016/S0300-9084(02)00013-5 (<https://doi.org/10.1016%2FS0300-9084%2802%2900013-5>). PMID 12595138 (<https://pubmed.ncbi.nlm.nih.gov/12595138>).
163. Curtis C, Hereward J (29 August 2017). "From the crime scene to the courtroom: the journey of a DNA sample" (<https://theconversation.com/from-the-crime-scene-to-the-courtroom-the-journey-of-a-dna-sample-82250>). *The Conversation*. Archived (<https://web.archive.org/web/20171022033110/http://theconversation.com/from-the-crime-scene-to-the-courtroom-the-journey-of-a-dna-sample-82250>) from the original on 22 October 2017. Retrieved 22 October 2017.
164. Collins A, Morton NE (June 1994). "Likelihood ratios for DNA identification" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC44126>). *Proceedings of the National Academy of Sciences of the United States of America*. **91** (13): 6007–11. Bibcode:1994PNAS...91.6007C (<https://ui.adsabs.harvard.edu/abs/1994PNAS...91.6007C>). doi:10.1073/pnas.91.13.6007 (<https://doi.org/10.1073%2Fpnas.91.13.6007>). PMC 44126 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC44126>). PMID 8016106 (<https://pubmed.ncbi.nlm.nih.gov/8016106>).
165. Weir BS, Triggs CM, Starling L, Stowell LI, Walsh KA, Buckleton J (March 1997). "Interpreting DNA mixtures". *Journal of Forensic Sciences*. **42** (2): 213–22. doi:10.1520/JFS14100J (<https://doi.org/10.1520%2FJFS14100J>). PMID 9068179 (<https://pubmed.ncbi.nlm.nih.gov/9068179>). S2CID 14511630 (<https://api.semanticscholar.org/CorpusID:14511630>).
166. Jeffreys AJ, Wilson V, Thein SL (1985). "Individual-specific 'fingerprints' of human DNA" (<https://doi.org/10.1038%2F316076a0>). *Nature*. **316** (6023): 76–79. Bibcode:1985Natur.316...76J (<https://ui.adsabs.harvard.edu/abs/1985Natur.316...76J>). doi:10.1038/316076a0 (<https://doi.org/10.1038%2F316076a0>). PMID 2989708 (<https://pubmed.ncbi.nlm.nih.gov/2989708>). S2CID 4229883 (<https://api.semanticscholar.org/CorpusID:4229883>).
167. "Colin Pitchfork" (https://web.archive.org/web/20061214004903/http://www.forensic.gov.uk/forensic_t/inside/news/list_casefiles.php?case=1). 14 December 2006. Archived from the original (http://www.forensic.gov.uk/forensic_t/inside/news/list_casefiles.php?case=1) on 14 December 2006. Retrieved 27 March 2023.

168. "DNA Identification in Mass Fatality Incidents" (<https://web.archive.org/web/20061112000837/http://massfatality.dna.gov/Introduction/>). National Institute of Justice. September 2006. Archived from the original (<http://massfatality.dna.gov/Introduction/>) on 12 November 2006.
169. Pollack A (19 June 2012). "Before Birth, Dad's ID" (<https://www.nytimes.com/2012/06/20/health/paternity-blood-tests-that-work-early-in-a-pregnancy.html>). *The New York Times*. ISSN 0362-4331 (<https://search.worldcat.org/issn/0362-4331>). Archived (<https://web.archive.org/web/20170624231639/http://www.nytimes.com/2012/06/20/health/paternity-blood-tests-that-work-early-in-a-pregnancy.html>) from the original on 24 June 2017. Retrieved 27 March 2023.
170. Breaker RR, Joyce GF (December 1994). "A DNA enzyme that cleaves RNA". *Chemistry & Biology*. **1** (4): 223–29. doi:10.1016/1074-5521(94)90014-0 (<https://doi.org/10.1016%2F1074-5521%2894%2990014-0>). PMID 9383394 (<https://pubmed.ncbi.nlm.nih.gov/9383394>).
171. Chandra M, Sachdeva A, Silverman SK (October 2009). "DNA-catalyzed sequence-specific hydrolysis of DNA" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2746877>). *Nature Chemical Biology*. **5** (10): 718–20. doi:10.1038/nchembio.201 (<https://doi.org/10.1038%2Fncmbio.201>). PMC 2746877 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2746877>). PMID 19684594 (<https://pubmed.ncbi.nlm.nih.gov/19684594>).
172. Carmi N, Shultz LA, Breaker RR (December 1996). "In vitro selection of self-cleaving DNAs" (<https://doi.org/10.1016%2FS1074-5521%2896%2990170-2>). *Chemistry & Biology*. **3** (12): 1039–46. doi:10.1016/S1074-5521(96)90170-2 (<https://doi.org/10.1016%2FS1074-5521%2896%2990170-2>). PMID 9000012 (<https://pubmed.ncbi.nlm.nih.gov/9000012>).
173. Torabi SF, Wu P, McGhee CE, Chen L, Hwang K, Zheng N, Cheng J, Lu Y (May 2015). "In vitro selection of a sodium-specific DNAzyme and its application in intracellular sensing" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4434688>). *Proceedings of the National Academy of Sciences of the United States of America*. **112** (19): 5903–08. Bibcode:2015PNAS..112.5903T (<https://ui.adsabs.harvard.edu/abs/2015PNAS..112.5903T>). doi:10.1073/pnas.1420361112 (<https://doi.org/10.1073%2Fpnas.1420361112>). PMC 4434688 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4434688>). PMID 25918425 (<https://pubmed.ncbi.nlm.nih.gov/25918425>).
174. Baldi P, Brunak S (2001). *Bioinformatics: The Machine Learning Approach*. MIT Press. ISBN 978-0-262-02506-5. OCLC 45951728 (<https://search.worldcat.org/oclc/45951728>).
175. Gusfield D (15 January 1997). *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press. ISBN 978-0-521-58519-4.
176. Sjölander K (January 2004). "Phylogenomic inference of protein molecular function: advances and challenges". *Bioinformatics*. **20** (2): 170–79. CiteSeerX 10.1.1.412.943 (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.412.943>). doi:10.1093/bioinformatics/bth021 (<https://doi.org/10.1093%2Fbioinformatics%2Fbth021>). PMID 14734307 (<https://pubmed.ncbi.nlm.nih.gov/14734307>).
177. Mount DM (2004). *Bioinformatics: Sequence and Genome Analysis* (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. ISBN 0-87969-712-1. OCLC 55106399 (<https://search.worldcat.org/oclc/55106399>).
178. Strong M (March 2004). "Protein nanomachines" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC368168>). *PLOS Biology*. **2** (3): E73. doi:10.1371/journal.pbio.0020073 (<https://doi.org/10.1371%2Fjournal.pbio.0020073>). PMC 368168 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC368168>). PMID 15024422 (<https://pubmed.ncbi.nlm.nih.gov/15024422>). S2CID 13222080 (<https://api.semanticscholar.org/CorpusID:13222080>).
179. Rothemund PW (March 2006). "Folding DNA to create nanoscale shapes and patterns" (<https://authors.library.caltech.edu/22244/3/nature04586-s2.pdf>) (PDF). *Nature*. **440** (7082): 297–302. Bibcode:2006Natur.440..297R (<https://ui.adsabs.harvard.edu/abs/2006Natur.440..297R>). doi:10.1038/nature04586 (<https://doi.org/10.1038%2Fnature04586>). PMID 16541064 (<https://pubmed.ncbi.nlm.nih.gov/16541064>). S2CID 4316391 (<https://api.semanticscholar.org/CorpusID:4316391>).

180. Andersen ES, Dong M, Nielsen MM, Jahn K, Subramani R, Mamdouh W, Golas MM, Sander B, Stark H, Oliveira CL, Pedersen JS, Birkedal V, Besenbacher F, Gothelf KV, Kjems J (May 2009). "Self-assembly of a nanoscale DNA box with a controllable lid". *Nature*. **459** (7243): 73–76. Bibcode:2009Natur.459...73A (<https://ui.adsabs.harvard.edu/abs/2009Natur.459...73A>). doi:10.1038/nature07971 (<https://doi.org/10.1038%2Fnature07971>). hdl:11858/00-001M-0000-0010-9362-B (<https://hdl.handle.net/11858%2F00-001M-0000-0010-9362-B>). PMID 19424153 (<https://pubmed.ncbi.nlm.nih.gov/19424153>). S2CID 4430815 (<https://api.semanticscholar.org/CorpusID:4430815>).
181. Ishitsuka Y, Ha T (May 2009). "DNA nanotechnology: a nanomachine goes live". *Nature Nanotechnology*. **4** (5): 281–82. Bibcode:2009NatNa...4..281I (<https://ui.adsabs.harvard.edu/abs/2009NatNa...4..281I>). doi:10.1038/nnano.2009.101 (<https://doi.org/10.1038%2Fnnano.2009.101>). PMID 19421208 (<https://pubmed.ncbi.nlm.nih.gov/19421208>).
182. Aldaye FA, Palmer AL, Sleiman HF (September 2008). "Assembling materials with DNA as the guide". *Science*. **321** (5897): 1795–99. Bibcode:2008Sci...321.1795A (<https://ui.adsabs.harvard.edu/abs/2008Sci...321.1795A>). doi:10.1126/science.1154533 (<https://doi.org/10.1126%2Fscience.1154533>). PMID 18818351 (<https://pubmed.ncbi.nlm.nih.gov/18818351>). S2CID 2755388 (<https://api.semanticscholar.org/CorpusID:2755388>).
183. Dunn MR, Jimenez RM, Chaput JC (2017). "Analysis of aptamer discovery and technology" (<https://www.nature.com/articles/s41570-017-0076>). *Nature Reviews Chemistry*. **1** (10). doi:10.1038/s41570-017-0076 (<https://doi.org/10.1038%2Fs41570-017-0076>). Retrieved 30 June 2022.
184. Wray GA (2002). "Dating branches on the tree of life using DNA" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC150454>). *Genome Biology*. **3** (1): REVIEWS0001. doi:10.1186/gb-2001-3-1-reviews0001 (<https://doi.org/10.1186%2Fgb-2001-3-1-reviews0001>). PMC 150454 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC150454>). PMID 11806830 (<https://pubmed.ncbi.nlm.nih.gov/11806830>).
185. Panda D, Molla KA, Baig MJ, Swain A, Behera D, Dash M (May 2018). "DNA as a digital information storage device: hope or hype?" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5935598>). *3 Biotech*. **8** (5): 239. doi:10.1007/s13205-018-1246-7 (<https://doi.org/10.1007%2Fs13205-018-1246-7>). PMC 5935598 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5935598>). PMID 29744271 (<https://pubmed.ncbi.nlm.nih.gov/29744271>).
186. Akram F, Haq IU, Ali H, Laghari AT (October 2018). "Trends to store digital data in DNA: an overview". *Molecular Biology Reports*. **45** (5): 1479–1490. doi:10.1007/s11033-018-4280-y (<https://doi.org/10.1007%2Fs11033-018-4280-y>). PMID 30073589 (<https://pubmed.ncbi.nlm.nih.gov/30073589>). S2CID 51905843 (<https://api.semanticscholar.org/CorpusID:51905843>).
187. Miescher F (1871). "Ueber die chemische Zusammensetzung der Eiterzellen" (<https://books.google.com/books?id=YJRTAAAcAAJ&pg=PA441>) [On the chemical composition of pus cells]. *Medicinisch-chemische Untersuchungen* (in German). **4**: 441–60. "[p. 456] Ich habe mich daher später mit meinen Versuchen an die ganzen Kerne gehalten, die Trennung der Körper, die ich einstweilen ohne weiteres Präjudiz als lösliches und unlösliches Nuclein bezeichnen will, einem günstigeren Material überlassend. (Therefore, in my experiments I subsequently limited myself to the whole nucleus, leaving to a more favorable material the separation of the substances, that for the present, without further prejudice, I will designate as soluble and insoluble nuclear material ("Nuclein"))"
188. Dahm R (January 2008). "Discovering DNA: Friedrich Miescher and the early years of nucleic acid research". *Human Genetics*. **122** (6): 565–81. doi:10.1007/s00439-007-0433-0 (<https://doi.org/10.1007%2Fs00439-007-0433-0>). PMID 17901982 (<https://pubmed.ncbi.nlm.nih.gov/17901982>). S2CID 915930 (<https://api.semanticscholar.org/CorpusID:915930>).

189. See:

- Kossel A (1879). "Ueber Nuclein der Hefe" (<https://books.google.com/books?id=4H5NAAAYAAJ&pg=PA284>) [On nuclein in yeast]. *Zeitschrift für physiologische Chemie* (in German). **3**: 284–91.
 - Kossel A (1880). "Ueber Nuclein der Hefe II" (<https://books.google.com/books?id=u4s1AQAAMAAJ&pg=PA290>) [On nuclein in yeast, Part 2]. *Zeitschrift für physiologische Chemie* (in German). **4**: 290–95.
 - Kossel A (1881). "Ueber die Verbreitung des Hypoxanthins im Thier- und Pflanzenreich" (<https://books.google.com/books?id=xYs1AQAAMAAJ&pg=PA267>) [On the distribution of hypoxanthins in the animal and plant kingdoms]. *Zeitschrift für physiologische Chemie* (in German). **5**: 267–71.
 - Kossel A (1881). *Untersuchungen über die Nucleine und ihre Spaltungsprodukte* [Investigations into nuclein and its cleavage products] (in German). Strassburg, Germany: K.J. Trübner. p. 19.
 - Kossel A (1882). "Ueber Xanthin und Hypoxanthin" (<https://books.google.com/books?id=z4s1AQAAMAAJ&pg=PA422>) [On xanthin and hypoxanthin]. *Zeitschrift für physiologische Chemie*. **6**: 422–31.
 - Albrecht Kossel (1883) "Zur Chemie des Zellkerns" (<https://books.google.com/books?id=2os1AQAAMAAJ&pg=PA7>) Archived (<https://web.archive.org/web/20171117235430/http://books.google.com/books?id=2os1AQAAMAAJ&pg=PA7>) 17 November 2017 at the Wayback Machine (On the chemistry of the cell nucleus), *Zeitschrift für physiologische Chemie*, **7**: 7–22.
 - Kossel A (1886). "Weitere Beiträge zur Chemie des Zellkerns" (http://vlp.mpiwg-berlin.mpg.de/library/data/lit16615/index_html?pn=1&ws=1.5) [Further contributions to the chemistry of the cell nucleus]. *Zeitschrift für Physiologische Chemie* (in German). **10**: 248–64. "On p. 264, Kossel remarked presciently: Der Erforschung der quantitativen Verhältnisse der vier stickstoffreichen Basen, der Abhängigkeit ihrer Menge von den physiologischen Zuständen der Zelle, verspricht wichtige Aufschlüsse über die elementaren physiologisch-chemischen Vorgänge. (The study of the quantitative relations of the four nitrogenous bases—[and] of the dependence of their quantity on the physiological states of the cell—promises important insights into the fundamental physiological-chemical processes.)"
190. Jones ME (September 1953). "Albrecht Kossel, a biographical sketch" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2599350>). *The Yale Journal of Biology and Medicine*. **26** (1): 80–97. PMC 2599350 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2599350>). PMID 13103145 (<https://pubmed.ncbi.nlm.nih.gov/13103145>).
191. Levene PA, Jacobs WA (1909). "Über Inosinsäure" (<https://babel.hathitrust.org/cgi/pt?id=iau.31858002459620&view=1up&seq=1054>). *Berichte der Deutschen Chemischen Gesellschaft* (in German). **42**: 1198–203. doi:10.1002/cber.190904201196 (<https://doi.org/10.1002%2Fcber.190904201196>).
192. Levene PA, Jacobs WA (1909). "Über die Hefe-Nucleinsäure" (<https://zenodo.org/record/2175598>). *Berichte der Deutschen Chemischen Gesellschaft* (in German). **42** (2): 2474–78. doi:10.1002/cber.190904202148 (<https://doi.org/10.1002%2Fcber.190904202148>).
193. Levene P (1919). "The structure of yeast nucleic acid" (<https://doi.org/10.1016%2FS0021-9258%2818%2987254-4>). *J Biol Chem*. **40** (2): 415–24. doi:10.1016/S0021-9258(18)87254-4 (<https://doi.org/10.1016%2FS0021-9258%2818%2987254-4>).
194. Cohen JS, Portugal FH (1974). "The search for the chemical structure of DNA" (<https://profiles.nlm.nih.gov/ps/access/CCAAHW.pdf>) (PDF). *Connecticut Medicine*. **38** (10): 551–52, 554–57. PMID 4609088 (<https://pubmed.ncbi.nlm.nih.gov/4609088>).

195. Koltsov proposed that a cell's genetic information was encoded in a long chain of amino acids. See:

- Koltsov HK (12 December 1927). *Физико-химические основы морфологии* [*The physical-chemical basis of morphology*] (Speech). 3rd All-Union Meeting of Zoologist, Anatomists, and Histologists (in Russian). Leningrad, U.S.S.R.
- Reprinted in: Koltsov HK (1928). "Физико-химические основы морфологии" [*The physical-chemical basis of morphology*]. *Успехи экспериментальной биологии* (*Advances in Experimental Biology*) series B (in Russian). **7** (1): ?.
- Reprinted in German as: Koltzoff NK (1928). "Physikalisch-chemische Grundlagen der Morphologie" [*The physical-chemical basis of morphology*]. *Biologisches Zentralblatt* (in German). **48** (6): 345–69.
- In 1934, Koltsov contended that the proteins that contain a cell's genetic information replicate. See: Koltzoff N (October 1934). "The structure of the chromosomes in the salivary glands of *Drosophila*". *Science*. **80** (2075): 312–13. Bibcode:1934Sci....80..312K (<https://ui.adsabs.harvard.edu/abs/1934Sci....80..312K>). doi:10.1126/science.80.2075.312 (<https://doi.org/10.1126/science.80.2075.312>). PMID 17769043 (<https://pubmed.ncbi.nlm.nih.gov/17769043>). "From page 313: "I think that the size of the chromosomes in the salivary glands [of *Drosophila*] is determined through the multiplication of *genomes*. By this term I designate the axial thread of the chromosome, in which the geneticists locate the linear combination of genes; ... In the normal chromosome there is usually only one *genome*; before cell-division this *genome* has become divided into two strands." "

196. Soyfer VN (September 2001). "The consequences of political dictatorship for Russian science". *Nature Reviews Genetics*. **2** (9): 723–29. doi:10.1038/35088598 (<https://doi.org/10.1038/35088598>). PMID 11533721 (<https://pubmed.ncbi.nlm.nih.gov/11533721>). S2CID 46277758 (<https://api.semanticscholar.org/CorpusID:46277758>).

197. Griffith F (January 1928). "The Significance of Pneumococcal Types" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2167760>). *The Journal of Hygiene*. **27** (2): 113–59. doi:10.1017/S0022172400031879 (<https://doi.org/10.1017/S0022172400031879>). PMC 2167760 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2167760>). PMID 20474956 (<https://pubmed.ncbi.nlm.nih.gov/20474956>).

198. Lorenz MG, Wackernagel W (September 1994). "Bacterial gene transfer by natural genetic transformation in the environment" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372978>). *Microbiological Reviews*. **58** (3): 563–602. doi:10.1128/MMBR.58.3.563-602.1994 (<https://doi.org/10.1128/MMBR.58.3.563-602.1994>). PMC 372978 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372978>). PMID 7968924 (<https://pubmed.ncbi.nlm.nih.gov/7968924>).

199. Brachet J (1933). "Recherches sur la synthese de l'acide thymonucleique pendant le developpement de l'oeuf d'Oursin". *Archives de Biologie* (in Italian). **44**: 519–76.

200. Burian R (1994). "Jean Brachet's Cytochemical Embryology: Connections with the Renovation of Biology in France?" (<http://www.histcnrs.fr/ColloqDijon/Burian-Brachet.pdf>) (PDF). In Debru C, Gayon J, Picard JF (eds.). *Les sciences biologiques et médicales en France 1920–1950*. Cahiers pour l'histoire de la recherche. Vol. 2. Paris: CNRS Editions. pp. 207–20.

201. See:

- Astbury WT, Bell FO (1938). "Some recent developments in the X-ray study of proteins and related structures" (https://web.archive.org/web/20140714204539/http://www.leeds.ac.uk/heritage/Astbury/bibliography/CSHSQB_Astbury_and_Bell_1938.pdf) (PDF). *Cold Spring Harbor Symposia on Quantitative Biology*. **6**: 109–21. doi:10.1101/sqb.1938.006.01.013 (<https://doi.org/10.1101%2Fsqb.1938.006.01.013>). Archived from the original (http://www.leeds.ac.uk/heritage/Astbury/bibliography/CSHSQB_Astbury_and_Bell_1938.pdf) (PDF) on 14 July 2014.
 - Astbury WT (1947). "X-ray studies of nucleic acids" (<https://web.archive.org/web/20140705132403/http://scarc.library.oregonstate.edu/coll/pauling/dna/papers/astbury-xray.html>). *Symposia of the Society for Experimental Biology* (1): 66–76. PMID 20257017 (<https://pubmed.ncbi.nlm.nih.gov/20257017>). Archived from the original (<http://scarc.library.oregonstate.edu/coll/pauling/dna/papers/astbury-xray.html>) on 5 July 2014.
202. Avery OT, Macleod CM, McCarty M (February 1944). "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from *Pneumococcus* Type III" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2135445>). *The Journal of Experimental Medicine*. **79** (2): 137–158. doi:10.1084/jem.79.2.137 (<https://doi.org/10.1084%2Fjem.79.2.137>). PMC 2135445 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2135445>). PMID 19871359 (<https://pubmed.ncbi.nlm.nih.gov/19871359>).
203. Chargaff E (June 1950). "Chemical specificity of nucleic acids and mechanism of their enzymatic degradation". *Experientia*. **6** (6): 201–209. doi:10.1007/BF02173653 (<https://doi.org/10.1007%2FBF02173653>). PMID 15421335 (<https://pubmed.ncbi.nlm.nih.gov/15421335>). S2CID 2522535 (<https://api.semanticscholar.org/CorpusID:2522535>).
204. Kresge N, Simoni RD, Hill RL (June 2005). "Chargaff's Rules: the Work of Erwin Chargaff" (<https://doi.org/10.1016%2FS0021-9258%2820%2961522-8>). *Journal of Biological Chemistry*. **280** (24): 172–174. doi:10.1016/S0021-9258(20)61522-8 (<https://doi.org/10.1016%2FS0021-9258%2820%2961522-8>).
205. Hershey AD, Chase M (May 1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2147348>). *The Journal of General Physiology*. **36** (1): 39–56. doi:10.1085/jgp.36.1.39 (<https://doi.org/10.1085%2Fjgp.36.1.39>). PMC 2147348 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2147348>). PMID 12981234 (<https://pubmed.ncbi.nlm.nih.gov/12981234>).
206. "Pictures and Illustrations: Crystallographic photo of Sodium Thymonucleate, Type B. "Photo 51." May 1952" (<http://scarc.library.oregonstate.edu/coll/pauling/dna/pictures/sci9.001.5.html>). *scarc.library.oregonstate.edu*. Retrieved 18 May 2023.
207. Schwartz J (2008). *In pursuit of the gene: from Darwin to DNA* (<https://archive.org/details/inpursuitofgenef00schw>). Cambridge, Mass.: Harvard University Press. ISBN 978-0-674-02670-4.
208. Pauling L, Corey RB (February 1953). "A Proposed Structure For The Nucleic Acids" (<http://scarc.library.oregonstate.edu/coll/pauling/dna/papers/1953p.9-084.html>). *Proceedings of the National Academy of Sciences of the United States of America*. **39** (2): 84–97. Bibcode:1953PNAS...39...84P (<https://ui.adsabs.harvard.edu/abs/1953PNAS...39...84P>). doi:10.1073/pnas.39.2.84 (<https://doi.org/10.1073%2Fpnas.39.2.84>). PMC 1063734 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1063734>). PMID 16578429 (<https://pubmed.ncbi.nlm.nih.gov/16578429>).
209. Regis E (2009). *What Is Life?: investigating the nature of life in the age of synthetic biology*. Oxford: Oxford University Press. p. 52. ISBN 978-0-19-538341-6.
210. "Double Helix of DNA: 50 Years" (<https://web.archive.org/web/20150405140401/http://www.nature.com/nature/dna50/archive.html>). *Nature Archives*. Archived from the original (<http://www.nature.com/nature/dna50/archive.html>) on 5 April 2015.

211. "Original X-ray diffraction image" (<http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/dna/pictures/franklin-typeBphoto.html>). Oregon State Library. Archived (<https://web.archive.org/web/20090130111849/http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/dna/pictures/franklin-typeBphoto.html>) from the original on 30 January 2009. Retrieved 6 February 2011.
212. Burakoff M (25 April 2023). "Rosalind Franklin's role in DNA discovery gets a new twist" (<https://apnews.com/article/dna-double-helix-roosalind-franklin-watson-crick-69ec8164c720e0b23374da69a1d3708d>). *AP News*. Retrieved 25 April 2023.
213. Anthes E (25 April 2023). "Untangling Rosalind Franklin's Role in DNA Discovery, 70 Years On – Historians have long debated the role that Dr. Franklin played in identifying the double helix. A new opinion essay argues that she was an "equal contributor." " (<https://www.nytimes.com/2023/04/25/science/rosalind-franklin-dna.html>). *The New York Times*. Archived (<https://archive.today/20230425182515/https://www.nytimes.com/2023/04/25/science/rosalind-franklin-dna.html>) from the original on 25 April 2023. Retrieved 26 April 2023.
214. Cobb M, Comfort N (25 April 2023). "What Rosalind Franklin truly contributed to the discovery of DNA's structure – Franklin was no victim in how the DNA double helix was solved. An overlooked letter and an unpublished news article, both written in 1953, reveal that she was an equal player" (<https://doi.org/10.1038%2Fd41586-023-01313-5>). *Nature*. **616** (7958): 657–660. Bibcode:2023Natur.616..657C (<https://ui.adsabs.harvard.edu/abs/2023Natur.616..657C>). doi:10.1038/d41586-023-01313-5 (<https://doi.org/10.1038%2Fd41586-023-01313-5>). PMID 37100935 (<https://pubmed.ncbi.nlm.nih.gov/37100935>). S2CID 258314143 (<https://api.semanticscholar.org/CorpusID:258314143>).
215. "The Nobel Prize in Physiology or Medicine 1962" (http://nobelprize.org/nobel_prizes/medicine/laureates/1962/). *Nobelprize.org*.
216. Maddox B (January 2003). "The double helix and the 'wronged heroine' " (http://www.biomath.nyu.edu/index/course/hw_articles/nature4.pdf) (PDF). *Nature*. **421** (6921): 407–08. Bibcode:2003Natur.421..407M (<https://ui.adsabs.harvard.edu/abs/2003Natur.421..407M>). doi:10.1038/nature01399 (<https://doi.org/10.1038%2Fnature01399>). PMID 12540909 (<https://pubmed.ncbi.nlm.nih.gov/12540909>). S2CID 4428347 (<https://api.semanticscholar.org/CorpusID:4428347>). Archived (https://web.archive.org/web/20161017011403/http://www.biomath.nyu.edu/index/course/hw_articles/nature4.pdf) (PDF) from the original on 17 October 2016.
217. Crick FH (1955). *A Note for the RNA Tie Club* (<https://web.archive.org/web/20081001223217/http://genome.wellcome.ac.uk/assets/wtx030893.pdf>) (PDF) (Speech). Cambridge, England. Archived from the original (<http://genome.wellcome.ac.uk/assets/wtx030893.pdf>) (PDF) on 1 October 2008.
218. Meselson M, Stahl FW (July 1958). "The Replication of DNA in Escherichia Coli" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC528642>). *Proceedings of the National Academy of Sciences of the United States of America*. **44** (7): 671–82. Bibcode:1958PNAS...44..671M (<https://ui.adsabs.harvard.edu/abs/1958PNAS...44..671M>). doi:10.1073/pnas.44.7.671 (<https://doi.org/10.1073%2Fpnas.44.7.671>). PMC 528642 (<https://www.ncbi.nlm.nih.gov/pmc/article/s/PMC528642>). PMID 16590258 (<https://pubmed.ncbi.nlm.nih.gov/16590258>).
219. "The Nobel Prize in Physiology or Medicine 1968" (http://nobelprize.org/nobel_prizes/medicine/laureates/1968/). *Nobelprize.org*.
220. Pray L (2008). "Discovery of DNA structure and function: Watson and Crick". *Nature Education*. **1** (1): 100.
221. Panneerchelvam S, Norazmi MN (2003). "Forensic DNA Profiling and Database" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3561883>). *The Malaysian Journal of Medical Sciences*. **10** (2): 20–26. PMC 3561883 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3561883>). PMID 23386793 (<https://pubmed.ncbi.nlm.nih.gov/23386793>).

Further reading

- Berry A, Watson J (2003). *DNA: the secret of life* (<https://archive.org/details/dnasecretoflife00wats>). New York: Alfred A. Knopf. ISBN 0-375-41546-7.
- Calladine CR, Drew HR, Luisi BF, Travers AA (2003). *Understanding DNA: the molecule & how it works*. Amsterdam: Elsevier Academic Press. ISBN 0-12-155089-3.
- Carina D, Clayton J (2003). *50 years of DNA* (<https://archive.org/details/50yearsofdna00clay>). Basingstoke: Palgrave Macmillan. ISBN 1-4039-1479-6.
- Judson HF (1979). *The Eighth Day of Creation: Makers of the Revolution in Biology* (2nd ed.). Cold Spring Harbor Laboratory Press. ISBN 0-671-22540-5.
- Olby RC (1994). *The path to the double helix: the discovery of DNA*. New York: Dover Publications. ISBN 0-486-68117-3. First published in October 1974 by MacMillan, with foreword by Francis Crick; the definitive DNA textbook, revised in 1994 with a nine-page postscript.
- Olby R (January 2003). "Quiet debut for the double helix" (<https://doi.org/10.1038%2Fnature01397>). *Nature*. **421** (6921): 402–05. Bibcode:2003Natur.421..402O (<https://ui.adsabs.harvard.edu/abs/2003Natur.421..402O>). doi:10.1038/nature01397 (<https://doi.org/10.1038%2Fnature01397>). PMID 12540907 (<https://pubmed.ncbi.nlm.nih.gov/12540907>).
- Olby RC (2009). *Francis Crick: A Biography*. Plainview, N.Y: Cold Spring Harbor Laboratory Press. ISBN 978-0-87969-798-3.
- Micklas D (2003). *DNA Science: A First Course*. Cold Spring Harbor Press. ISBN 978-0-87969-636-8.
- Ridley M (2006). *Francis Crick: discoverer of the genetic code*. Ashland, OH: Eminent Lives, Atlas Books. ISBN 0-06-082333-X.
- Rosenfeld I (2010). *DNA: A Graphic Guide to the Molecule that Shook the World*. Columbia University Press. ISBN 978-0-231-14271-7.
- Schultz M, Cannon Z (2009). *The Stuff of Life: A Graphic Guide to Genetics and DNA* (<http://archive.org/details/stuffoflifegraph00schu>). Hill and Wang. ISBN 978-0-8090-8947-5.
- Stent GS, Watson J (1980). *The Double Helix: A Personal Account of the Discovery of the Structure of DNA* (https://archive.org/details/doublehelixpers00wats_0). New York: Norton. ISBN 0-393-95075-1.
- Watson J (2004). *DNA: The Secret of Life*. Random House. ISBN 978-0-09-945184-6.
- Wilkins M (2003). *The third man of the double helix the autobiography of Maurice Wilkins*. Cambridge, England: University Press. ISBN 0-19-860665-6.

External links

- DNA binding site prediction on protein (<https://web.archive.org/web/20070306082905/http://pipe.scs.fsu.edu/displar.html>)
- DNA the Double Helix Game (https://web.archive.org/web/20100223035803/http://nobelprize.org/educational_games/medicine/dna_double_helix/) From the official Nobel Prize web site
- DNA under electron microscope (http://www.fidelitysystems.com/Unlinked_DNA.html)
- Dolan DNA Learning Center (<http://www.dnalc.org/>)
- Double Helix: 50 years of DNA (<http://www.nature.com/nature/dna50/archive.html>), *Nature*
- *Proteopedia* DNA (<http://www.proteopedia.org/wiki/index.php/DNA>)
- *Proteopedia* Forms of DNA (http://www.proteopedia.org/wiki/index.php/Forms_of_DNA)

- ENCODE threads explorer (<http://www.nature.com/encode/>) ENCODE home page at Nature
- Double Helix 1953–2003 (<https://web.archive.org/web/20070213030135/http://www.ncbe.reading.ac.uk/DNA50/>) National Centre for Biotechnology Education
- Genetic Education Modules for Teachers (<http://www.genome.gov/10506718>) – *DNA from the Beginning* Study Guide
- PDB Molecule of the Month *DNA* (<https://pdb101.rcsb.org/motm/23>)
- "Clue to chemistry of heredity found" (<https://www.nytimes.com/packages/pdf/science/dna-article.pdf>). *The New York Times*, June 1953. First American newspaper coverage of the discovery of the DNA structure
- DNA from the Beginning (<http://www.dnafb.org/>) Another DNA Learning Center site on DNA, genes, and heredity from Mendel to the human genome project.
- The Register of Francis Crick Personal Papers 1938 – 2007 (<https://web.archive.org/web/20070825101712/http://orpheus.ucsd.edu/speccoll/testing/html/mss0660a.html#abstract>) at Mandeville Special Collections Library, University of California, San Diego
- Seven-page, handwritten letter that Crick sent to his 12-year-old son Michael in 1953 describing the structure of DNA. (http://www.nature.com/polopoly_fs/7.9746!/file/Crick%20letter%20to%20Michael.pdf) See *Crick's medal goes under the hammer* (<http://www.nature.com/news/crick-s-medal-goes-under-the-hammer-1.12705>), Nature, 5 April 2013.

Retrieved from "<https://en.wikipedia.org/w/index.php?title=DNA&oldid=1268524198>"