The Concept of the Gene

Stavros Ioannidis
MA Cont. Phil.

Key Questions:

i) what is a gene?

ii) where did the gene concept come from?

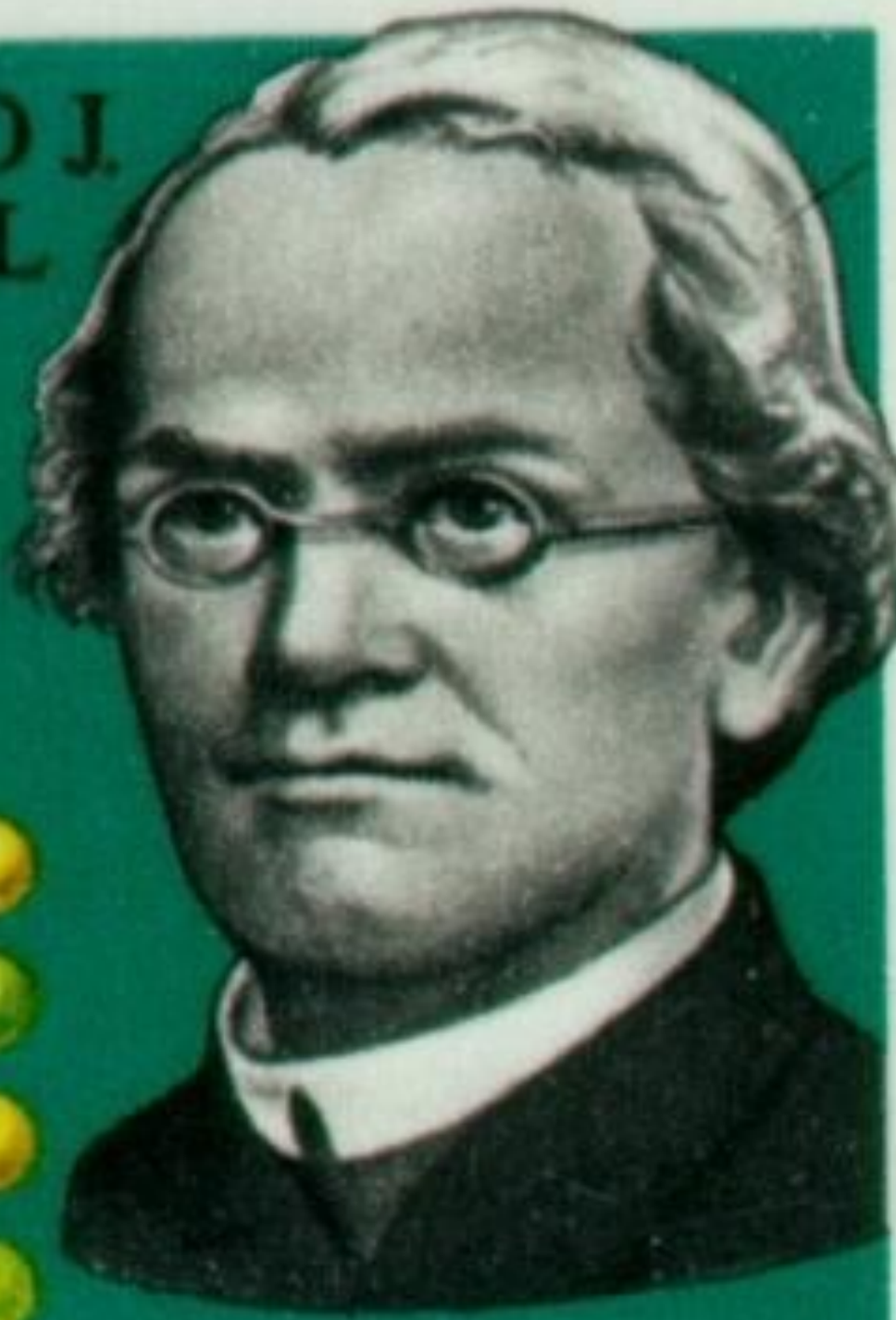
iii) is there more than one gene concept in modern biology?

iv) what is the relation between classical and molecular genetics?

v) does the notion of ‘genetic information’ make sense?

GREGORIO J.
MENDEL

1884
1984



450 POSTE VATICANE

M. CODONI

I.P.Z.S.-ROMA-1984

Mendelian Genetics

-1860s: Gregor Mendel experimented with **pea-plants**
-two states of a single **character**: round versus wrinkled seeds
-pattern of transmission across generations
-crossed true-breeding lineages and found:

i) **F1 hybrids** all had round seeds
ii) **F2 hybrids** had 3/4 round but 1/4 wrinkled

-explanation?

-trait is controlled by **pairs of factors**
-each factor is **one of two possible types**
-**gametes** of an organism (pollen or ova) contain just one of the factors
-new organism **inherits one factor from each parent**
-factor for round seeds **dominates** factor for wrinkled seeds
-when an F1 hybrid forms gametes, **50%** of them contain factor for wrinkled, **50%** for round

-perfectly explains patterns of inheritance observed

-Mendel's 'factors' are now called '**genes**' or '**alleles**' 'allelomorphs'

-they are **units of inheritance**

Mendelian Genetics

Mendel's Laws

1)*Laws of Independent Segregation of Alleles*:
for each individual, each of the two alleles at a locus segregates to a germ cell with equal probability

2) *Law of Independent Assortment* (of Alleles at Different Loci):
when the germ cells are formed, the allele pairs at different loci assort independently of each other

-Mendel illustrated (2) by looking at two different traits simultaneously:
seed shape (round/wrinkled) and flower colour (yellow/green)

-key observation: if a gamete contains factor for **wrinkled**, that tells us nothing about whether it contains factor for **yellow** or **green**

-> 1) is almost always true, 2) regularly violated

Blending inheritance

Red and yellow make orange

Verhandlungen

des

naturforschenden Vereines

in Brünn.

IV. Band

1865.

Brünn, 1866.

Im Verlage des Vereines.

Versuche über Pflanzen-Hybriden.

Von

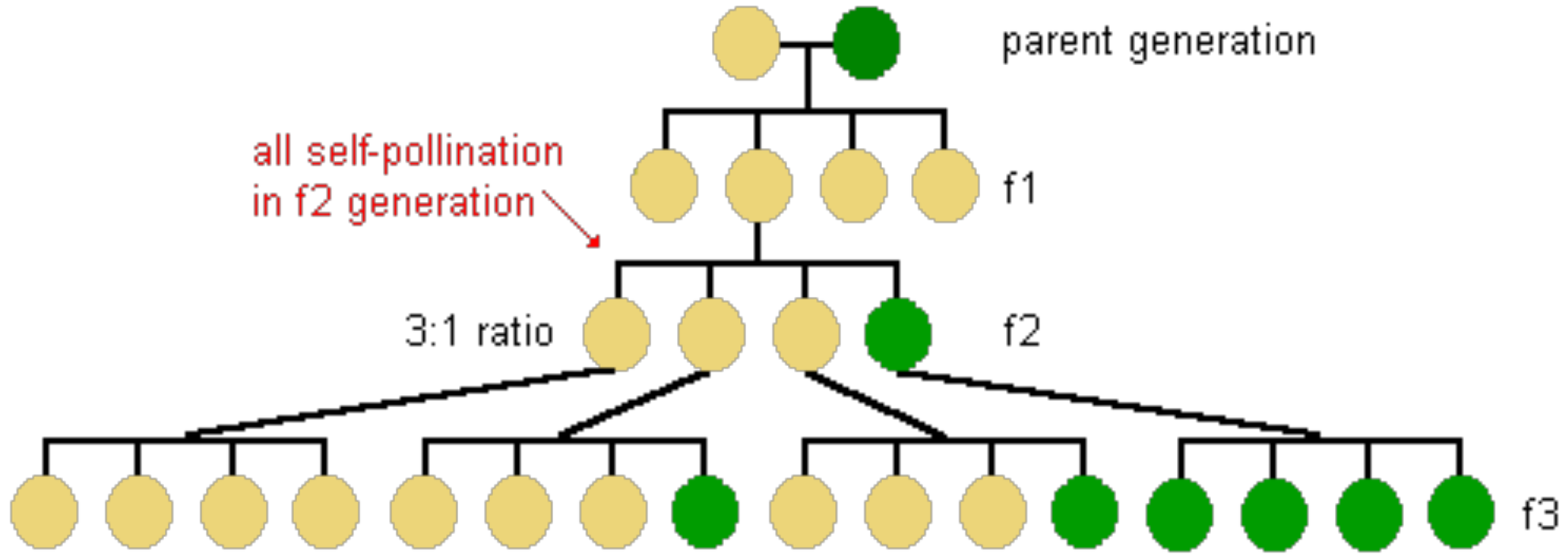
Gregor Mendel.

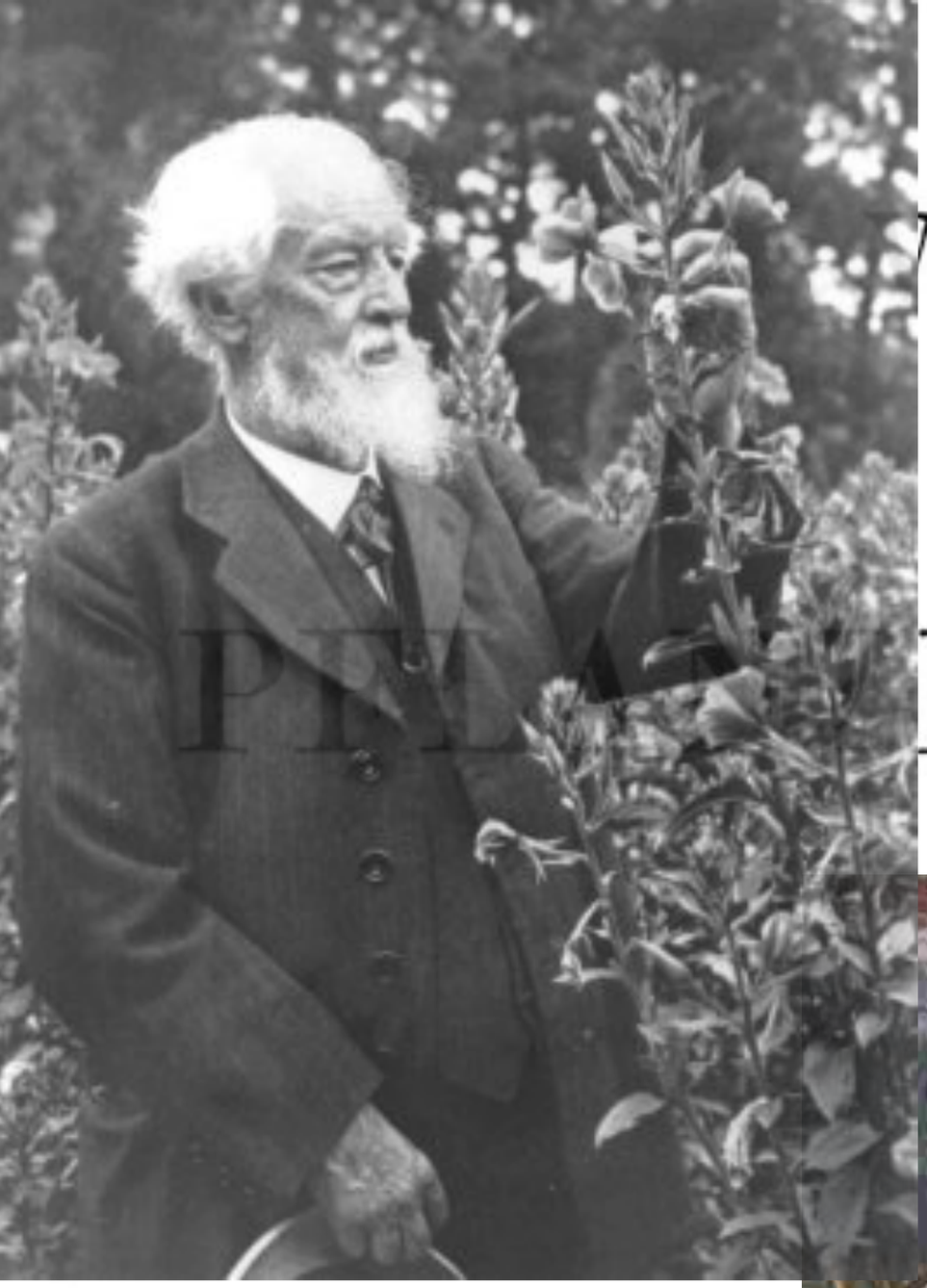
(Vorgelegt in den Sitzungen vom 8. Februar und 8. März 1865.)

Einleitende Bemerkungen.

Künstliche Befruchtungen, welche an Zierpflanzen desshalb vorgenommen wurden, um neue Farben-Varianten zu erzielen, waren die Veranlassung zu den Versuchen, die her besprochen werden sollen. Die auffallende Regelmässigkeit, mit welcher dieselben Hybridformen immer wiederkehrten, so oft die Befruchtung zwischen gleichen Arten geschah, gab die Anregung zu weiteren Experimenten, deren Aufgabe es war, die Entwicklung der Hybriden in ihren Nachkommen zu verfolgen.

Dieser Aufgabe haben sorgfältige Beobachter, wie Kölreuter, Gärtner, Herbert, Lecocq, Wichura u. a. einen Theil ihres Lebens mit unermüdlicher Ausdauer geopfert. Namentlich hat Gärtner in seinem Werke „die Bastarderzeugung im Pflanzenreiche“ sehr schätzbare Beobachtungen niedergelegt, und in neuester Zeit wurden von Wichura gründliche Untersuchungen über die Bastarde der Weiden veröffentlicht. Wenn es noch nicht gelungen ist, ein allgemein giltiges Gesetz für die Bildung und Entwicklung der Hybriden aufzustellen, so kann das Niemanden Wunder nehmen, der den Umfang der Aufgabe kennt und die Schwierigkeiten zu würdigen weiss, mit denen Versuche dieser Art zu kämpfen haben. Eine endgiltige Entscheidung kann erst dann erfolgen, bis Detail-Versuche aus den verschiedensten Pflanzen-Familien vorliegen. Wer die Ar-


parent generation
all self-pollination
in f2 generation
f1
3:1 ratio
f2
f3
3:1 ratio
3:1 ratio
parent generation
Y = yellow allele
G = green allele
YY GG
YG YG YG YG
f1 generation
YY YG YG GG
f2 generation
YY YY YY YY YY YG YG GG YY YG YG GG GG GG GG GG



VERSUCHE

ÜBER

ENHYBI

Top: de Vries, Correns
Bottom: Tschermak

Gen. 37

GREGOR MENDEL, 1866

Enlarged from a group of the brethren of the Königskloster

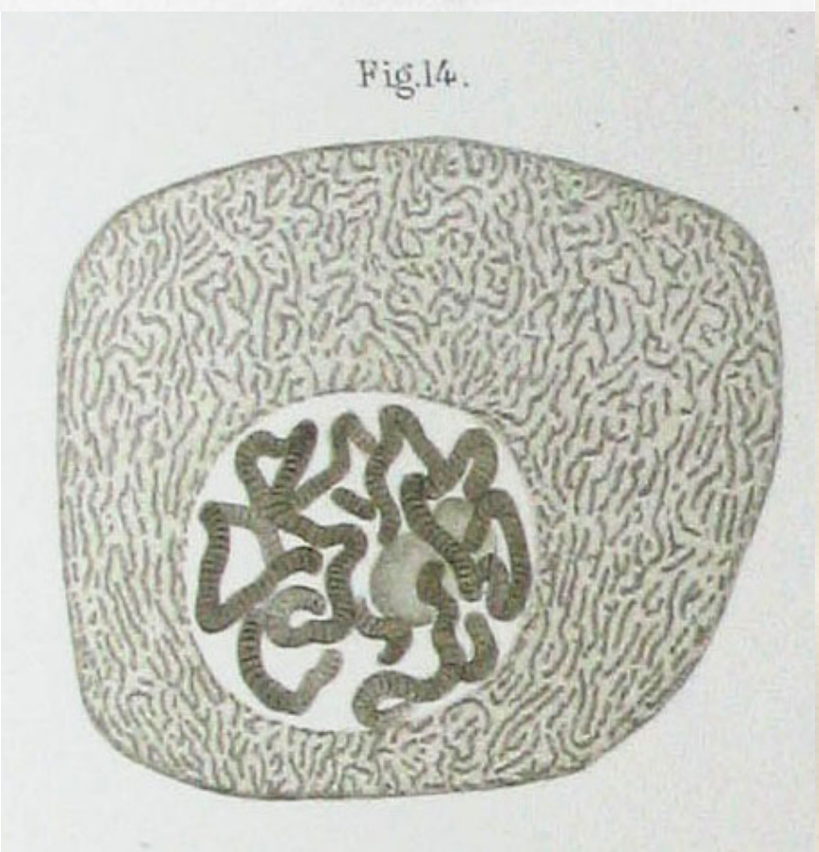
MENDEL'S
PRINCIPLES OF HEREDITY

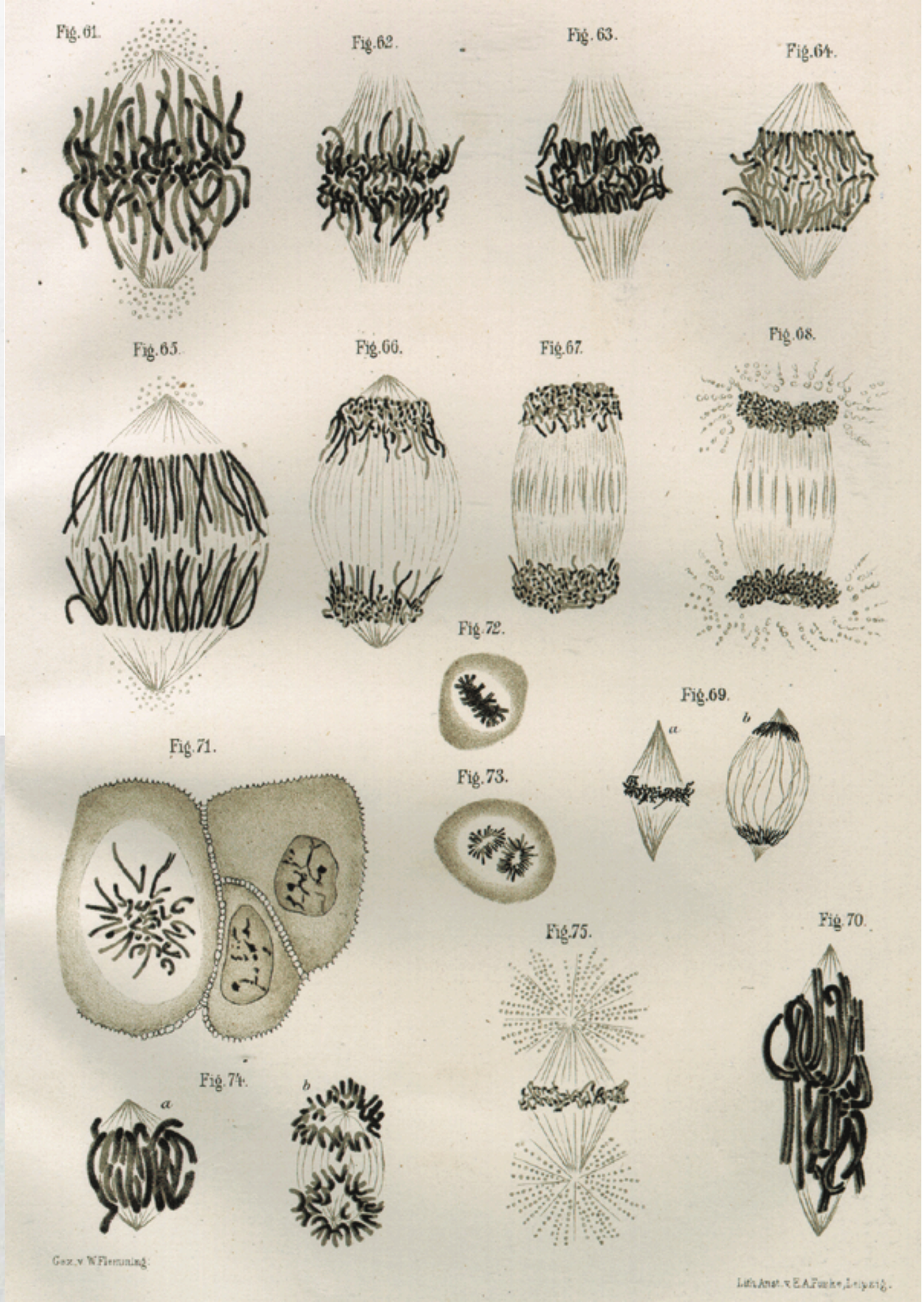
BY

W. BATESON, M.A., F.R.S., V.M.H.

HONORARY FELLOW OF ST JOHN'S COLLEGE,
DIRECTOR OF THE JOHN INNES HORTICULTURAL INSTITUTION

Cambridge:
at the University Press
1913





Left: Walther Flemming *Bottom*: Polytene chromosomes from a cell from the salivary glands of *Chironimus*

Right: the process of mitosis, from Flemming *Zellsubstanz, Kern und Zelltheilung* (1885)

T. H. Morgan

Classical Genetics

-Mendel's work forgotten and 're-discovered' around turn of the century

-after a long struggle, eventually came to be accepted
(-> opposition between **Mendelians** and **biometricians**)

-research program of T. H. Morgan at Columbia University
-two main aspects:

i) **segregation** analysis

ii) **linkage** analysis

-segregation analysis extended Mendel's own techniques
-used to infer genetic origin of a trait

Classical Genetics

-linkage analysis based on recognition that **Law of Independent Assortment** is **often violated**
-observed that some loci tend to be **inherited together**
-such loci were in the same ***linkage group***
-why? Morgan (1910): because they are on the **same chromosome**
-> genes on same chromosome violate Mende's second law

-complicated by ***crossing-over***
-> exchange of parts between homologous chromosomes during meiosis
-means that alleles at loci on same chromosome may be **separated**, hence assort **independently**
-this is called '**recombination**' of genes
-depends on ***degree of linkage***

-**linkage analysis**: start with a trait known to satisfy **Mendel's first law**
-then, look for a trait that is **linked** to it
-then, try to estimate **degree of linkage**

-led to ***linkage mapping***: trying to map entire chromosomes
-**loci** placed in **different linkage groups** with a **distance** assigned between each pair of loci

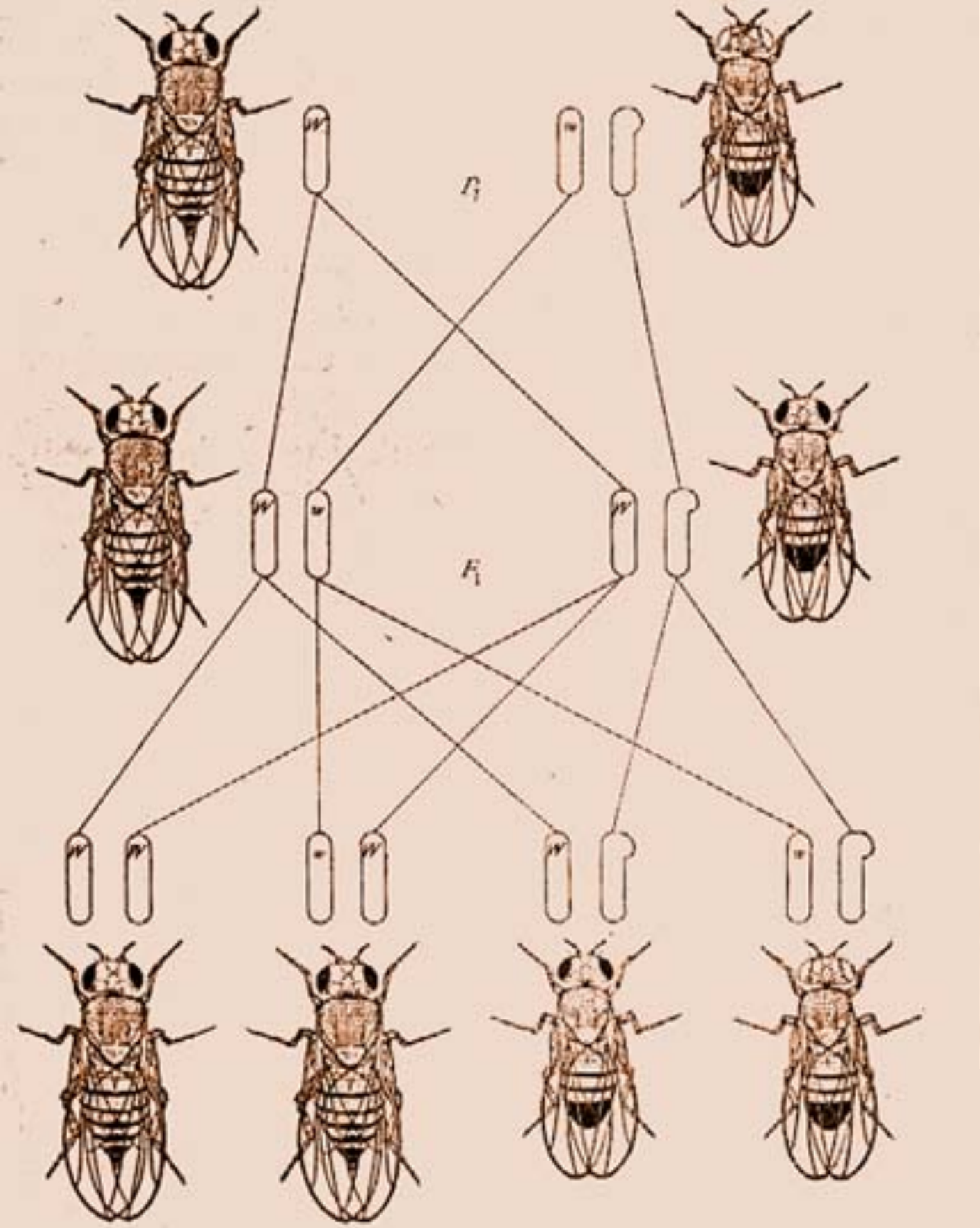
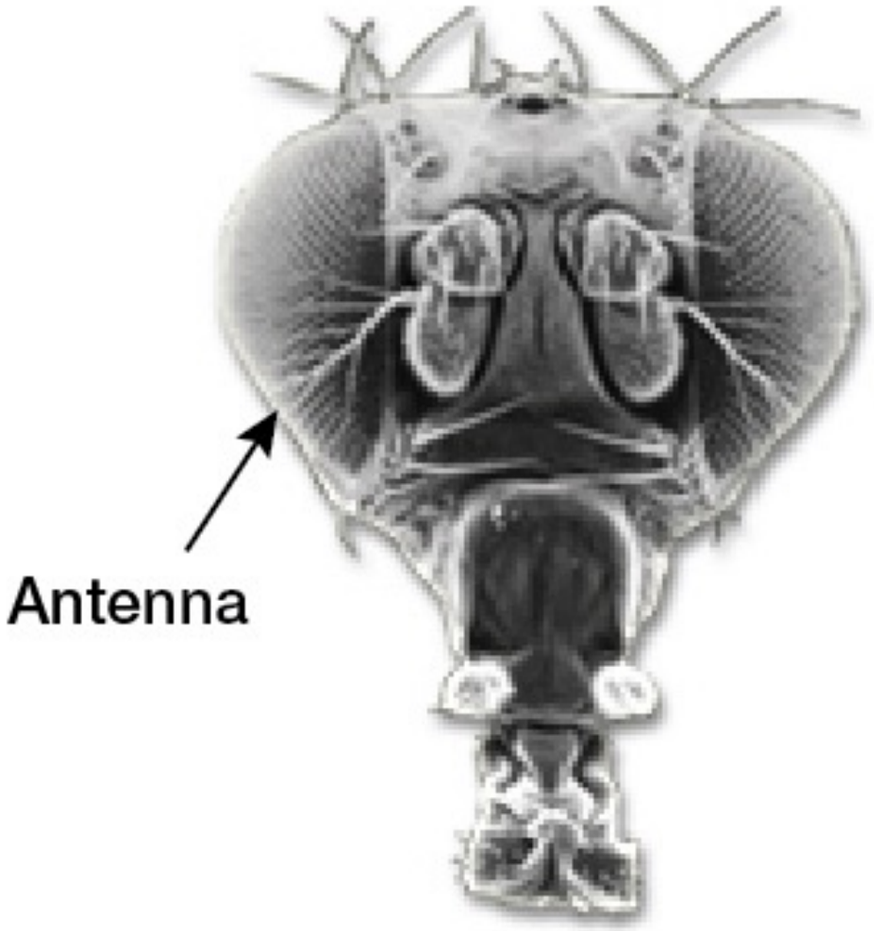


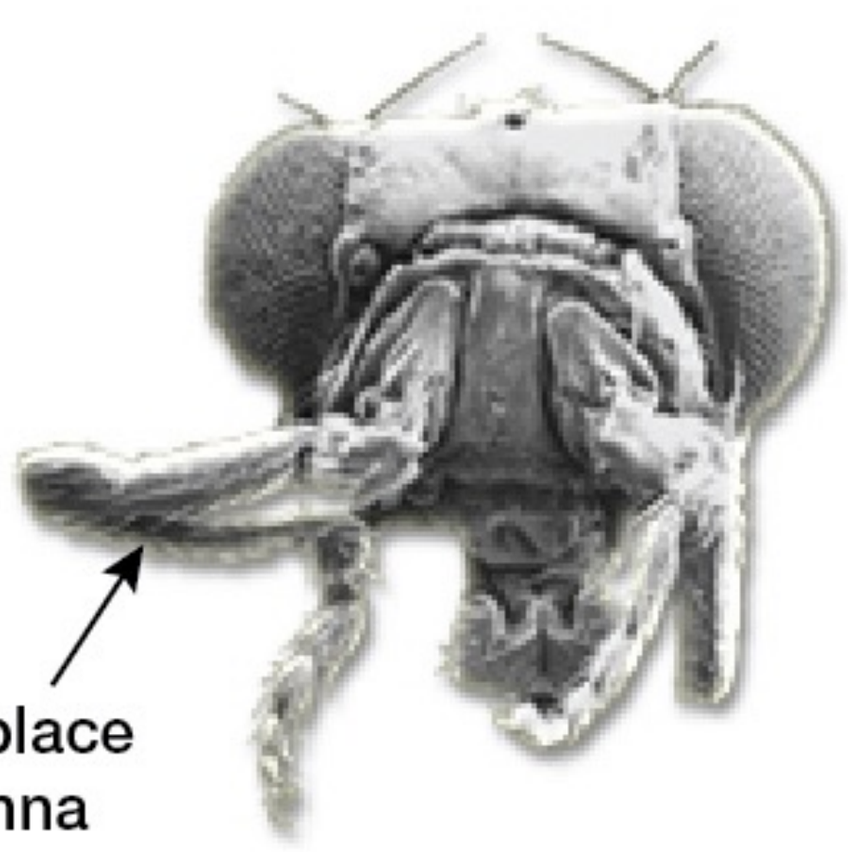
Fig. 9.—Red-eyed female by white-eyed male (D. melanogaster). This is the reciprocal of the cross shown in Fig. 10.

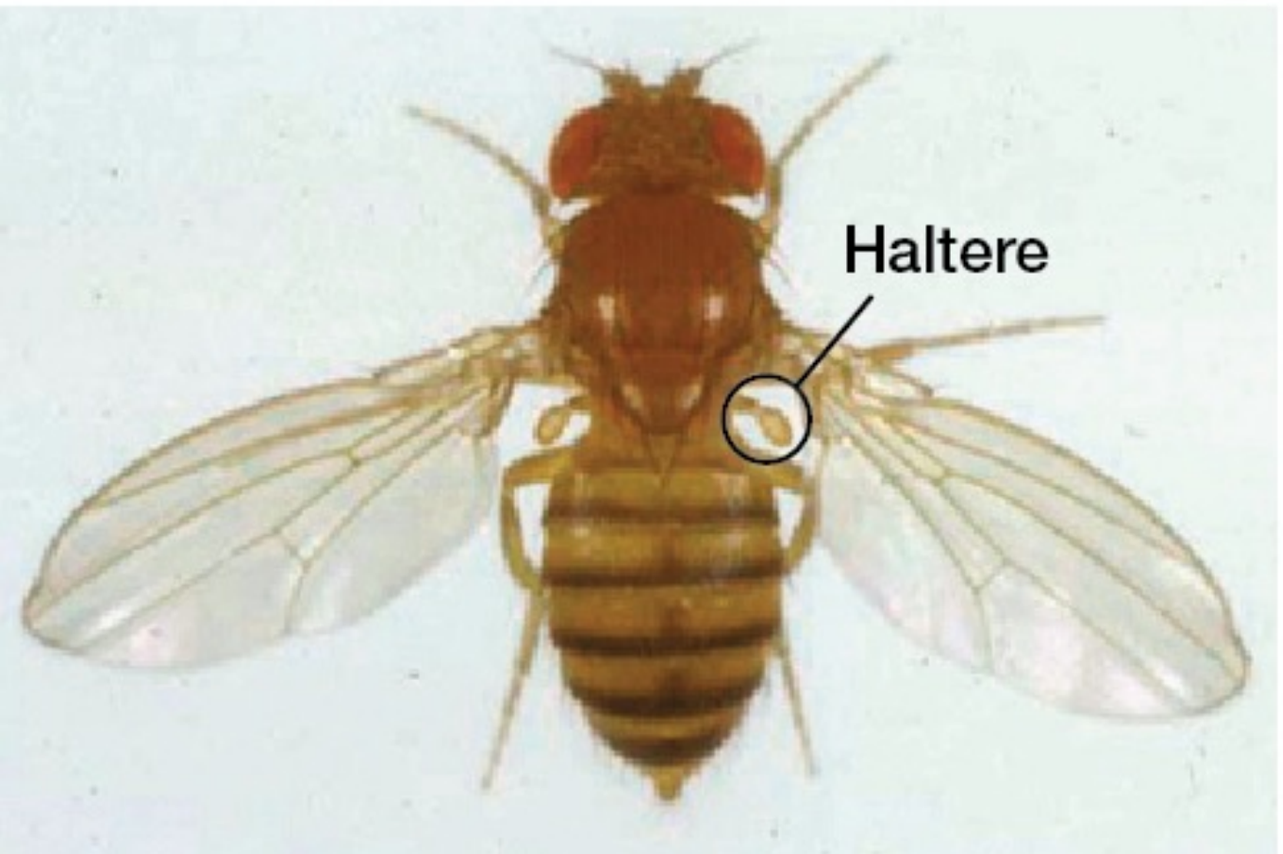
Left: From Morgan, Sturtevant, Muller & Bridges (1915) *The Mechanism of Mendelian Heredity*
Top: Morgan and drosophilas

White-eyed mutant fly

Red-eyed wild-type fly


Antenna


Leg in place
of antenna


Haltere

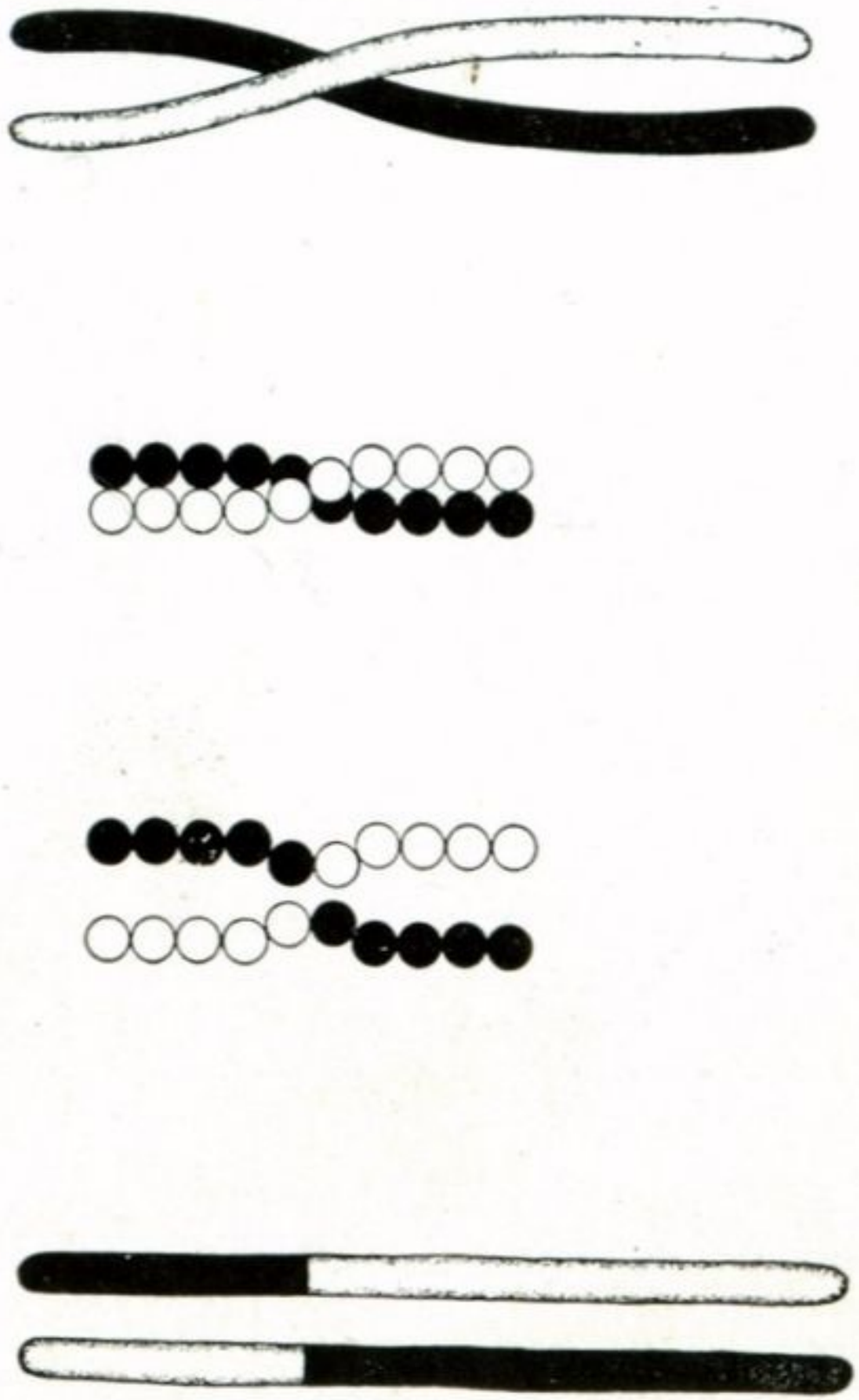
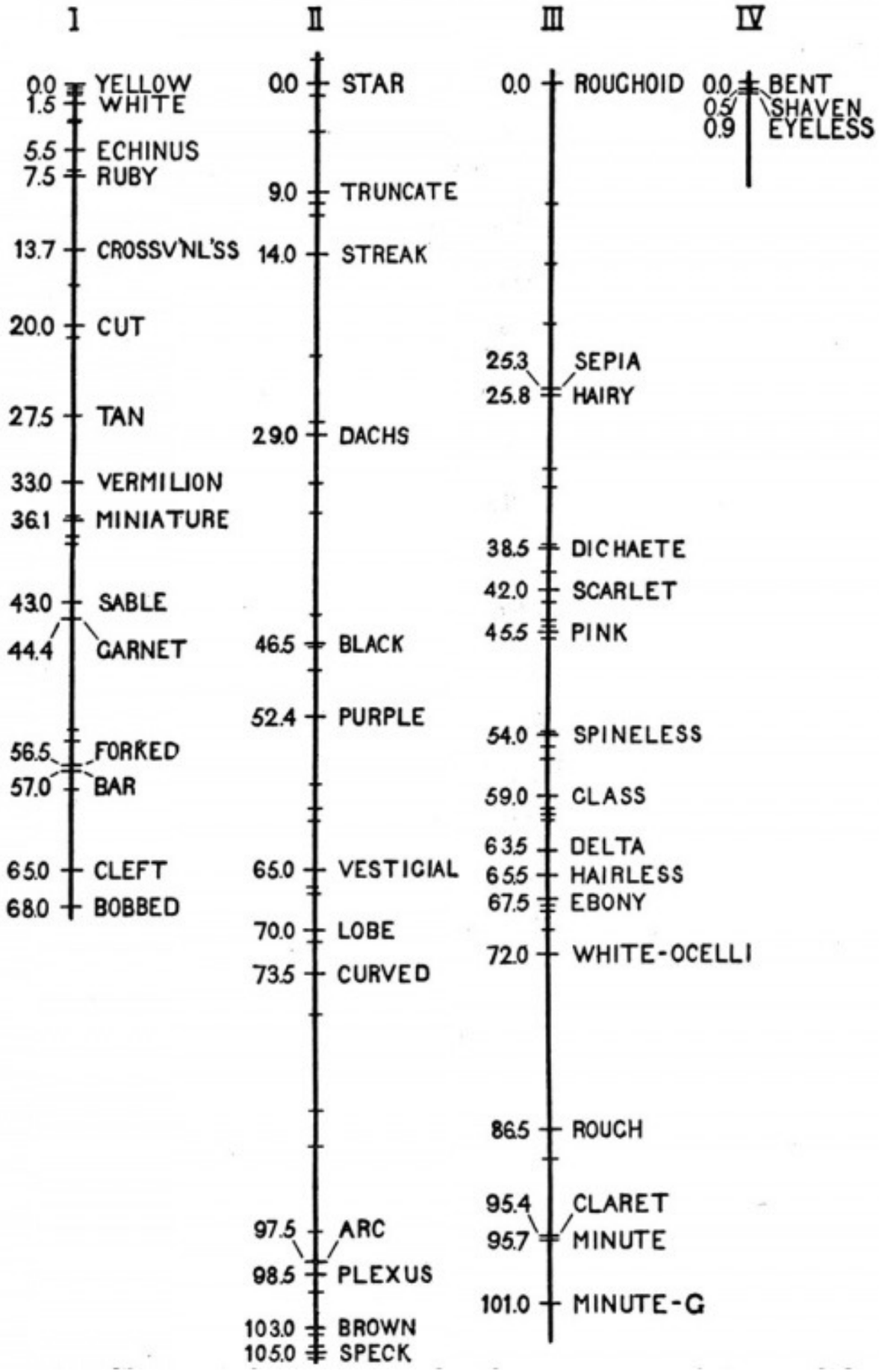
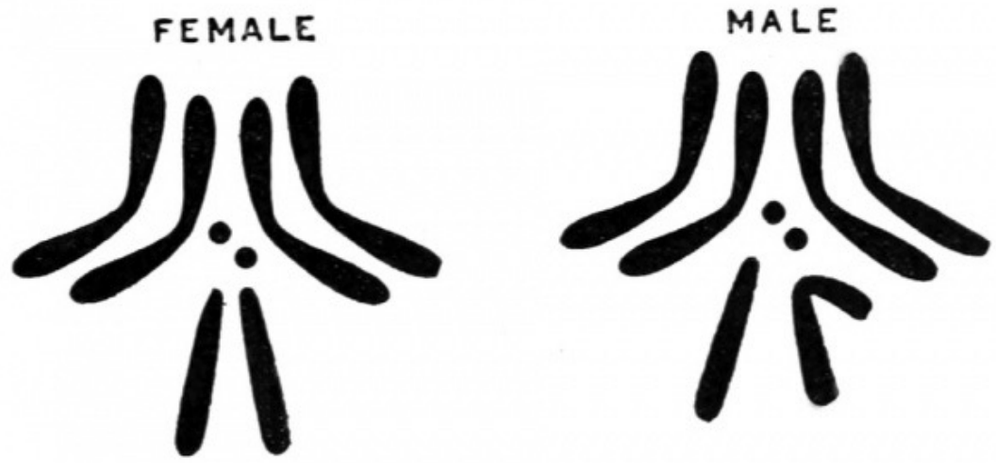


FIG. 64. Scheme to illustrate a method of crossing over of the chromosomes.





Left: Drosophila gene map
Top: the 4 pairs of chromosomes

MUTANTS

short aristae

dumpy wings

short legs

black body

purple eyes

vestigial wings

brown eyes

0 — 13.0 — 31.0 — 48.5 — 54.5 — 67.0 — 104.0

chromosome II

long aristae

long wings

long legs

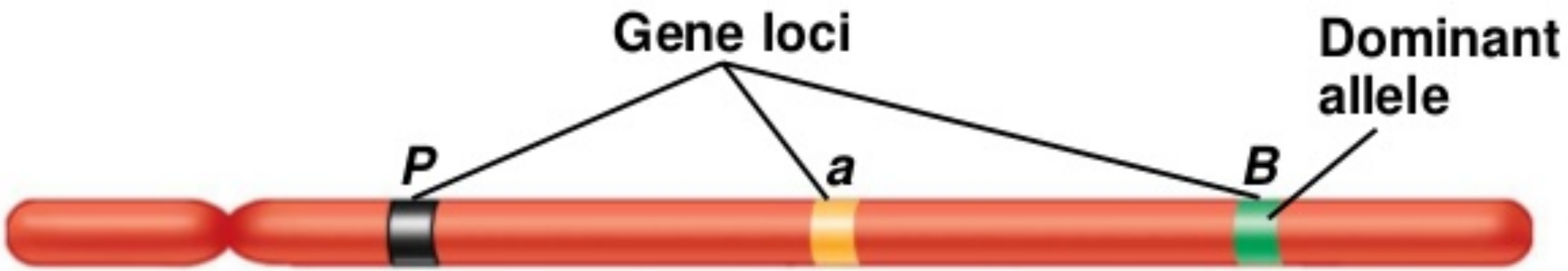
gray body

red eyes

long wings

red eyes

WILD TYPE



Genotype: *PP* *aa* *Bb*

Homozygous for the dominant allele

Homozygous for the recessive allele

Heterozygous

Classical Genetics

Importance of Mendelian/Classical Genetics:

-quite striking **predictive** and **explanatory** success
-integrated with Darwinian principles in the *Modern Synthesis* (Fisher, Haldane, Wright)

-**particulate** nature of inheritance critically important for natural selection
-why? *because it shows that variation can be maintained even with sexual reproduction*
-solved a **major puzzle** for Darwin's theory

-> How many traits are controlled by single Mendelian (or classical) genes?
-Not many
-in some ways, Mendel was **lucky** in his choice of traits
-most traits are affected by **many genes**, and by environmental factors too
-still, often convenient to speak of genes 'for' traits

-central theoretical problem: **how** are genes replicated?
-classical geneticists **had no answer to this**

-> the 'gene' of classical genetic was ***posited*** to explain observed data
-in a sense, the gene was a **hypothetical** entity

-> The classical gene as a difference-maker

Morgan:
Suppose, for instance, to take perhaps an extreme case, **all the genes** are instrumental in producing each organ of the body. This may only mean that they all produce chemical substances essential for the normal course of development

Morgan:
If now **one gene is changed** so that it produces some substance different from that which it produced before, the end-result may be **affected**, and if the change affects **one** organ predominatingly **it may appear the one gene alone has produced this effect**. In a strictly causal sense this is true, but the effect is produced only in conjunction with all the other genes. In other words, they are all still contributing, as before, to the end-result which is different in so far as one of them is different.

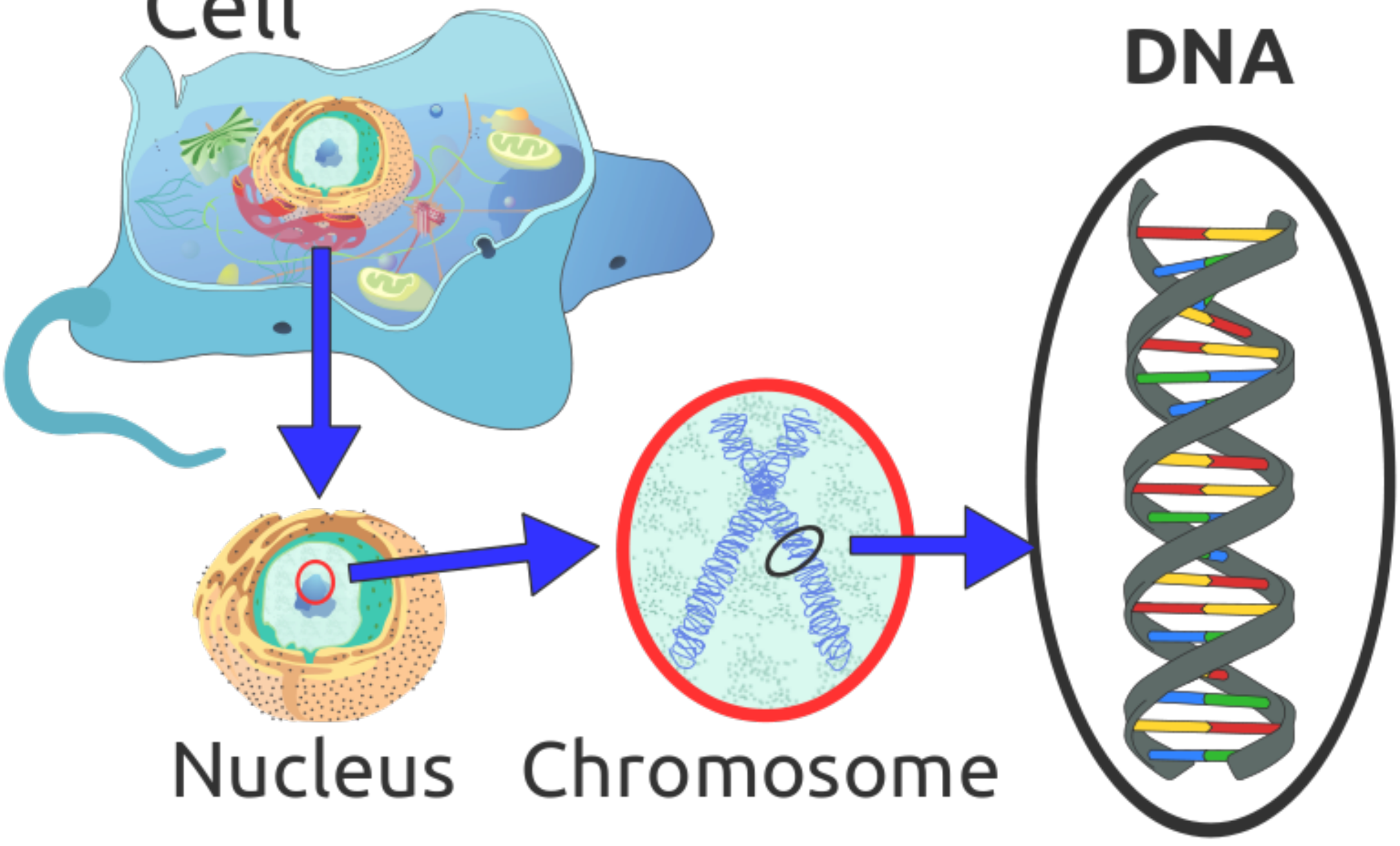
-Waters:
Difference principle: differences in a gene cause uniform phenotypic differences in particular genetic and environmental contexts.

There is **no consensus** of opinion amongst geneticists as to **what the genes are** - whether they are **real or purely fictitious** - because at the level at which the genetic experiments lie, it does not make the **slightest difference** whether the gene is a **hypothetical** unit, or whether the gene is a **material** particle. In either case the unit is associated with a specific chromosome, and can be localized there by purely genetic analysis. Hence, if the gene is a **material** unit, it is a piece of a chromosome; if it is a **fictitious** unit, it must be referred to a definite location in a chromosome - the same place as on the other hypothesis. **Therefore, it makes no difference in the actual work in genetics which point of view is taken**.

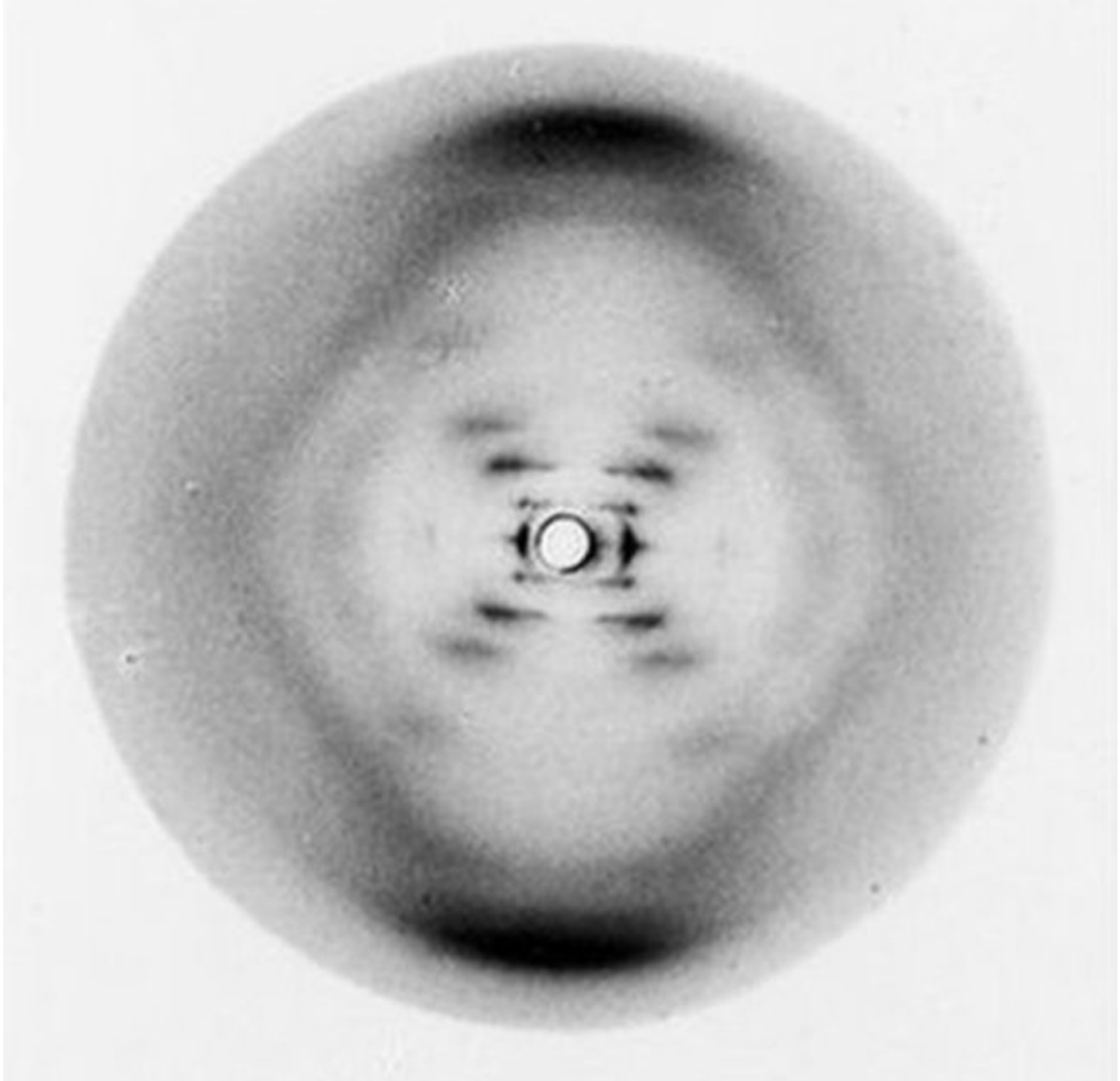
Thomas Hunt Morgan, 1934 (Nobel lecture)

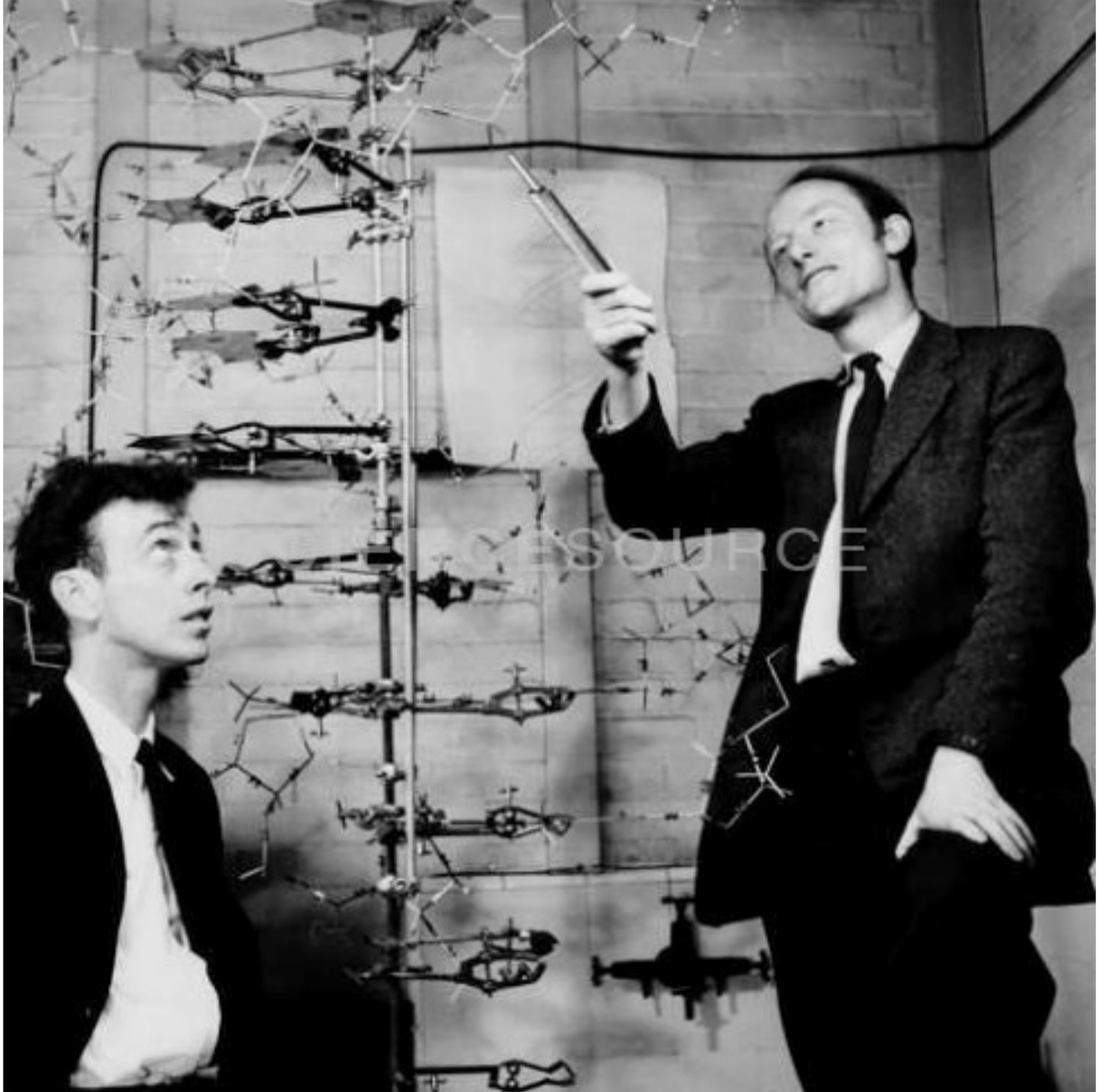
Cell

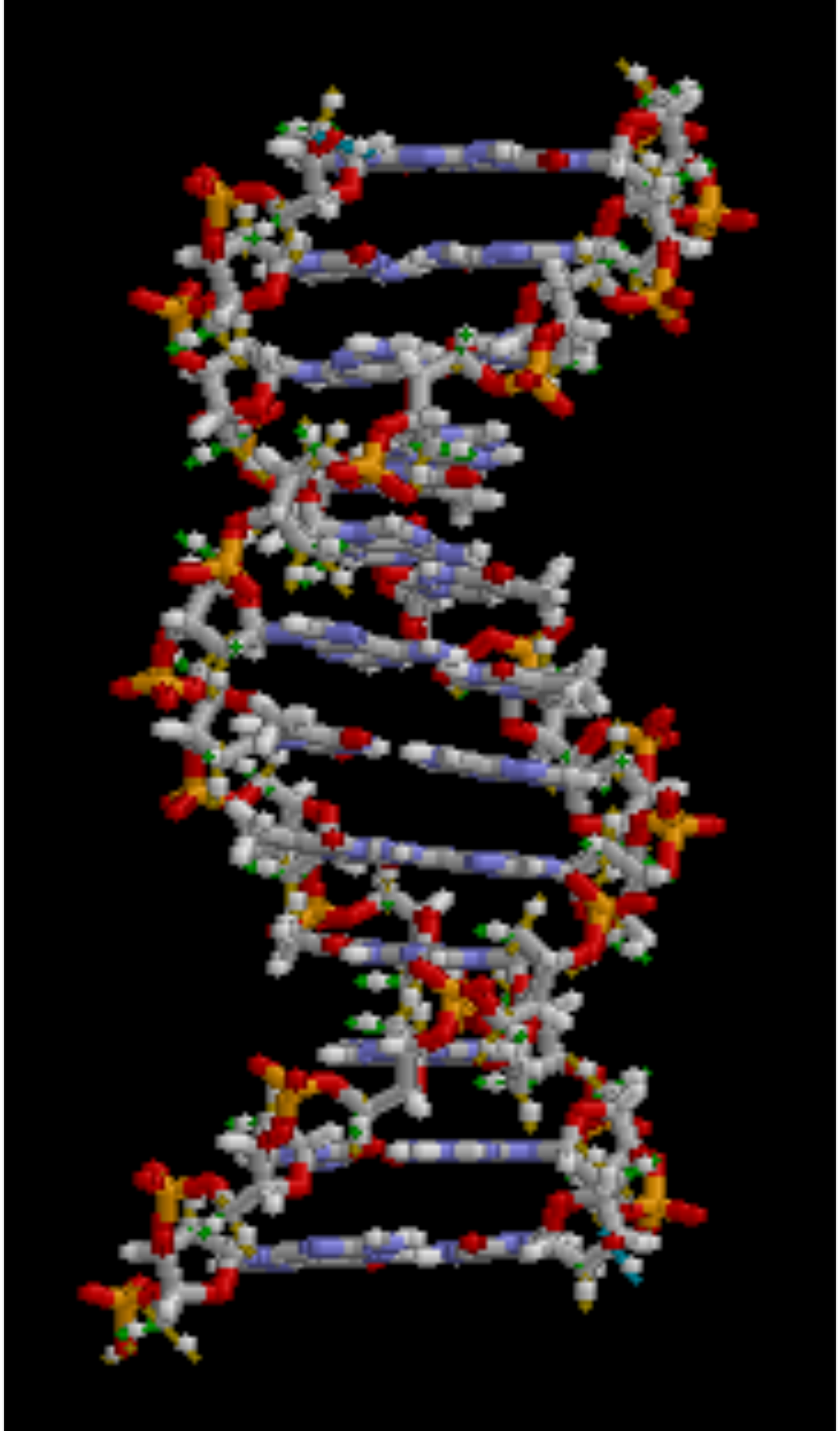
DNA



Nucleus Chromosome







equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik.* (Stockholm), **2** (11) (1905).

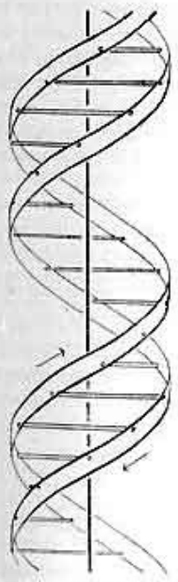
MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

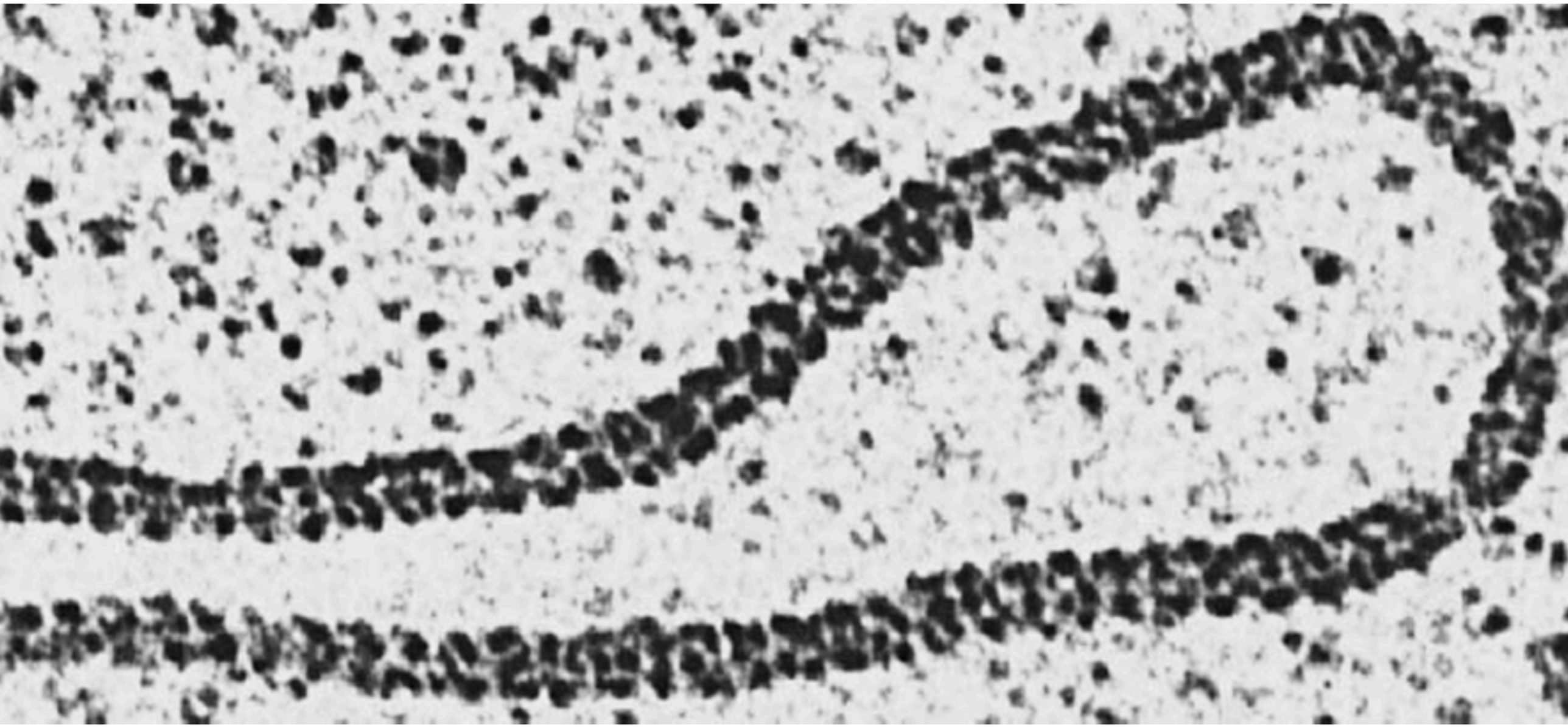
It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

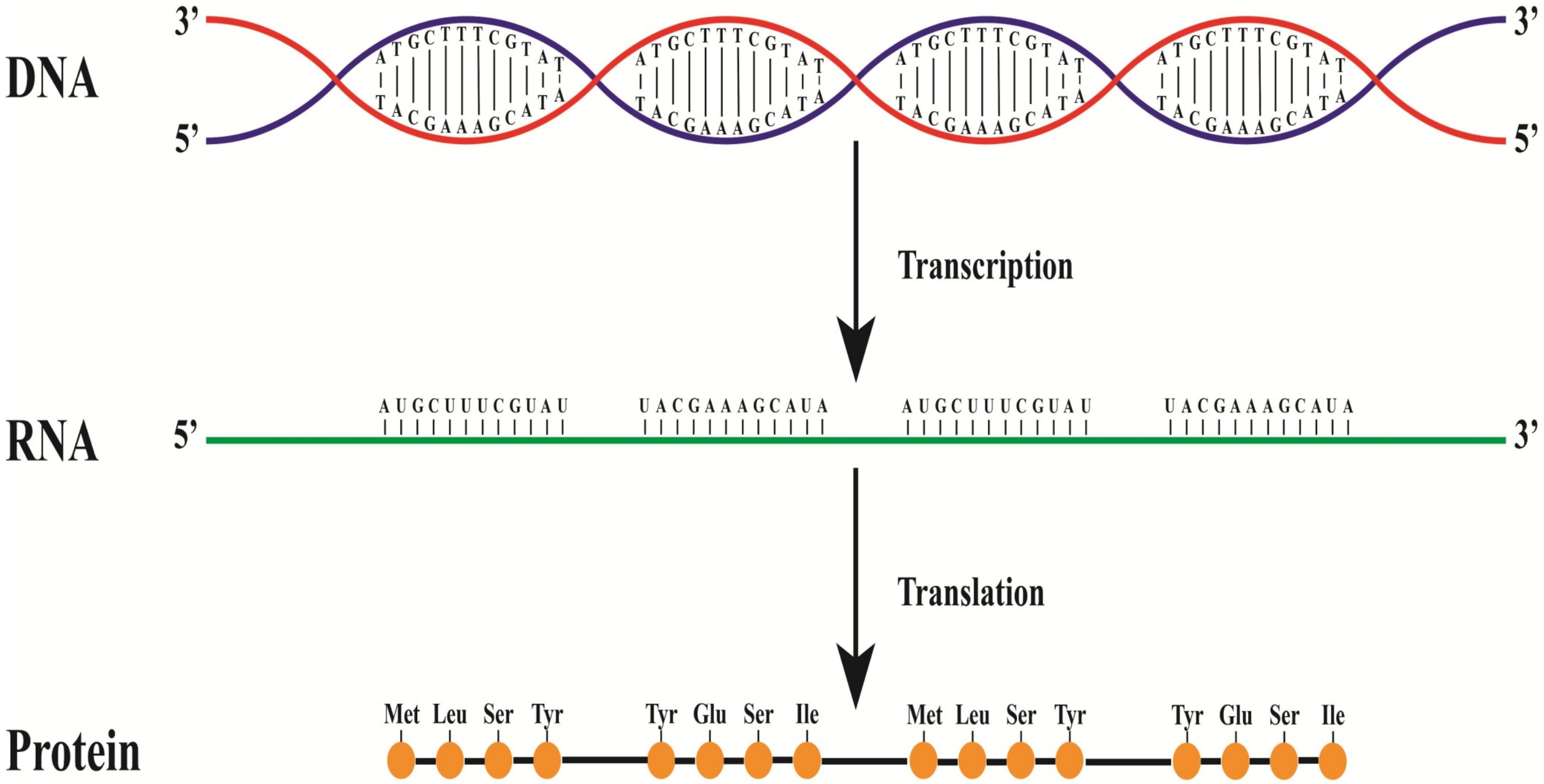
Nobel Prize in Physiology or Medicine (1962) -Crick, Watson & Wilkins

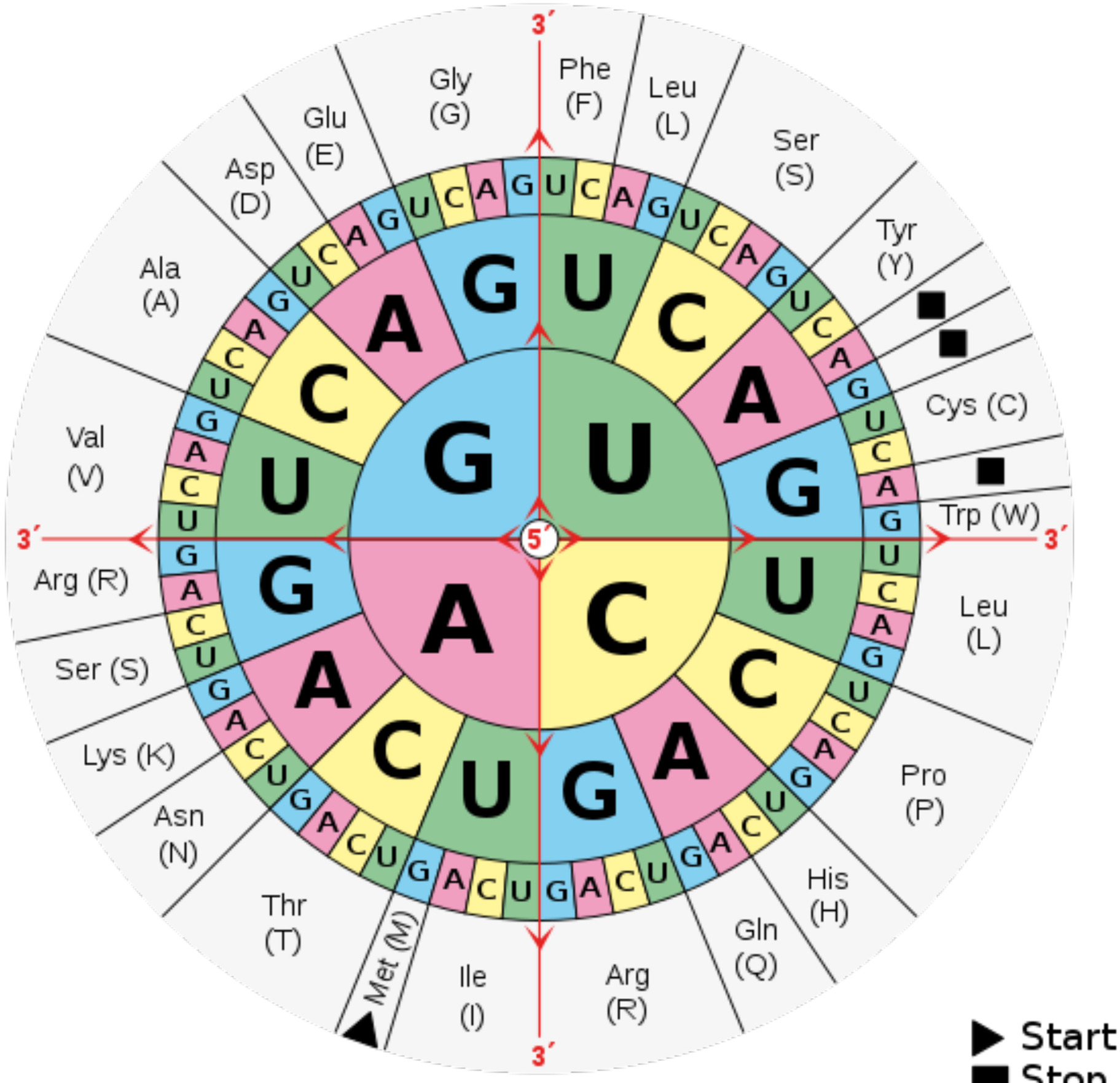
"for their discoveries concerning the **molecular structure of nucleic acids** and its significance for **information transfer in living material**"



-> the molecular gene (1960s)

-> Crick: ‘sequence hypothesis’




3′
Phe (F)
Leu (L)
Ser (S)
Tyr (Y)
Cys (C)
Trp (W)
Leu (L)
Pro (P)
His (H)
Gln (Q)
Arg (R)
Ile (I)
Met (M)
Thr (T)
Asn (N)
Lys (K)
Ser (S)
Arg (R)
Val (V)
Ala (A)
Asp (D)
Glu (E)
Gly (G)
5′
3′
3′
3′
U C A G
C A G U
A G U C
G U C A
Start
Stop

www.dnalc.org

Molecular Genetics

-often dated to **Watson and Crick's (1953)** discovery of structure of DNA

-1900-1953: investigation into cellular and molecular basis of heredity

-molecular basis not resolved till mid-century (nucleic acid vs protein)

-but known since 1903 that in meiosis, each gamete receives **one of a pair of homologous chromosomes**

-relevance?
-> explains Mendel's first law

Molecular Genetics

-molecular genetics **sheds much light** on classical genetics:

i) **structure** of DNA explains how it can be **reliably replicated**
-> DNA composed of two ***complementary*** strands
-each strand **uniquely specifies** the other
-explains gene replication

ii) DNA (indirectly) makes the proteins that make up structural and functional elements of cells
-DNA is ***transcribed*** into mRNA, which itself is ***translated*** into protein
-explains (in theory) how **different genes** can have **different phenotypic effects**

iii) ***appears*** to tell us what a gene is ***in physical terms***
-> a gene is **a length of DNA** that **makes a single protein**
-> genes are **no longer hypothetical** entities, but **physical** entities

-no-one doubts that discoveries in molecular biology **shed lots of light** on classical genetics
-***but what exactly is the relation?***

-not entirely **simple**
-> classical and molecular geneticists use 'gene' in somewhat **different** ways

-**not** obvious that everything that is a **classical** gene is a **molecular** gene
-nor vice-versa

structure

(protein)

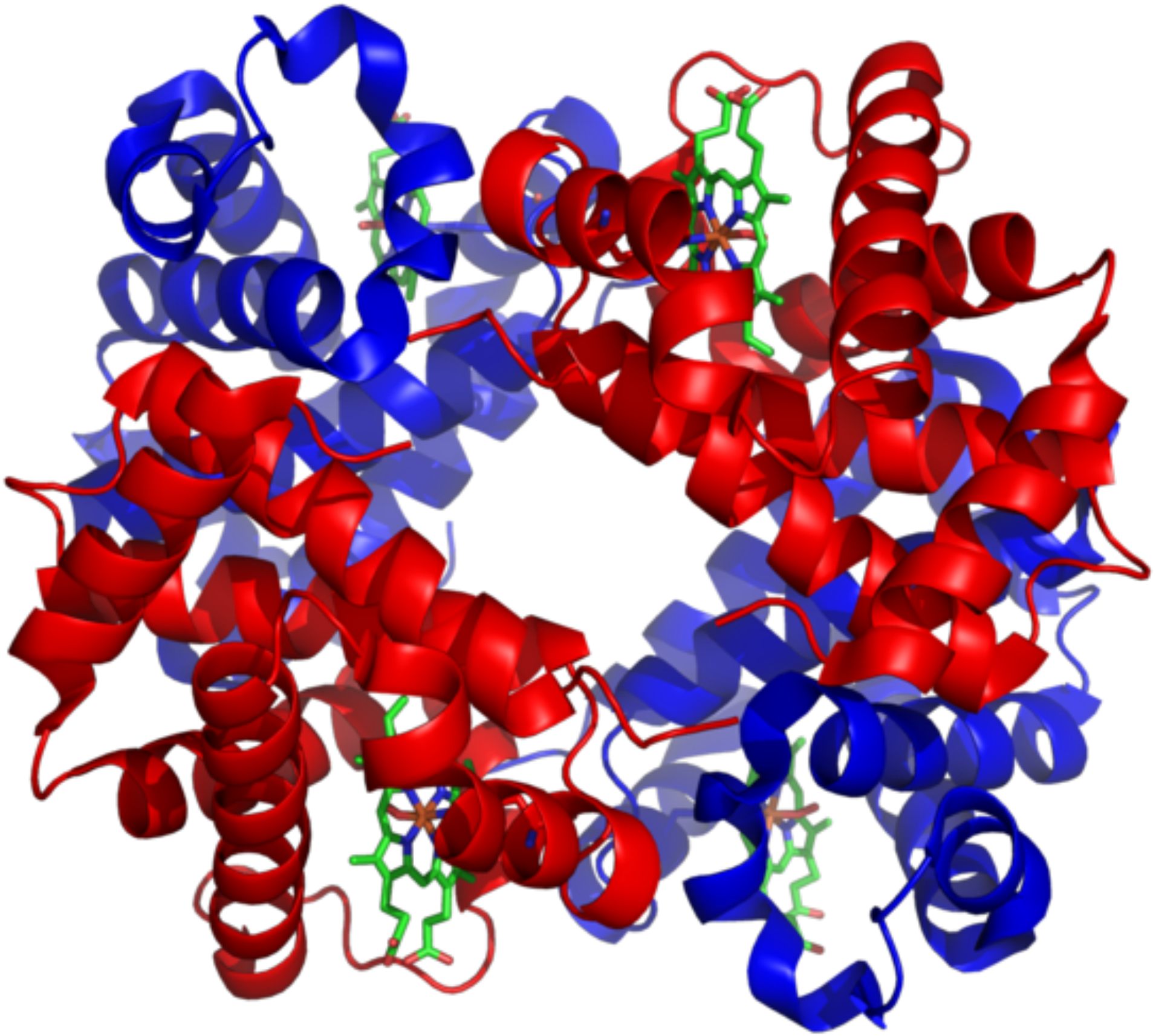
encodes

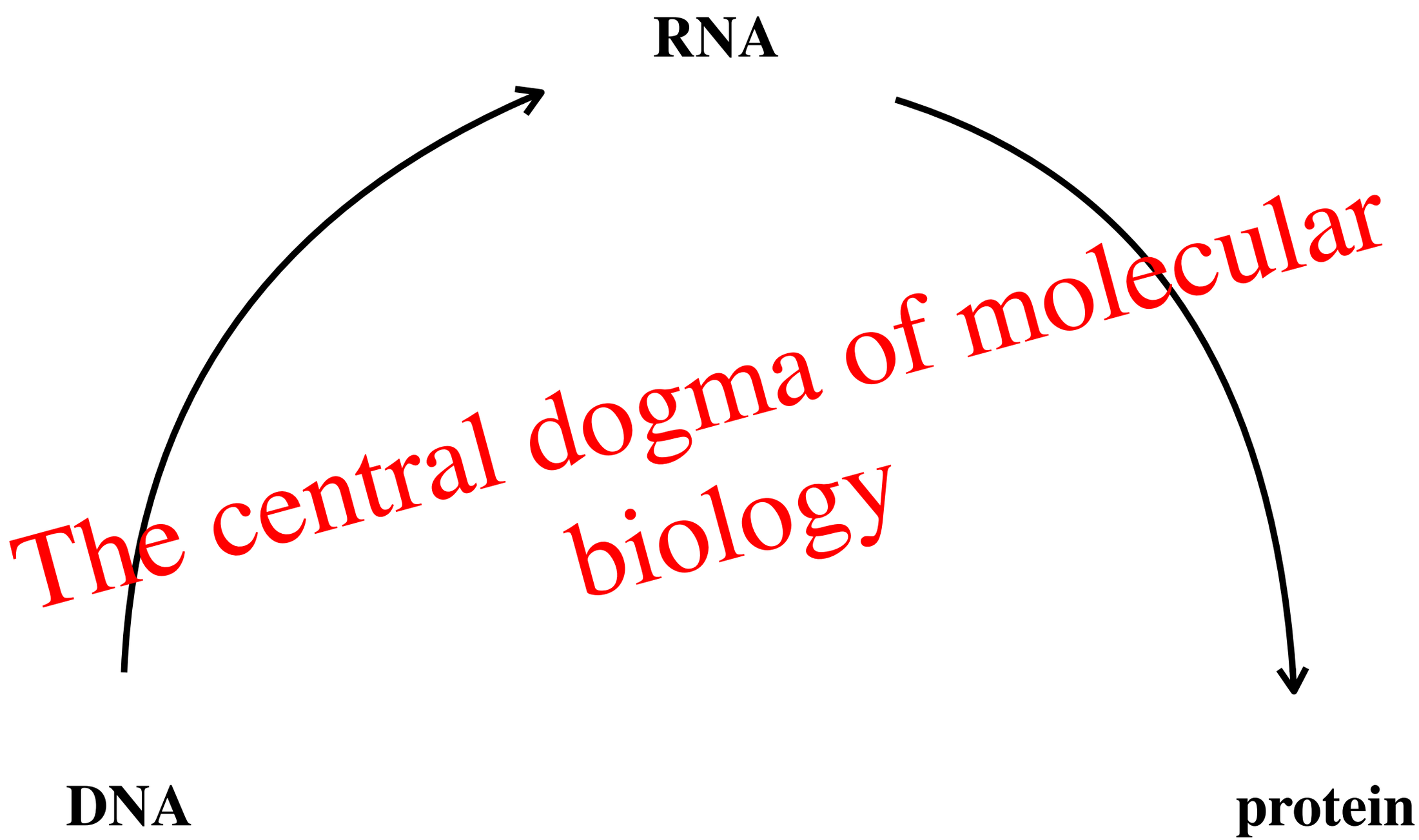
enables

information

(gene)

function




RNA
The central dogma of molecular biology
DNA
protein

a message

encodes

to build

a gene

a protein

16 February 2001

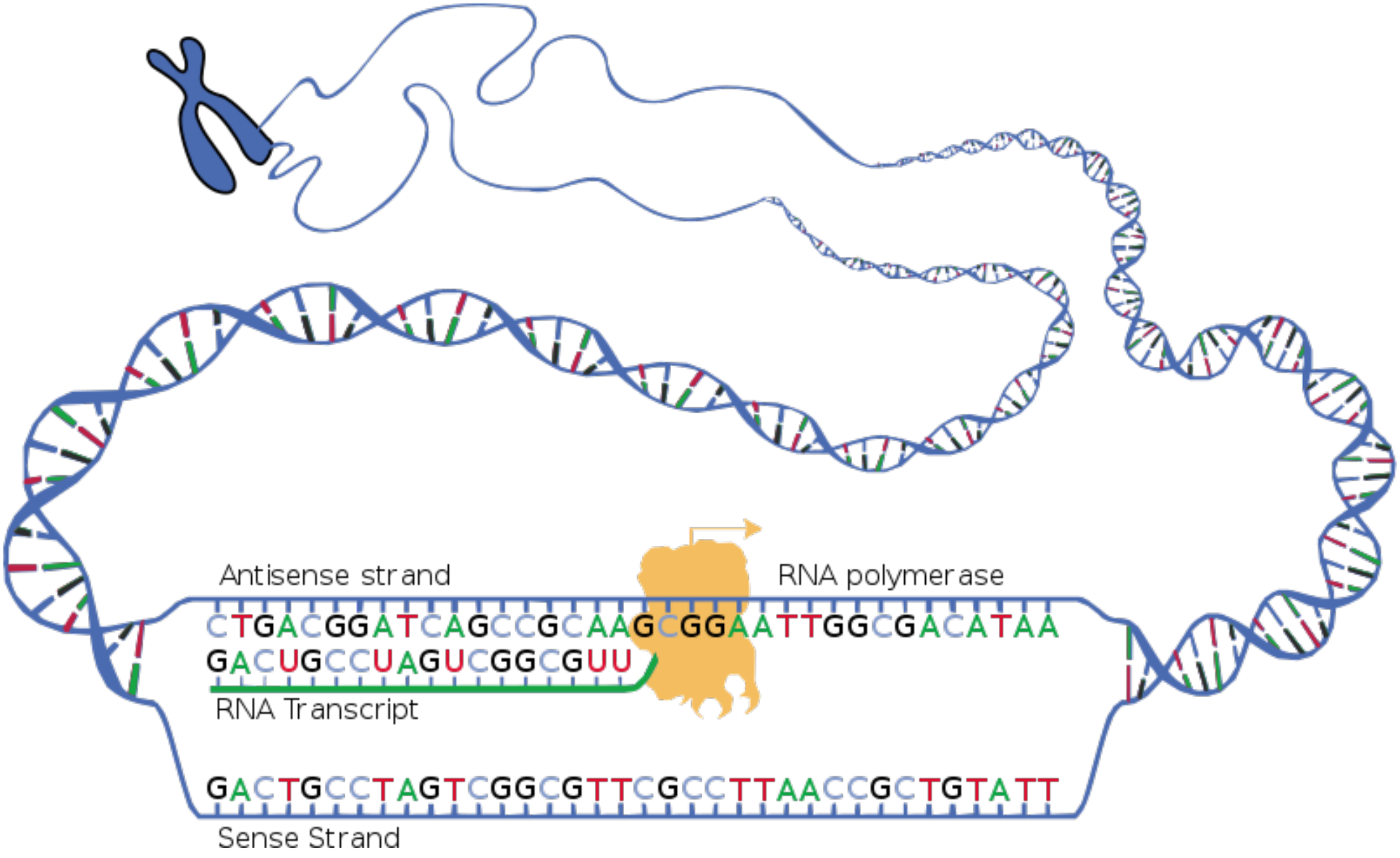
Science

Vol. 291 No. 5507
Pages 1145–1434 $9

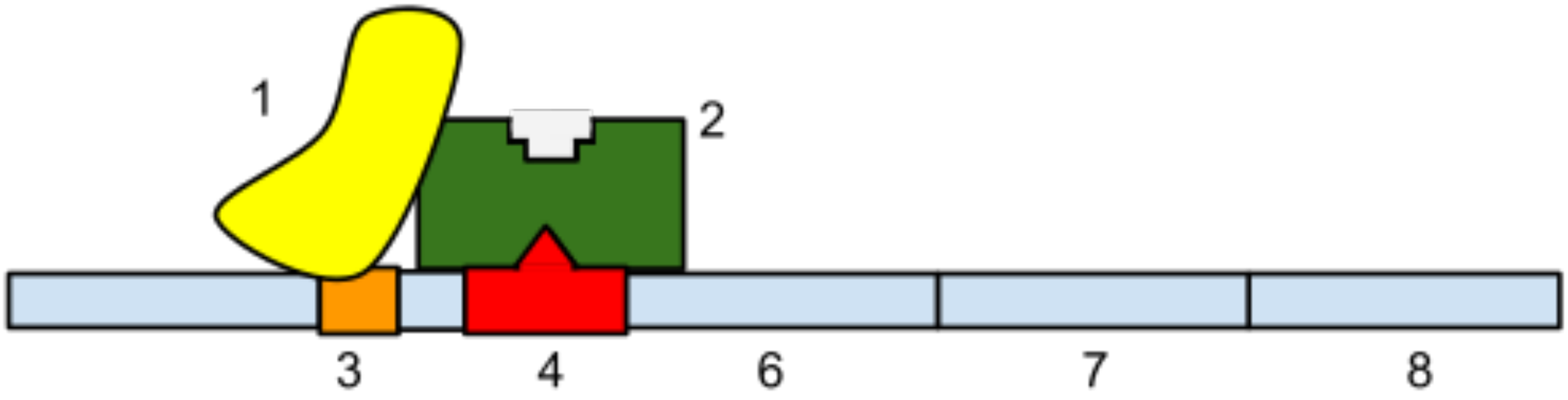
THE HUMAN GENOME

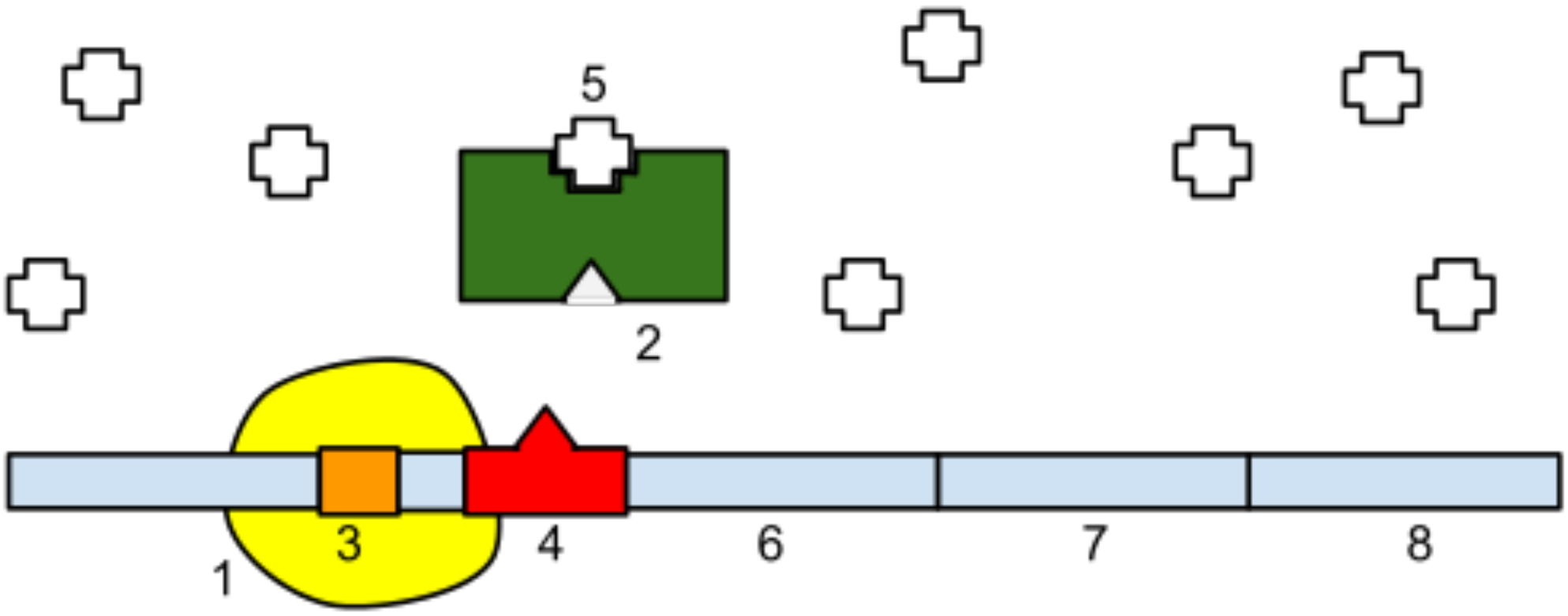
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

GATCCCTGCC TGGGCTTTGC CTCTGCAGCC CCCCGCCCCA CAGGTTCACA CCTCGGGTCT
TCTCCACCGC TGCCACACGC CAGAGCCTGT AGCGGGGCCT CAGAGTCTGG GAGGTGGGAC
TCCTGCACCT CAGCCATCAT CAGACCCATG GGGCCACCCA GGGAACCTTG GCAGGGACCA
TTACCAGTGA CCTGCCGAGG CCCCGGACTC TGCCAGCCAG CTGTGCCGGC CACCCTGCCC
CGGACAGTGC CGGTTCATGT GGGAACTAGG GGACGATGTG GTTCTTCGCA TCTGATGATG
AAGGCCCTGG GCCACTTGGC ACGGGCGGGC GCTCCCGAGA TGGATATGAG GAGCCCCCTC
TGCCCAACTC CCAGAAAGGC CGAGGCTCTG CAGCGGGAGG AAGTCCTGCG ATGTCCTGGG
GGGCAGCAGC GCAGGGCACA GGGACAGCCC CCCTCCACAG CTCTTCCTGG CCAGCCCTCC
CCACTATCTG CCAGGAGGTT GCTTCTTCCA GGAGGCTTTT CCCGACCAGC CCAGGGGTCC
AGGGTCTGGG GCTCCCAGCT GCTGTGAGTG CTGCACATTC TCTTGAGGAC AGCCCCCTCC
CTCCCCCACC CACTTCTGGT GCCCACTGTG GCCACAGCAA GCACTGGGGC CTGCACTCAG
GGACCTCGGG GCCTCCTGGG GAGCTGCTGA CCCTAGGCAG AGAGATTGCA CATCCCTAAG
AGTCTACAGA CACCCCAGTG TTTGCCAGTG TTTGCCCGTG TTCACCAGTG TTTGCCAGTG
TTTGCCAGTA TTTGCTCGCC AGTGTTCGCC ACTTGTCCCT CTGGCTGCAA GAGTGACTGG
GTTTGGGCGG GAAGTTGCAG GTCCCTCCAG GACAGTTGGC CGATGACGTG GAGACAGACC
CACCCCCCAA TCCTGGCTCC CTGCAGGACG CGGGGCCCCC CGAGATCCTG GCGGTGCTCA
GCACGACGGG CACCTCCGTG TTCACCAGTC CAATGGGCAC GGAGCGTGGC TTTATTTGCA
TGTCTGGATT CCTAACGACT TCAGCCTCTG CACCTCCTGG GTTTTCCCTG CTGCAAATTG
CCATTTGGCG TCGTCCCCAA TTTCCGGCCA AGGCCGCGTC GTCGTGCTGC TGTGTAATTT
GATGTGTGGA GTTCTAGATA CCAAGTGTCT GTCGGTTTTA GACATCGCAA ACGTCCTTCC
CAGTGTGGCC CGTCCATTCG CTTCTGTGCA GCAAAATCTT TAATTATTTG ATGGCATCAA
AATGTGTGTC CAGTTTTACC TTCTAGTTTA TACTTTCGAA CATTTGTTTG AGAAATCTTT
CTCCCACCTG TGGCTGATAG TGACGTCTTC TAACTTCCCA TTTACTATGT TACATTCAGA
CCCATCATCT TCAGGAAGAC GCTTGTGTGC GAGACGGGTA TGAGGCCCCC ACACCCCGCC


Antisense strand
RNA polymerase
CTGACGGATCAGCCGCAAGCGGAATTGGCGACATAA
GACUGCCUAGUCGGCGUU
RNA Transcript
GACTGCCTAGTCGGCGTTCGCCTTAACCGCTGTATT
Sense Strand

-> operon





a gene

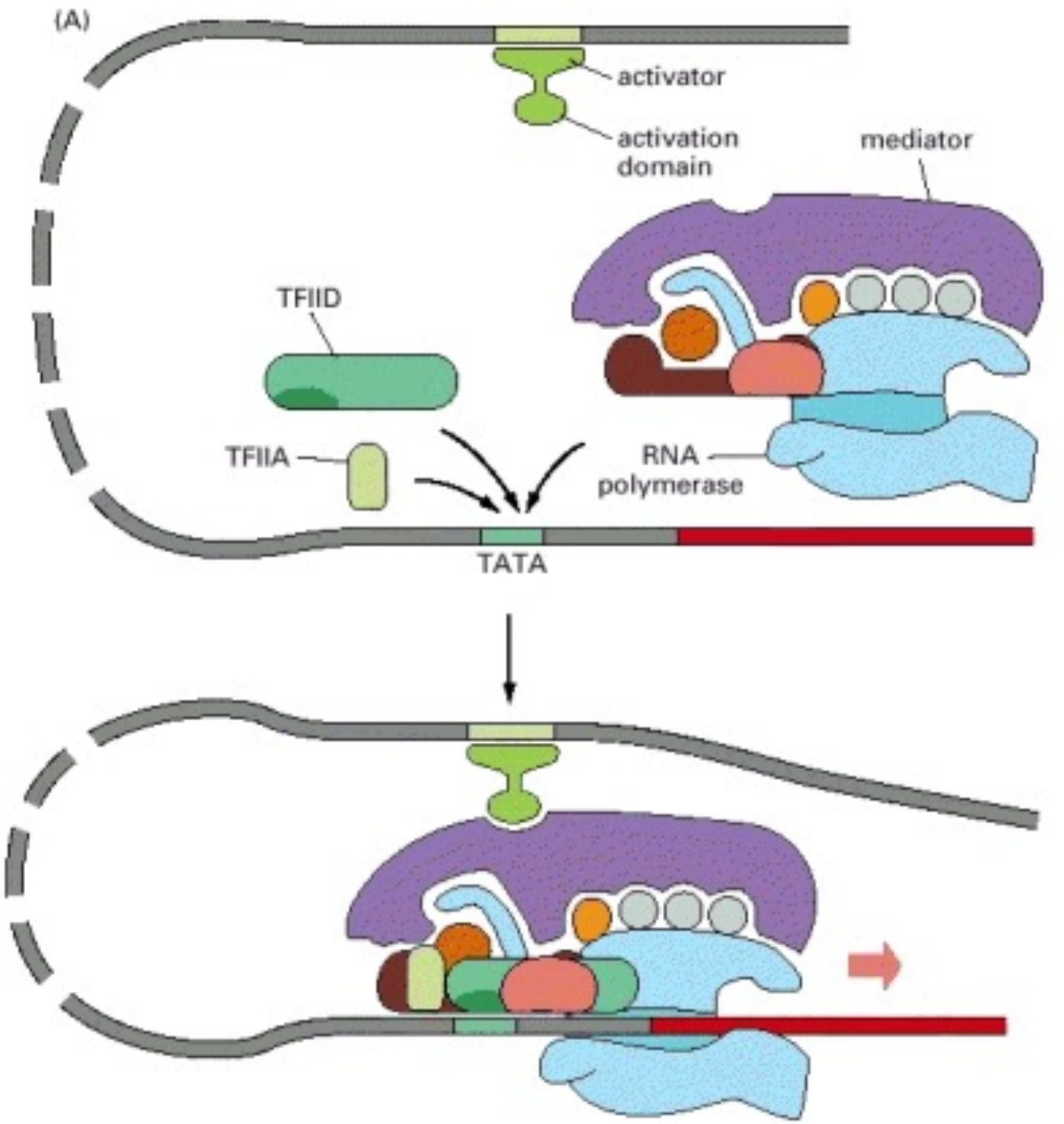
that regulates

encodes

a protein

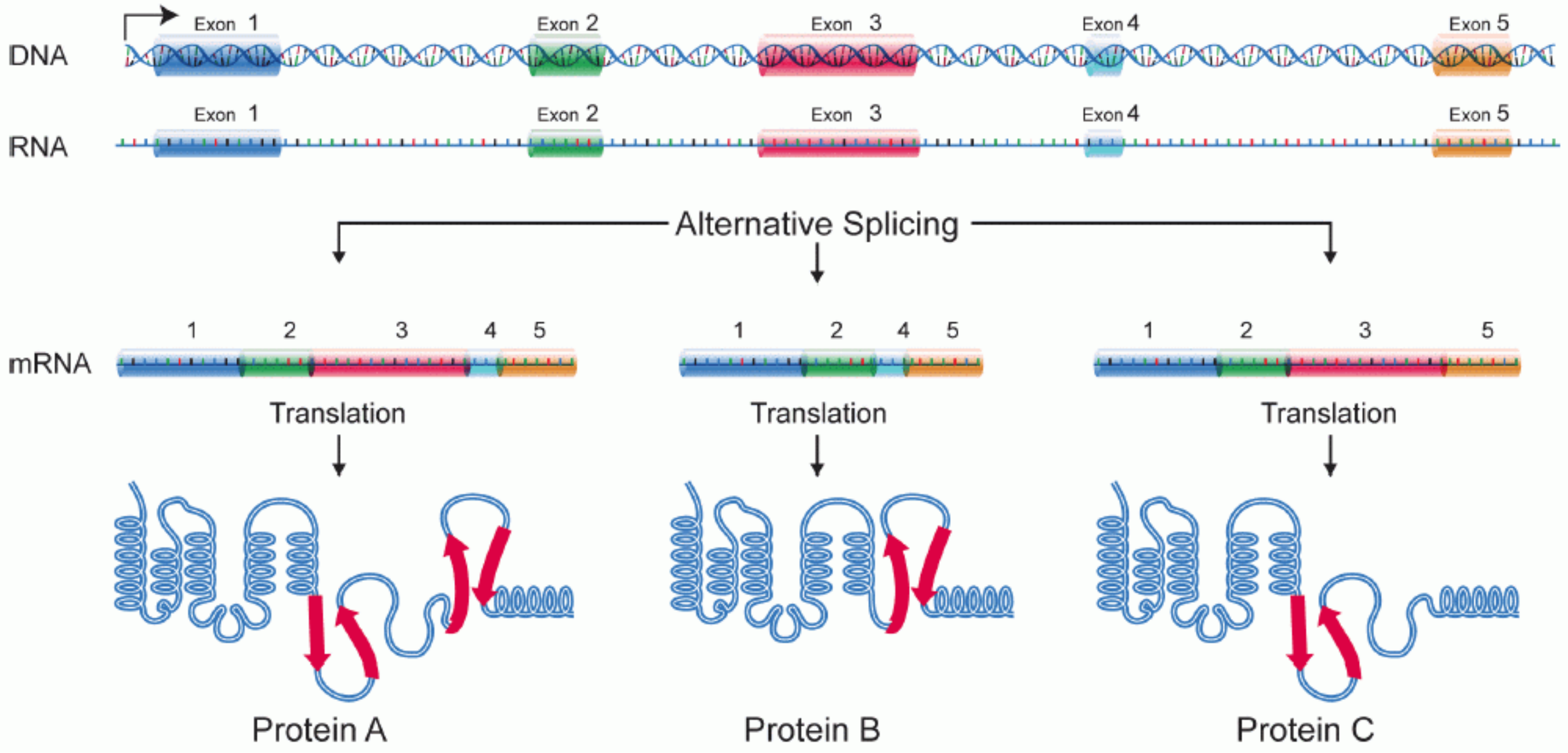
a message

to build


(A)
activator
activation domain
mediator
TFIID
TFIIA
RNA polymerase
TATA

www.dnalc.org

-> alternative splicing



Ovalbumin processing

DNA Genome

7,700 base pairs

transcription

Primary RNA Transcript

5' 3'

1. addition of 5'-cap
2. addition of poly [A] tail

5' cap

poly [A] tail

removal of 5 introns (splicing)

removal of 2 introns (splicing)

mature mRNA

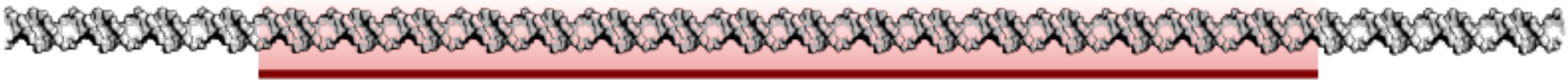
1,872 bases

export to cytoplasm

INTRON

EXON

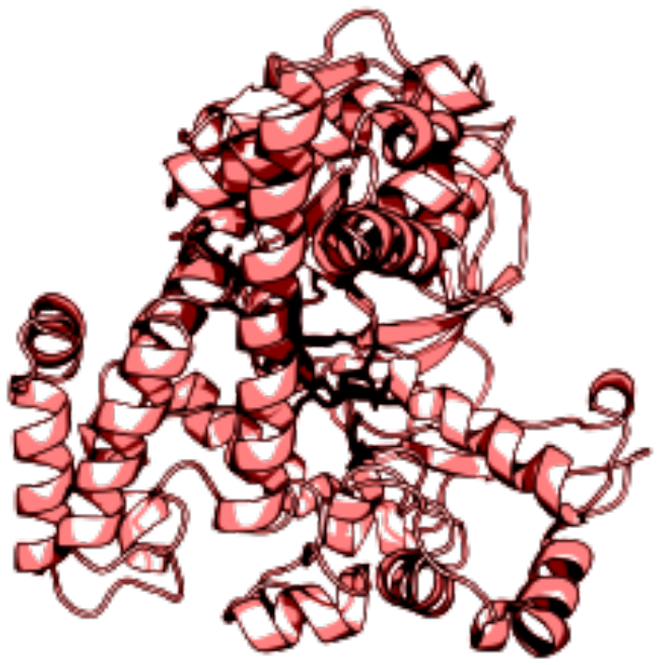
DNA



Gene

mRNA

Protein



Function

ncRNA

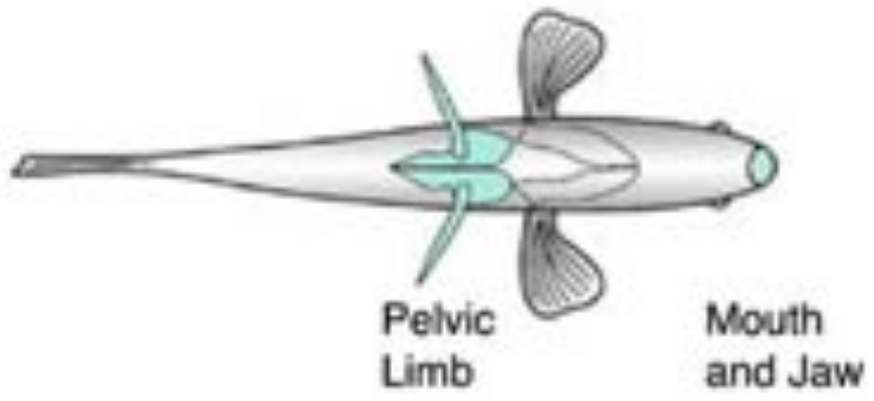
Function

Marine

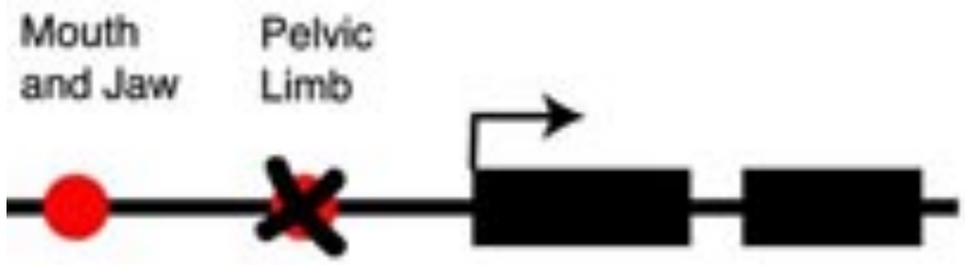
Freshwater

Marine Stickleback Pitx1 Gene

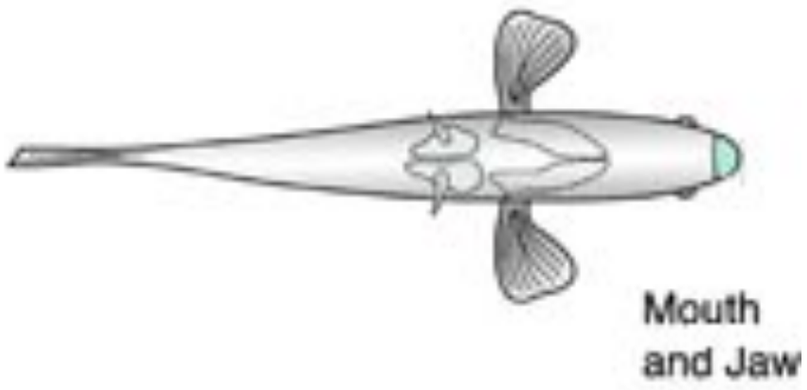
In Marine Sticklebacks Pitx1 Gene is turned on

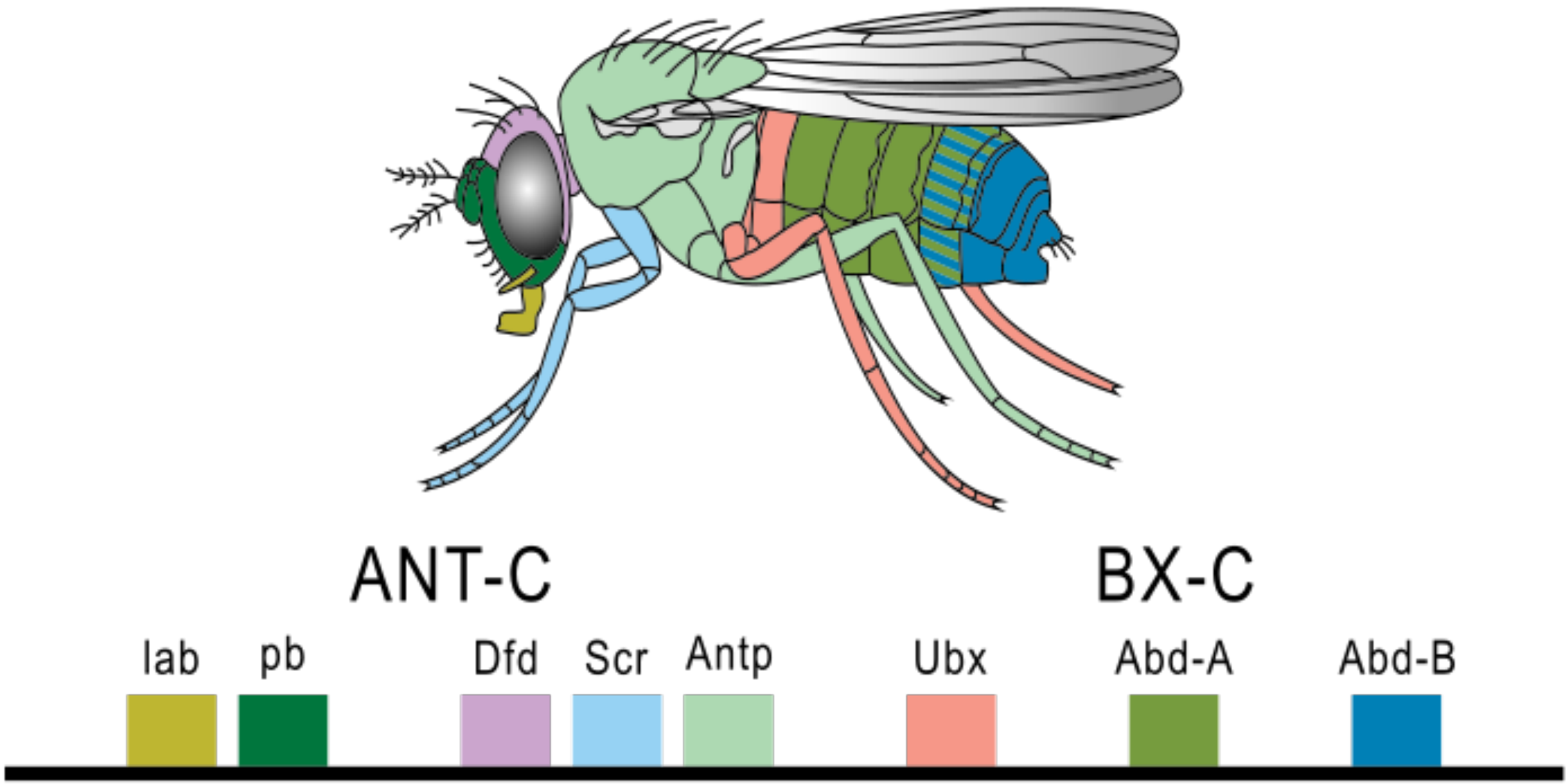


Freshwater Stickleback Pitx1 Gene

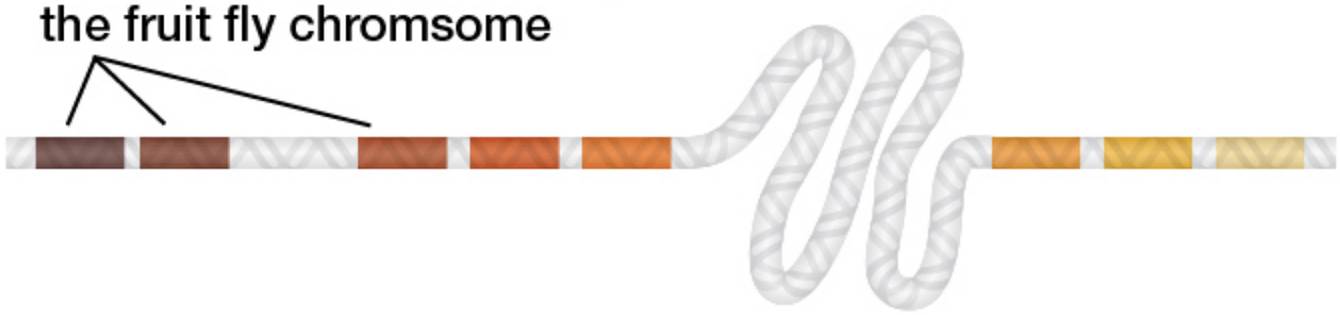


In Freshwater Sticklebacks Pitx1 Gene is turned on

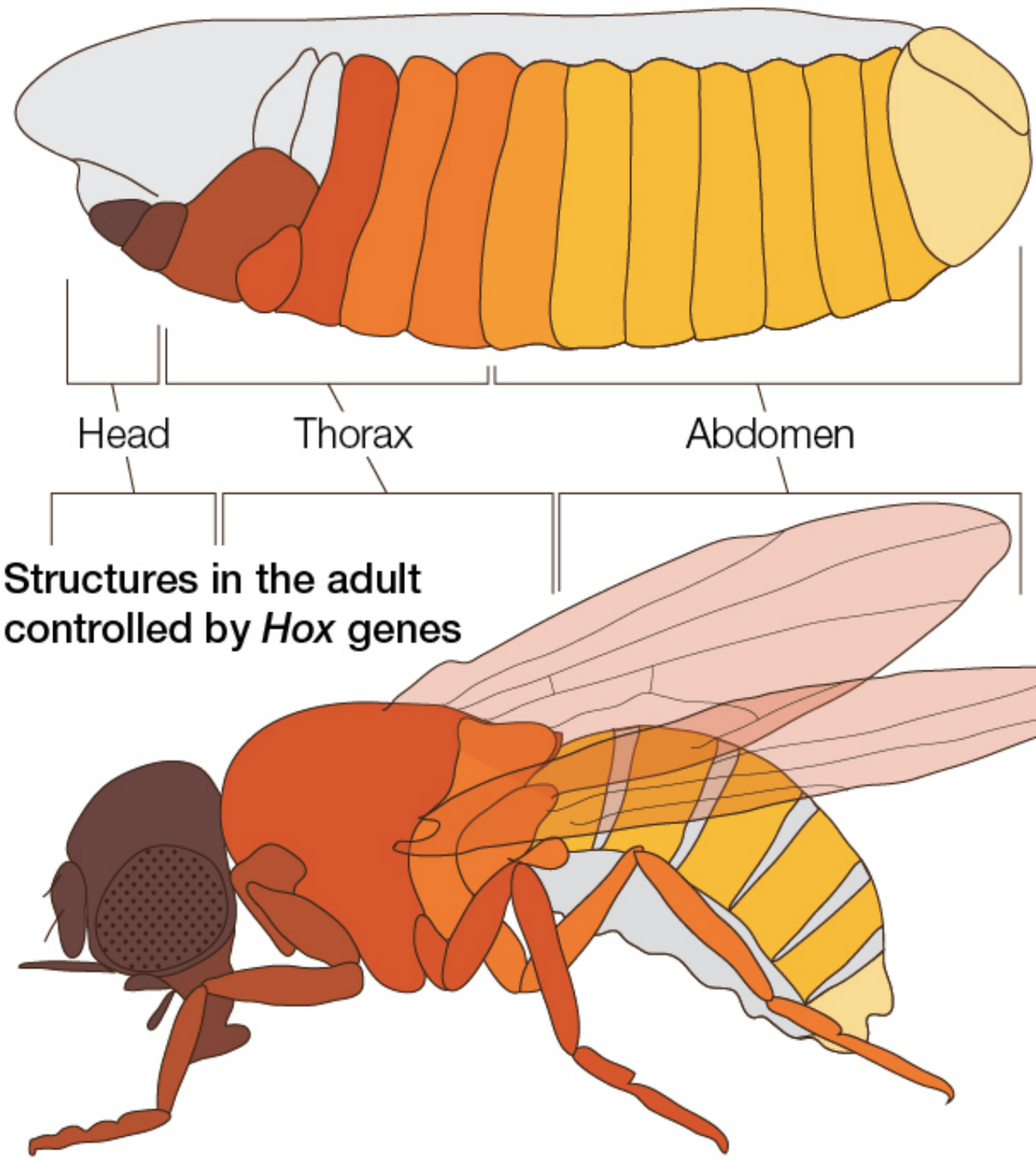


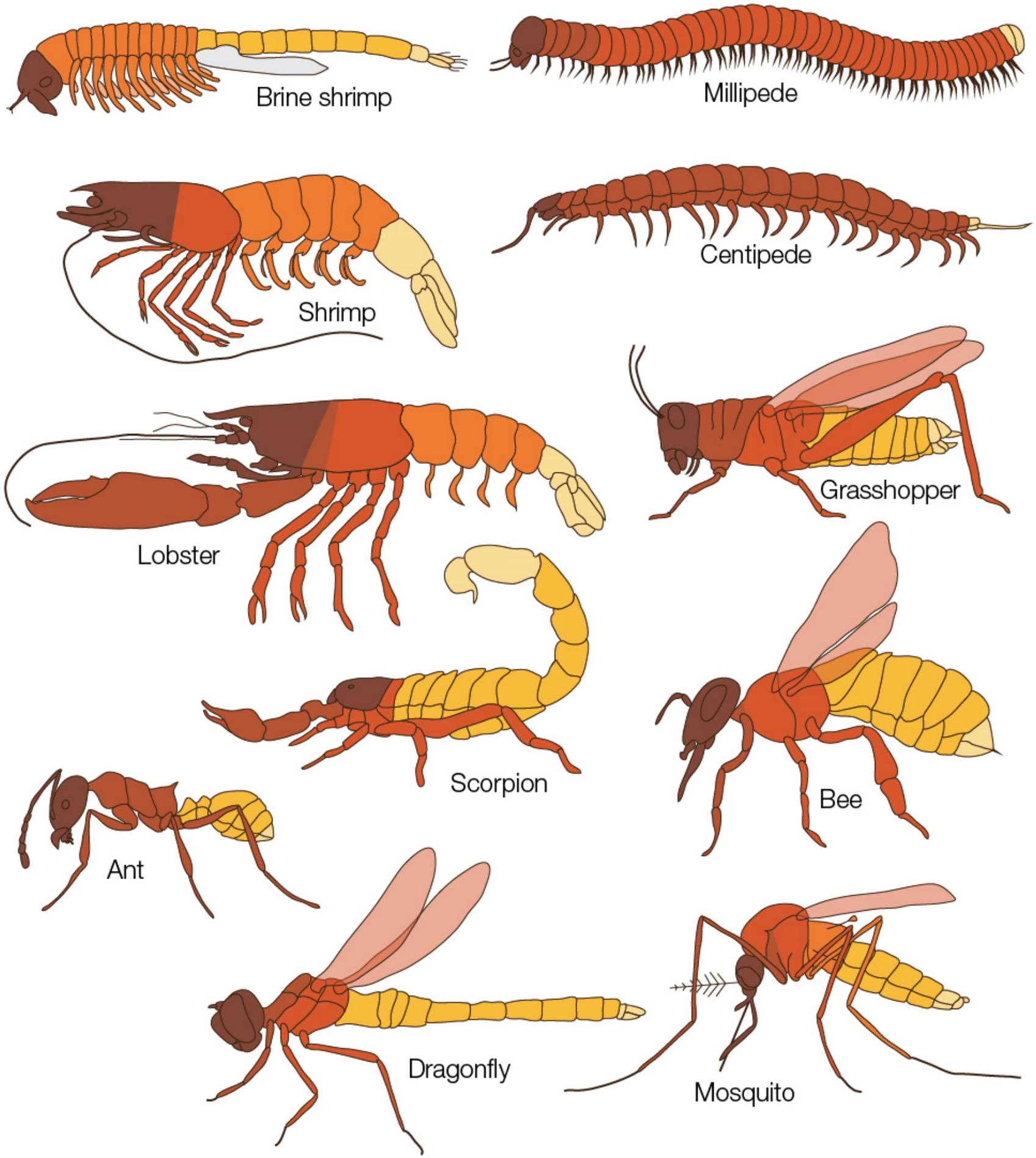

ANT-C
BX-C
lab
pb
Dfd
Scr
Antp
Ubx
Abd-A
Abd-B

Hox genes arranged along the fruit fly chromsome



Zones of *Hox* gene activity in the embryo




Brine shrimp
Millipede
Shrimp
Centipede
Grasshopper
Lobster
Scorpion
Bee
Ant
Dragonfly
Mosquito

lab pb Hox3 Dfd Scr ftz Antp Ubx abd-A Abd-B

Euperipatoides

Archegozetes

Phalangium

Cupiennius

Centruroides

Glomeris

Lithobius

Strigamia

Porcellio

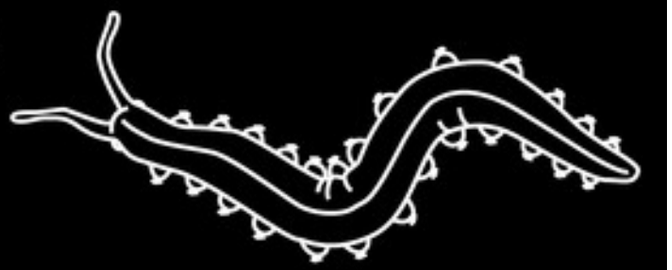
Artemia

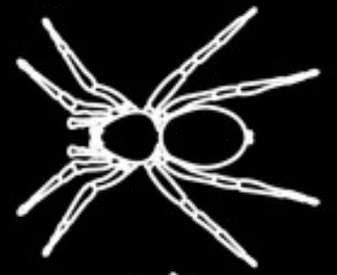
Thermobia

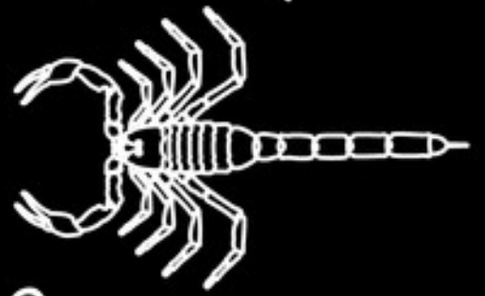
Oncopeltus

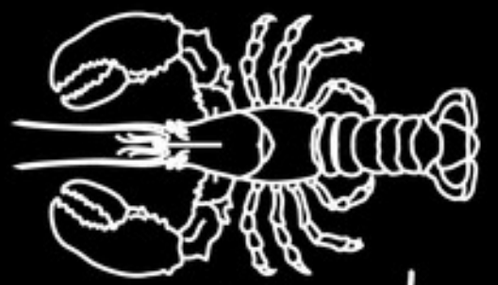
Tribolium

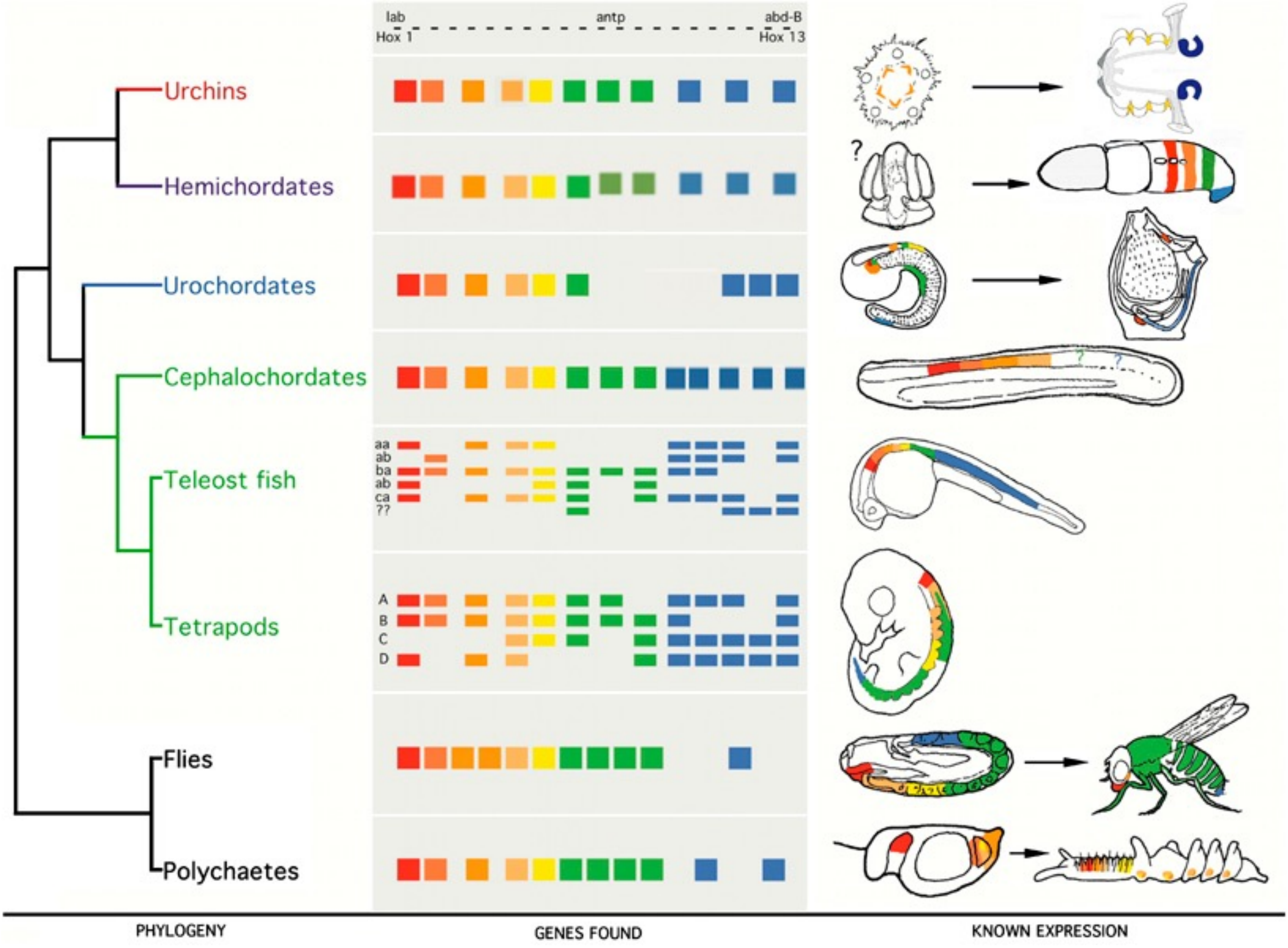
Drosophila

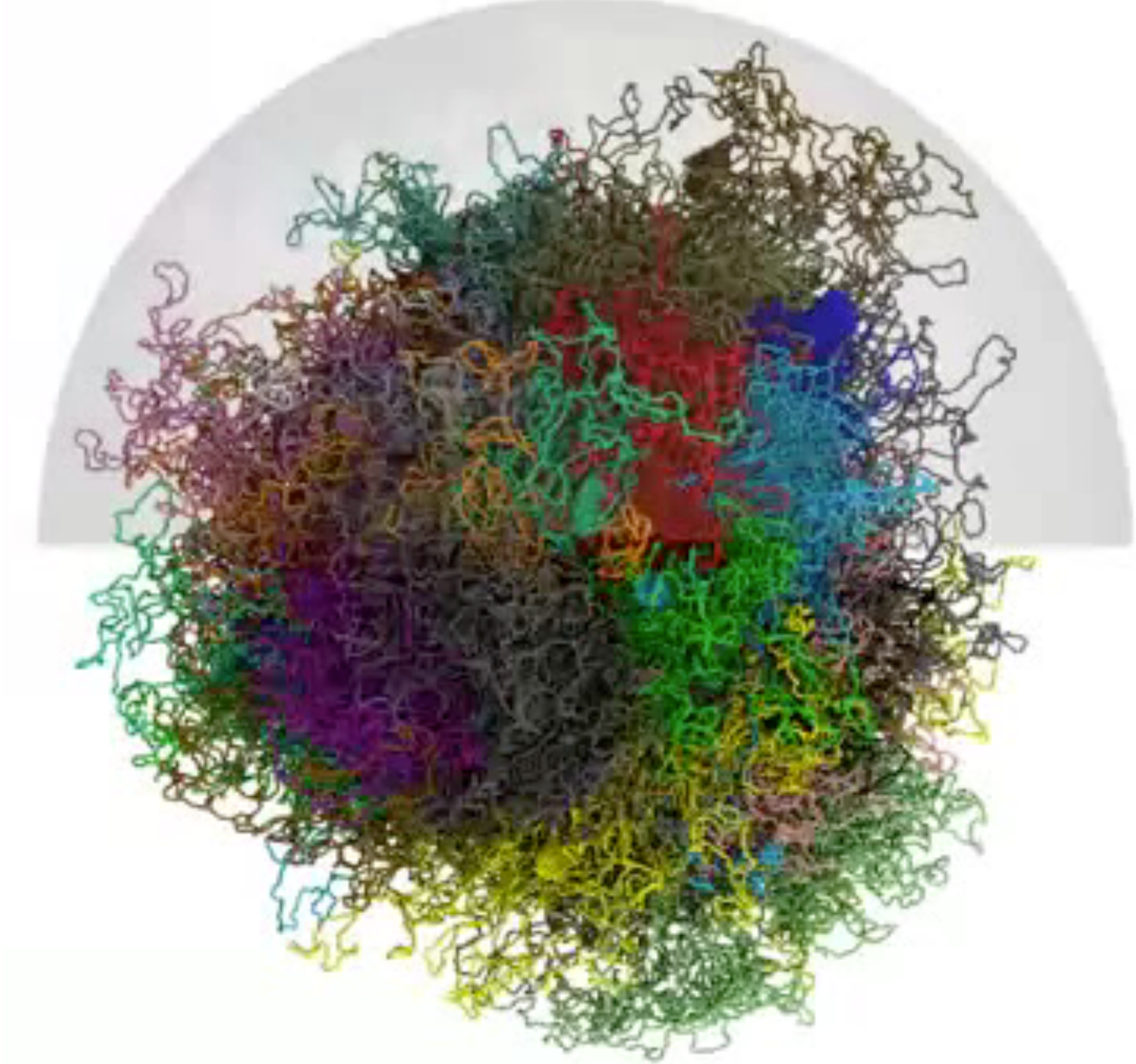









lab
Hox 1
antp
abd-B
Hox 13
Urchins
Hemichordates
Urochordates
Cephalochordates
Teleost fish
Tetrapods
Flies
Polychaetes
aa
ab
ba
ab
ca
??
A
B
C
D
?
PHYLOGENY
GENES FOUND
KNOWN EXPRESSION



For further study:

-Godfrey-Smith, *Philosophy of Biology*, ch. 6

-Griffiths & Neumann-Held (1999) 'The many faces of the gene'

-Sterelny & Griffiths, ch. 6 + 7

-Griffiths & Stotz (2013), *Genetics and philosophy*

-https://plato.stanford.edu/entries/gene/

-https://plato.stanford.edu/entries/molecular-genetics/

-https://plato.stanford.edu/entries/molecular-biology/

- polygeny
- pleiotropy
- introns exons
- Griffiths Stotz
- beanbag genetics