TRENDS IN GENETIC ANALYSIS

By

G. PONTECORVO

TRENDS IN GENETIC ANALYSIS

BY G. PONTECORVO, F.R.S.

Professor of Genetics, University of Glasgow

1958 NEW YORK

COLUMBIA UNIVERSITY PRESS

PUBLISHED IN GREAT BRITAIN, CANADA, INDIA, AND PAKISTAN
BY THE OXFORD UNIVERSITY PRESS
LONDON, TORONTO, BOMBAY, AND KARACHI

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 58-8805
MANUFACTURED IN THE UNITED STATES OF AMERICA

COLUMBIA BIOLOGICAL SERIES

Edited at Columbia University

CONTENTS

TABLES

PLATE

INTRODUCTION

THIS book is based on the Jesup Lectures delivered in April, 1956, in the Department of Zoology of Columbia University in the City of New York. A portion of the text is unchanged and the colloquial style generally retained.

The title—for short, "Trends in Genetic Analysis"—is an obvious overstatement: there was no intention, of course, of dealing in six lectures with all the directions along which genetics is advancing. The treatment was confined only to those fields with which the author has firsthand acquaintance. Cytoplasmic inheritance, biometrical genetics, and even the detailed study of mutation were completely omitted. Yet no one, least of all the author, will maintain that there are no trends there. These are all frontier fields expanding very vigorously. The scope of the lectures was a reappraisal, on the basis of present knowledge, of the theory of the gene.

A by-product of recent research is the realization that sexual reproduction—i.e., a regular alternation of karyogamy and meiosis as shown in higher organisms—is by no means the only process for the pooling and reassorting of genetic information from different lines of descent. Though known so far only in microorganisms, novel processes of genetic recombination make it clear that some modernized version of the theory of the gene is applicable in organisms or situations in which sexual reproduction (the basis of the original theory) does not occur. The two closing chapters of this book deal precisely with these novel processes.

To put the content of this book in its correct historical perspective, we have to remember a few landmarks. Bateson and Punnett discovered linkage in 1906. In 1912 Morgan and his group at Columbia University correlated the recombination of linked genes with the occurrence of reciprocal exchanges between homologous chromosomes at meiosis, which had been demonstrated by Janssens in 1909. The linear arrangement of genes and the first linkage map by Sturtevant soon followed in 1913. The next landmark is the series of papers by Muller on the mechanism of crossing over, published in 1916. With this work and with that by Bridges on "non-disjunction," all the elements for the theory of the gene had been gathered, though this wording was used by Morgan only in 1917. The theory was expounded by the Columbia University team in 1915 as "The Mechanism of Mendelian heredity." The term "gene," coined by Johannsen in 1911, has already become established.

It is no exaggeration to say that before about 1940 what was known on the nature and the mode of transmission of genetic specificity—*i.e.*, what was known about chromosomal heredity—was but a series of developments of the theory of the gene.

First came the studies of mutation started by Muller about 1920 that led in 1927 to the discovery by Muller and by Stadler of the mutagenic effects of radiations. It took about fifteen years before chemical mutagenesis was demonstrated by Auerbach and Robson and a vast new field opened up.

Second, in the 1920s Wilson, Belling, and later Darlington unified Mendelian inheritance and chromosome cytology. About the same period witnessed another major success: the unification of Mendelism and Darwinism, which the theoretical work of Fisher, Sewall Wright, Haldane, and Tchetverikov made possible. It also witnessed a substantial beginning: the analysis of the effects of genes in terms of biochemical and developmental processes, in which Haldane, Goldschmidt, and later Ephrussi played leading roles.

Third, in the years immediately preceding World War II, something quite new happened: the introduction of ideas (not techniques) from the realm of physics into the realm of genetics, particu-

larly applied to the problems of the size, mutability, and self-replication of genes. The names of Jordan, Frank-Kamenetski, Friedrich-Freska, Zimmer, and Delbrück, with Muller and Timofeef-Ressovsky as their biological interpreters, are linked to this development. Though this first application of physical ideas to a particular set of problems did not work out too well, the whole outlook in theoretical genetics has since been perfused with a physical flavour.

The debt of genetics to physics, and to physical chemistry, for ideas began to be substantial then, and it has been growing steadily all the time. Techniques from physics and physical chemistry, on the other hand, have contributed very little to genetics. This is in sharp contrast to the relations of genetics with chemistry and biochemistry, which have contributed innumerable techniques and facts, but few, if any, ideas.

The historical landscape sketched here is what appeared to the writer in 1939 when, at an age older than is usual for a neophyte, he started learning the elements of genetics. No doubt the picture is blurred and incomplete: some landmarks are forgotten and others are given too much prominence. But the main point, which can be hardly controversial, is the one made above, that up to 1940, or thereabouts, genetics was essentially a development of the "Theory of the Gene." Its impact has been profound, and has remoulded the present ways of looking at living things and investigating the living world. Its impact on politics of the extreme "left" or the extreme "right" has also had far-reaching results, but few would say for the better.

Since about 1940 there has been a gradual change in the outlook in genetics. One reason for this change is the realisation that the theory of the gene, though still indispensable in everyday genetics, is no longer of heuristic value at levels of further refinement, especially when it comes to the enquiry of what the genetic materials are and how they work. The analogy here—and it is almost a platitude—is with classical physics versus quantum mechanics. Unfortunately the quantum mechanics of genetics is not yet with us.

Another reason is technical: the developments brought about by

the use of microorganisms in genetics. This has enormously increased the resolving power of genetic analysis and has stimulated very fruitful general ideas from the study of relatively simpler genetic systems.

A third reason is the closer relationship between biochemistry and genetics. For this we are indebted mostly to Beadle and Tatum, who, quite independently, rediscovered Garrod's idea of "inborn errors of metabolism" and applied it to biosyntheses in microorganisms. By doing so they provided a technique of immense and versatile power: suffice it to say that most of the development of microbial genetics is based on this technique. Unfortunately, by and large, this technique has not been put to the best possible use in one of the directions for which it has exceptional value, *i.e.*, the study of the primary actions of the genetic material and their relations to its fine structure. In this respect, it has been made mainly into a tool for the unexciting description of intermediary metabolism, for which it competes or cooperates with half a dozen other more traditional techniques. Only occasionally has it been used for the study of the biochemical system of which the genetic material is one component. It is still full of promise but has not yet made fundamental achievements. The large amount of information collected since 1940 on the genetic control of biosyntheses has so far only descriptive value.

A fourth reason is the impact of ideas from information theory, especially in relation to molecular structure. It is too soon to assess its results, but I have no doubt about its decisive importance.

Clearly, if we are to free ourselves of the fetters of purely formal genetics, of genetics based on abstractions—though valid abstractions—of genetics as merely the mechanics of hereditary transmission, there is no doubt that we have to give physical, chemical, and physiological content to the processes of heredity, variation, and differentiation. We have to express such concepts as gene, allele, mutation, crossing over, dominance, etc., in terms of precise processes taking place in or on structures of the cell.

So far there is a certain reluctance both in approaches and in techniques to giving structure to the chemical processes within the cell: *i.e.*, trying to do something about the fact that spatial arrangement at the megamolecular or even higher levels is an essential part of the game. Coupled reactions on surfaces, relations of molecular structure to biological activity, arrangement of reactants in microvessels, are the things that matter for our purpose, rather than the elegant unravelling of pathways of intermediary metabolism. This is why physicists and physical chemists have contributed decisively to biological thought. This is why attempts like those of Watson and Crick, Pauling, Szent-Gyorgy, and Astbury, have such great appeal.

The fundamental problems in genetics with which we are faced are still the same as those which Muller so clearly stated in his Pilgrim Trust Lecture, "The Gene," at The Royal Society in 1945: among these, the nature of the self-duplicating process of genetic structures, the nature of gene effects, and the nature of the recombination process.

As to the first two, but not the third (see Chapter IV), we can now state the problems more specifically, and there are some testable models. For instance, on the one hand we can now test whether or not the atoms of an existing genetic structure are distributed at random between the two structures resulting from its self-duplication. The brilliant experiments with isotopes by Levinthal (1956) with bacteriophage, and by Plaut and Mazia (1956) and Taylor (1957) with chromosomes, show that they are not.

On the other hand Watson and Crick (1953) and Kacser (1956) have introduced an idea which does not seem to have occurred before (for example, see Muller, 1947a, p. 21), *i.e.*, that the self-duplicating structure itself has a complementary architecture like a positive and a negative photograph face to face.

According to this idea, the duplication process of a structure symbolised as $\overset{*}{A}\overset{*}{B}$ (* indicates parental structure) consists in $\overset{*}{A}$ com-

bining with B's building blocks to give a "daughter" $\overset{*}{A}B$, and $\overset{*}{B}$ combining with A's building blocks to give a "daughter" $A\overset{*}{B}$.

Watson and Crick applied this idea to the molecular (double helix) structure they proposed for DNA: in this case A and B stand for nucleotides able to pair by hydrogen bonds. Kacser applied it to a protein-DNA complementary architecture of supramolecular size in which the genetic specificity is a property of the interface between the two constituents.

Penrose and Penrose (1957 and unpublished), however, in a note which deserves more consideration than it has received, have shown that it is possible to devise even large mechanical models which are self-duplicating. An essential feature of these models is a symmetrical structure say $\overset{*}{A}\overset{*}{A}$, which can duplicate either by accretion on each side followed by splitting ($\overset{*}{A}\overset{*}{A}AA \rightarrow A\overset{*}{A} + \overset{*}{A}A$) or on one side only ($\overset{*}{A}\overset{*}{A}AA \rightarrow \overset{*}{A}\overset{*}{A} + AA$).

Models of these kinds have predictable consequences, and it is not too optimistic to believe that experiments of the types already mentioned will soon be able to discriminate between them.

As to the nature of gene effects, Beadle and Tatum's technique has provided material at will. The most promising line seems to be the analysis of cases in which a heterozygote produces both substances, say proteins, each of which is produced by each of the two homozygotes. Horowitz's work on tyrosinases in Neurospora is one of the best examples of this kind.

The identification of the precise difference between the two substances may give a clue as to what the primary gene action is. For instance, Ingram (1957) has shown that the haemoglobin of individuals homozygous for sickle cell anaemia (a gene-determined abnormality) differs from normal in one out of its 300-odd aminoacid residues. The heterozygotes have both types of haemoglobin.

The next fascinating step is that of finding out what it is, in the gene "code," that determines this difference. Again this aim is not as fantastic as it would have been in 1945. It requires the combination

of work like that of Ingram on a gene-determined protein with work like that of Benzer (1957) or Pritchard (1955) on the fine structure of a gene.

Certain aspects of these problems of genetics are the main subject of this book.

CHAPTER I

GENETIC ANALYSIS AND ITS RESOLVING POWER

"ANALYSIS," in the Oxford dictionary, is defined as "resolution into simple elements." In genetic analysis we must be clear about what we resolve and into what simpler elements.

Classical genetic analysis is based on the results of breeding and by means of them resolves the genome into linkage groups, and each linkage group into loci. By also making use of cytological techniques and combining them with breeding techniques it goes further: it establishes on which chromosome each linkage group has its structural basis and to which small section of the chromosome each locus corresponds.

Mainly as a consequence of the development of microbial genetics, genetic analysis has increased enormously its resolving power in recent years, so much so that it now goes beyond that of physical or chemical techniques applied to biological organisation. I hope to substantiate this contention and make it more precise than was possible in 1952 when it was first put forward.

The essential process on which genetic analysis is based is recombination. Consider the analogy with microscopy, which is based instead on diffraction. The resolving power attained in microscopy depends on the quality of the microscope and on other technical details, but we know that it has a theoretical limit set by the wavelength of the light used. So far, in genetic analysis the resolving power has been limited only by the refinement of techniques. What

the ultimate limit is we do not know, nor can we deduce from theory. Recent advances make it possible to venture a few guesses.

Recombination can be defined as any process which gives origin to cells or individuals associating in new ways two or more hereditary determinants in which their ancestors differed: for instance, cells with determinants *Ab* or *aB* descending from other cells with *AB* or *ab*.

Until less than fifteen years ago, only two processes of recombination were known: sexual reproduction and infection. Now we know that there are more. For instance, transformation by means of desoxyribonucleic acid (Avery, MacLeod, and McCarthy, 1944) and virus-mediated transduction (Zinder and Lederberg, 1952) in bacteria, the parasexual cycle (Pontecorvo, 1954) in fungi, etc.

We recognise recombination by observing in a line of descent certain cells or individuals—recombinants—which show new associations of properties. Recombination of properties, however, is only the detectable secondary effect of reassociation of subcellular structures determining differences in such properties.

In the type of recombination on which classical genetic analysis is based, *i.e.*, recombination in sexual reproduction, these structures are the chromosomes and their linearly arranged elements. The latter are recognised as genes as a consequence of their specific activities in metabolism and development.

In sexual reproduction recombination of chromosomes and their elements takes place at meiosis and it is the result of the independent segregation of nonhomologous chromosome pairs and of crossing over between members of a chromosome pair, respectively.

Crossing over (whatever its precise mechanism, see Chapter IV) can be formally described as the reciprocal exchange of linear bonds at corresponding positions along pairs of homologous chromosomes. These exchanges are microscopically observable in suitable material. In a population of cells going through meiosis, the incidence of exchanges between any two given points in one chromosome pair is highly correlated with the physical distance between these two