ADVANCES IN GENETICS

VOLUME 20

Edited by

E. W. CASPARI

ADVANCES IN GENETICS

VOLUME 20

Edited by

E. W. CASPARI

Department of Biology
University of Rochester
Rochester, New York

1979

ACADEMIC PRESS • NEW YORK

A Subsidiary of Harcourt Brace

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 47-30313

ISBN 0-12-017620-3

PRINTED IN THE UNITED STATES OF AMERICA

79 80 81 82 9 8 7 6 5 4 3 2 1

CONTRIBUTORS TO VOLUME 20

Numbers in parentheses indicate the pages on which the authors' contributions begin.

URSULA K. ABBOTT (217), *Department of Avian Sciences, University of California, Davis, California 95616*

MULKH R. AHUJA (127), *Genetics Institute, Justus-Liebig University, Giessen, West Germany*

JOHN R. MCCARREY (217), *Department of Avian Sciences, University of California, Davis, California 95616*

VEDPAL SINGH MALIK (37), *Research Laboratories, The Upjohn Company, Kalamazoo, Michigan 49001*

ELIZABETH S. RUSSELL (357), *The Jackson Laboratory, Bar Harbor, Maine 04609*

CHRISTOPH SCHOLTISSEK (1), *Institut für Virologie, Justus-Liebig-Universität, Giessen, Frankfurterstr. 107, 6300 Giessen, West Germany*

GEORGE D. SNELL (291), *The Jackson Laboratory, Bar Harbor, Maine 04609*

INDRA K. VASIL (127), *Department of Botany, University of Florida, Gainesville, Florida 32611*

VIMLA VASIL (127), *Department of Botany, University of Florida, Gainesville, Florida 32611*

ERRATUM

Advances in Genetics, Volume 19

David D. Perkins and Edward G. Barry, The Cytogenetics of *Neurospora*

Page 209

Within Figure 18C the label T(IR; IIR)4637 × Normal should be changed to read:

T(IR; IIR)4637 × T(IR; IIR)STL76

CONTENTS

Mechanisms of Genetic Sex Determination, Gonadal Sex Differentiation, and Germ-Cell Development in Animals

JOHN R. MCCARREY AND URSULA K. ABBOTT

Recent Advances in Histocompatibility Immunogenetics

GEORGE D. SNELL

Hereditary Anemias of the Mouse: A Review for Geneticists

ELIZABETH S. RUSSELL

INFLUENZA VIRUS GENETICS

Christoph Scholtissek

Institut für Virologie, Justus-Liebig-Universität, Giessen, West Germany

I. Introduction

Influenza is still an unresolved problem mainly because the agent causing the disease exhibits a high antigenic variability, which is

ADVANCES IN GENETICS, Vol. 20

ISBN 0-12-017620-8

unique for this virus family. With most other viruses, people become infected only once during their life and thereafter are immune against a second infection. For the latter viruses, only a very limited number of serotypes exist that are antigenically stable. Therefore in recent years the molecular genetic basis for the antigenic variation of influenza viruses has been studied extensively. Several review articles have appeared on this matter (Sugiura, 1975; Palese, 1977; Scholtissek, 1978; Webster and Bean, 1978). Because of the unusual structure and behavior of the influenza genome, these viruses became more and more attractive for geneticists. Since the field is in a stage of rapid development, it seems worthwhile again to sum up the most recent findings about the genetics of influenza viruses. Thus, this chapter will deal only briefly with what is known about the structure of the virion and the various gene products. Emphasis will be laid on temperature-sensitive mutants and on the structure of the RNA segments in terms of base sequence, variable and constant genes and regions within genes, which might explain the antigenic variability of the various influenza strains. Some new possibilities for practical applications of influenza genetics will be mentioned at the end of this article.

II. Nomenclature and Structure of Influenza Viruses

A. Nomenclature

Influenza viruses are divided into types A, B, and C according to serological differences of their nucleocapsid proteins, which can be regarded as group-specific antigens.

This review will concentrate mainly on influenza A viruses, which are subdivided according to the serological properties of their surface glycoproteins hemagglutinin (HA) and neuraminidase (NA). Thus the human influenza A virus strains are numbered through according to the various pandemic prototype strains with H0, H1, H2, and H3 (formerly A0, A1, etc.). "H" stands for hemagglutinin, and the numbered designation means that the various strains do not cross-react serologically in the hemagglutinin-inhibition test. Animal influenza strains are marked according to the species from which they were isolated, e.g., Hsw for swine, Heq for equine, Hav for avian. The strains are further classified by the serological properties of their neuraminidases (N1, N2, etc.). According to the WHO report (Chanock *et al.*, 1972), also the year and place, and number of isolation should be indicated. Thus the correct nomenclature of the prototype isolate of the

pandemic in 1933/34 is A/PR/8/34 (H0N1), for which the trivial name PR8 will be used.

B. Structure and Number of Virus-Specific Proteins

The two influenza virus glycoproteins located at the surface of the virus particles are the hemagglutinin (HA) and neuraminidase (NA), which are embedded in a lipid bilayer. This lipid bilayer is derived from the host cell during virus maturation at the cytoplasmic membrane by a budding process. At the inside of the bilayer the matrix (M) protein is located; it surrounds a helical structure composed of the nucleoprotein (NP) and three minor proteins (P1, P2, P3) (for reviews, see Choppin and Compans, 1975; Rott and Klenk, 1977). The single-stranded viral RNA (vRNA) segments are embedded into this nucleocapsid, which exhibits RNA polymerase activity. This enzyme complex synthesizes exclusively complementary RNA (cRNA) (for a review, see Simpson and Bean, 1975). Thus, in the virion, 7 virus-specific proteins can be recognized. In virus-infected cells at least one additional virus-coded protein is found, which is not incorporated into the virion and, therefore, is called nonstructural (NS) protein (for a review, see Scholtissek and Klenk, 1975). All together, we have to expect a minimum of 8 genes on the influenza genome.

III. The Segmented Genome of Influenza Viruses

During the early genetic work on influenza viruses by Burnet and Hirst and their collaborators, an unexpectedly high rate of recombination between different strains was observed. It was already suggested at this time that the influenza viruses might have a segmented genome. This means that recombinants were formed by reassortment of individual RNA molecules that behave like chromosomes [for early reviews, see Burnet (1959) and Hirst (1962)]. Further evidence in favor of a segmented genome came from observations on multiplicity reactivation (Henle and Liu, 1951) and on stepwise inactivation of influenza virus RNA by specific chemicals (Scholtissek and Rott, 1964). Even under the most cautious conditions influenza virus RNA could be isolated only as a heterogeneous mixture of molecules (Pons and Hirst, 1968; Duesberg, 1968), which could be resolved by polyacrylamide gel electrophoresis in the presence of 6 *M* urea (Floyd *et al.*, 1974) into 8 individual segments (Pons, 1976; Palese and Schulman, 1976a; Bean

and Simpson, 1976; Scholtissek *et al.*, 1976; McGeoch *et al.*, 1976; Rohde *et al.*, 1977). An example of such an electrophoregram of fowl plague virus [A/FPV/Rostock/34 (Hav1N1)] RNA is given in Fig. 1. With influenza B viruses, 8 segments were also resolved (Ritchey *et al.*, 1976a). McGeoch *et al.* (1976) have presented evidence by T1-oligonucleotide fingerprints that each of the 8 bands seen in the electrophoregram consists of a unique RNA molecule. The molecular weights of the RNA segments are between 3×10^5 and 1×10^6. Since the molecular weights of the 8 viral proteins are between 2.5×10^4 (M and NS) and 9×10^4 (P proteins), it has been suggested that each segment consists of a single gene (Skehel, 1972; Pons, 1976; Inglis *et al.*, 1976). A more detailed assignment of the various RNA segments to specific viral proteins will be given below (see Fig. 1).

IV. Base Sequence Studies on Individual RNA Segments

A. Virion RNA

Relatively little is known as yet about the exact base sequence of the various RNA segments, although this field is in rapid development. Earlier studies on influenza virus RNA have revealed that the vRNA carries at its 5′-terminus pppAp (Young and Content, 1971) and at its 3′-end uridine (Lewandowski *et al.*, 1971). In a recent analysis it has been demonstrated that the sequence of the first 13 nucleotides at the 5′-end is the same for all 8 individual vRNA segments of different influenza A strains. Thereafter, a triplet was found, which varies for most of the segments, followed by at least 6 U residues. Thus the sequence of the first 22 nucleotides, except for the triplet in positions 14 to 16, is identical in all segments. The sequence originally published by Skehel and Hay (1978a) has to be corrected in positions 9 and 10 as found by Barry *et al.* (1979). Thus, there is now agreement that the sequence at the 5′-end of the FPV vRNA segments is as follows:

$$\begin{array}{c} \text{UAG} \\ \text{GAG} \\ \text{GUA} \\ \text{5′AGUAGAAACAAGGAGAUUUUUU} \\ \text{UAG} \\ \text{GUG} \end{array}$$

The triplet UAG in positions 14 to 16 was found in RNA segments 1 to 3, the triplet GAG in segment 4, the triplet GUA in segment 5, etc.

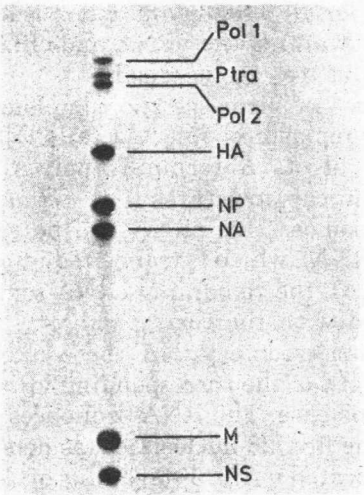

FIG. 1. Polyacrylamide gel electrophoresis of ^{32}P-labeled vRNA of fowl plague virus (FPV), strain "Rostock" (Hav1N1). The virus was grown in primary chick embryo cells in the presence of ^{32}P-orthophosphate. It was purified, and its isolated RNA was separated by electrophoresis on a polyacrylamide slab gel in the presence of 6 *M* urea. The RNA is visualized by exposure of an X-ray film to the gel (Scholtissek *et al.*, 1976). The RNA segments are numbered according to their migration rates, segment 1 (Pol 1) being the slowest moving one. Pol 1, polymerase 1 gene; Ptra, transport gene; Pol 2, polymerase 2 gene; HA, hemagglutinin gene; NP, nucleoprotein gene; NA, neuraminidase gene; M, matrix protein gene; NS, nonstructural protein gene. (By courtesy of V. von Hoyningen.)

(Skehel and Hay, 1978a; Moss *et al.*, 1978; Smith *et al.*, 1978; Barry *et al.*, 1979).

Also the sequence at the 5′-terminus of the 8 segments of an influenza B virus was determined by Skehel and Hay (1978a). Except for the positions 11, 13, and 22, they are the same as found for the two influenza A strains.

B. Complementary RNA

The segments of the 5′-terminus of the *in vitro* products (mRNAs) of the virion transcriptase reaction have been determined (Skehel and Hay, 1978a). The sequence of the first 12 nucleotides is the same for all 8 segments of fowl plague virus (FPV) and another influenza virus A strain (X-31): 5′-AGCAAAAGCAGG. . . .

Since all RNA segments have to be transcribed into cRNA, and cRNA again into vRNA, it is reasonable to assume that the recognition for the enzyme complex is at the very beginning of the 5′-termini,

which are the same for all 8 segments. There is also a certain similarity between the 5′-termini of the vRNA and cRNA segments: 8 out of 12 bases in that sequence are identical.

Two different types of virus-specific complementary RNA can be isolated from the cytoplasm of infected cells (Hay *et al.*, 1977b). (1) One type contains at its 3′-terminus poly(A) and a 5′-terminal 7-methylguanosine in cap structures (Krug *et al.*, 1976) and is found exclusively on polysomes. This cRNA can be regarded as the viral messenger RNA (mRNA), which is translated into viral proteins. After removal of the poly(A), the remaining cRNA segments are somewhat smaller as compared with the corresponding vRNA segments. Furthermore, when hybridized to vRNA these cRNA segments do not protect the total length of the corresponding vRNA segments against digestion with S1 nuclease. The RNA sequences accessible to S1 nuclease attack were the first 28 nucleotides (as determined for segments 5, 6, and 7 of FPV) located at the 5′-terminus of the vRNA. These data suggest that close to this position a termination signal for the synthesis of viral mRNA should be located (Hay *et al.*, 1977a; Skehel and Hay, 1978a). (2) The second type of cRNA does not contain poly(A) at its 3′-terminus, and is not located at the polysomes. The size of these cRNA segments is identical with that of the vRNA segments isolated from virus particles. It is assumed that this type of cRNA functions as a template for the synthesis of vRNA (Hay *et al.*, 1977b).

V. Synthesis of Viral RNA and Its Regulation

This matter has been reviewed recently by Skehel and Hay (1978b). Therefore, only the most important data on the synthesis of viral RNA will be summarized here.

As the first step after adsorption, penetration, and uncoating, mRNA is synthesized by the RNA polymerase complex found in virus particles. This type of synthesis of cRNA is called the primary transcription (Bean and Simpson, 1973; Taylor *et al.*, 1977). This cRNA messenger migrates to the polysomes, where it is translated into viral proteins. Thereafter, both species of cRNAs and some vRNA could be detected in infected cells.

The peak of synthesis of total cRNA precedes that of vRNA (Scholtissek and Rott, 1970; Taylor *et al.*, 1977). Of the two different types of cRNA found in the cytoplasm, the predominant species is the cRNA containing poly(A) at the 3′-terminus (mRNA), which is synthesized presumably by premature termination (see above). The various seg-

ments are present in unequal amounts, and their quantities correlate with the capacity to synthesize the corresponding viral proteins (Hay *et al.*, 1977b; Etkind *et al.*, 1977; Bosch *et al.*, 1978; Inglis *et al.*, 1978). Thus, the synthesis of viral proteins is regulated on the level of the availability of mRNA. In the other species of cRNA, which does not contain poly(A) at the 3′-terminus and is not found on polysomes, the various segments are present in equimolar amounts (Hay *et al.*, 1977a,b). In the presence of cycloheximide only mRNA is synthesized, but not the other cRNA species (template RNA) (Hay *et al.*, 1977b), and as a consequence also no vRNA (Scholtissek and Rott, 1970; Pons, 1973). This suggests that, for the synthesis of template RNA (i.e., to read through beyond the premature termination signal), synthesis of virus-coded protein is required (Skehel and Hay, 1978b). For the transcription of viral RNA the continued transcription of cellular DNA seems to be necessary, since the synthesis of vRNA can be inhibited by actinomycin D or α-amanitin (Scholtissek and Rott, 1970; Rott and Scholtissek, 1970; Pons, 1973, 1977; Lamb and Choppin, 1977; Spooner and Barry, 1977; Hay *et al.*, 1977b; Taylor *et al.*, 1977). The exact mechanisms involved in the regulation of vRNA synthesis and the production of the two types of cRNA are not yet known.

VI. Temperature-Sensitive Mutants

For a virus with a limited number of genes, it can be assumed that at least some of their gene products might exhibit more than one function. By isolating and studying as many independently isolated temperature-sensitive (ts) mutants as possible, not only the exact number of recombination groups, but also the potential multifunctional properties of a certain gene product should be recognizable. Thus, a mutation in a certain region of the gene might abolish one function, leaving the other function(s) intact and vice versa.

A. Number of Recombination Groups

Ts mutants of influenza A viruses were isolated first by Simpson and Hirst (1968) and subsequently by many other groups (Mackenzie, 1970; Mills and Chanock, 1971; Sugiura *et al.*, 1972; Ueda, 1972; Hirst, 1973; Mackenzie and Dimmock, 1973; Markushin and Ghendon, 1973; Ghendon *et al.*, 1973, 1975; Scholtissek *et al.*, 1974; Sugiura *et al.*, 1975; Spring *et al.*, 1975b; Scholtissek and Bowles, 1975; Nakajima and Sugiura, 1977; Almond *et al.*, 1977). Since different influenza A strains

were investigated and in many instances the biological defects were either not clearly defined or could not be related to a certain RNA segment, a comparison of the various recombination–complementation groups obtained by the various groups was rendered very difficult. Hirst (1973) found 8 recombination groups with the WSN (H0N1) strain without characterizing them extensively in biological terms. The ts mutants of the WSN strain isolated by Sugiura *et al.* (1972, 1975) can be placed into 7 clearly defined recombination groups, since the ts defects can be correlated to the various RNA segments (Ritchey and Palese, 1977). Ts mutants of 6 recombination groups of FPV isolated in our laboratory (Scholtissek *et al.*, 1974; Scholtissek and Bowles, 1975), also have been assigned to corresponding RNA segments (Scholtissek *et al.*, 1976) and, therefore, can be compared directly to those described by Sugiura *et al.* (1972, 1975). Almond *et al.* (1977) and recently also Könnecke and Scholtissek (1979) isolated a ts mutant of FPV, which could be correlated to RNA segment 8. Such a group had not been found previously.

Thus, there is no doubt that there exist 8 different groups, which is in agreement with the number of RNA segments found in virus particles and the number of virus-specific proteins (gene products).

B. Defects in Biological Functions

Attention has to be paid to the fact that comparable recombination groups are numbered by the various laboratories in different ways. Furthermore, the numbering of the RNA segments (genes) is different from that of the recombination groups. Therefore, it is time now to find a generally acceptable nomenclature, one that includes a definition of the influenza virus genes on the basis of the specific functions exerted by their gene products. However, such a definition has to be left for an international committee. In this chapter, comparison of the various genes and recombination groups has been done on the basis of rescue experiments using ts mutants, the defects of which can be assigned to specific RNA segments (Scholtissek *et al.*, 1976; Ritchey and Palese, 1977; Almond *et al.*, 1977).

Although many ts mutants have been studied biologically, in the following pages the biological defects of ts mutant groups of only two different influenza A strains [FPV (Hav1N1) and WSN (H0N1)] can be compared directly (Sugiura *et al.*, 1972, 1975; Krug *et al.*, 1975; Ueda and Kilbourne, 1976; Mowshowitz and Ueda, 1976; Palese, 1977; Ritchey and Palese, 1977; Scholtissek *et al.*, 1974, 1976; Scholtissek

and Bowles, 1975; Scholtissek, 1978; Almond *et al.*, 1977). A summary is presented in Table 1.

Ts mutants of group I of FPV correspond to group III mutants of WSN. They have a defect in the synthesis of virus-specific RNA: After shifting the temperature from the permissive to the nonpermissive one, synthesis of both types of viral RNA (cRNA as well as vRNA) is inhibited in cells infected with all mutants of this group so far tested (Scholtissek *et al.*, 1974; Scholtissek and Bowles, 1975; Krug *et al.*, 1975; Palese *et al.*, 1977a).

One ts mutant of group II of FPV, which has been studied in more detail, is identical in its phenotype to mutants of group I (Scholtissek and Bowles, 1975). A corresponding ts mutant of the WSN strain seems to have a defect only in the synthesis of vRNA, but not of cRNA. However, the synthesis of the former has been determined only indirectly; therefore this defect is still somewhat questionable (Krug *et al.*, 1975; Ritchey and Palese, 1977).

Recombination group III of FPV corresponds to group I of WSN. For the gene product of this recombination group it is quite clear that it has more than one function. Two ts mutants of FPV of this group have a defect in the transport of the polymerase complex from the nucleus to the cytoplasm. Virus-specific RNA synthesis itself is not impaired at the nonpermissive temperature. Another mutant belonging to this group (*ts 236*) is, in its phenotype, identical to mutants of group 1 (Scholtissek and Bowles, 1975). Again another ts mutant (*ts 18*) of FPV of group III has been found recently (Heller and Scholtissek, 1979) to belong to this group. It has no defect in viral RNA synthesis at the nonpermissive temperature, although the RNA polymerase activity cannot be detected in cell extracts under these conditions. After double infection with *ts 18* and other ts mutants of the same group, plaques are found at the nonpermissive temperature; however, these plaques cannot be passaged at 40°C (Scholtissek and Bowles, 1975). This appears to be an intracistronic complementation of the corresponding gene product. A ts mutant of WSN belonging to this group was found to be unable to synthesize vRNA as well as cRNA after shifting up the temperature (Krug *et al.*, 1975).

Ts mutants belonging to group IV have a defect in synthesizing a functional neuraminidase at the nonpermissive temperature. The corresponding ts mutants of the WSN strain contain a heat-labile enzyme and are not released from MDBK cells under restrictive conditions. They form clusters of virus particles at the cell surface, which can be broken up by external neuraminidase. Therefore, it has been suggested

TABLE 1

Correlation between Temperature-Sensitive (ts) Defects, Recombination Groups, and Number of RNA Segments according to Their Electrophoretic Migration Rates on Polyacrylamide-Urea Gels of Two Influenza A Viruses[a]

FPV Rostock (Hav1N1)			WSN (H0N1)		
Ts defects	Recombination group	Segment No.	Segment No.	Recombination group	Ts defects
cRNA synthesis	I	1	1	I	cRNA synthesis
Transport, cRNA synthesis	III	2	2	III	cRNA synthesis
cRNA synthesis	II	3	3	II	vRNA synthesis (?)
Hemagglutinin synthesis	V	4	4	VI	Hemagglutinin synthesis
Nucleoprotein function, vRNA synthesis, maturation (?)	VI	5	5	V	Nucleoprotein function, vRNA synthesis (?)
Neuraminidase synthesis	IV	6	6	IV	Heat-labile neuraminidase
	—	7	7	VII	Matrix protein
Nonstructural protein	VIII	8	8	—	—

[a] The arrows indicate that RNA segment 1 of FPV corresponds in function to RNA segment 2 of WSN and vice versa, as determined by corresponding rescue experiments.