Initial Genetic Characterization of the 1918 "Spanish" Influenza Virus

Jeffery K. Taubenberger,*† Ann H. Reid,† Amy E. Krafft, Karen E. Bijwaard, Thomas G. Fanning

The "Spanish" influenza pandemic killed at least 20 million people in 1918–1919, making it the worst infectious pandemic in history. Understanding the origins of the 1918 virus and the basis for its exceptional virulence may aid in the prediction of future influenza pandemics. RNA from a victim of the 1918 pandemic was isolated from a formalin-fixed, paraffin-embedded, lung tissue sample. Nine fragments of viral RNA were sequenced from the coding regions of hemagglutinin, neuraminidase, nucleoprotein, matrix protein 1, and matrix protein 2. The sequences are consistent with a novel H1N1 influenza A virus that belongs to the subgroup of strains that infect humans and swine, not the avian subgroup.

The influenza pandemic of 1918 was exceptional in both breadth and depth. Outbreaks of the disease swept not only North America and Europe but spread as far as the Alaskan wilderness and the most remote islands of the Pacific. Large proportions of the population became ill; 28% of the U.S. population is estimated to have been infected (*1*). The disease was also exceptionally severe, with mortality rates among the infected of over 2.5%, as compared with less than 0.1% in other influenza epidemics (*2*, *3*). Furthermore, in the 1918 pandemic, most deaths occurred among young adults, a group that usually has a very low death rate from influenza. Influenza and pneumonia death rates for 15- to 34-year-olds were more than 20 times higher in 1918 than in previous years (*4*). It has been estimated that the influenza epidemic of 1918 killed 675,000 Americans, including 43,000 servicemen mobilized for World War I (*5*). The impact was so profound as to depress the average life expectancy in the United States by more than 10 years (*6*).

The unusual severity of the 1918 pandemic and the exceptionally high mortality it caused among young adults have stimulated great interest in the influenza strain responsible for the 1918 outbreak. Characterization of this virus may help to elucidate the mechanisms whereby novel influenza viruses evolve and circulate in humans. Because the first human influenza viruses were not isolated until the early 1930s (*7*, *8*), characterization of the 1918 strain has had to rely on indirect evidence. The natural reservoir for influenza virus is thought to be wild waterfowl. Periodically, genetic material from avian strains emerges in strains infectious to humans. Because pigs can be infected with both avian and human strains, they are thought to be an intermediary in this process. Influenza strains with recently acquired genetic material are responsible for pandemic influenza outbreaks (*9*). Analysis of survivor antibody titers from the late 1930s and historical projection of phylogenetic analyses suggest that the 1918 strain was an H1N1-subtype virus, probably closely related to what is now known as classic swine influenza virus (*10*, *11*), which may have emerged from an avian reservoir before the 1918 outbreak.

The Armed Forces Institute of Pathology in Washington, D.C., has autopsy material consisting of formalin-fixed paraffin-embedded tissue and hematoxylin- and eosin-stained sections from U.S. servicemen killed in the 1918 pandemic. We randomly selected 28 cases for pathological review. Of these, the majority died of acute bacterial pneumonia, one of the most common sequelae of the pandemic (*12*). Sections in these cases demonstrated acute lobar pneumonia with massive neutrophilic infiltrates. As influenza virus replication peaks within 2 days in the respiratory tract, with little virus being shed after 6 days (*13*, *14*), we judged it unlikely that these cases would retain influenza virus. Several other cases, in which the victim died within 1 week after initial symptoms, demonstrated a distinct histology noted by pathologists performing autopsies in 1918 (*15*, *16*). These cases often showed massive pulmonary edema and alveolar hemorrhage with acute bronchopneumonia. On the assumption that virus may still have been present in these cases, we chose 14 formalin-fixed, paraffin-embedded tissue specimens representing seven such individuals for further analysis. The pathologic diagnoses noted in these cases were alveolar hemorrhage and bronchopneumonia (three cases), bronchopneumonia (three cases), and one case showing asynchrony between the disease manifestations in the left and right lungs. This case (1918 case 1) showed acute bacterial lobar pneumonia in the left lung but focal acute bronchiolitis and alveolitis in the right lung, which is indicative of a primary viral pneumonia. This was the only case among the 28 examined in which the histologic features of early viral pneumonia were conclusively present.

Control amplification of reverse-transcribed fragments by polymerase chain reaction (RT-PCR) for β-actin was positive in 11 of 14 formalin-fixed, paraffin-embedded tissue samples examined. However, RNA templates larger than 200 nucleotides were not amplifiable. We therefore designed nine degenerate, consensus RT-PCR primer sets to amplify small (under 200 nucleotides) fragments of the genes encoding hemagglutinin (four fragments), neuraminidase, nucleoprotein (two fragments), matrix protein 1, and matrix protein 2. Two influenza strains—one avian [A/Duck/Alberta/35/76 (H1N1)] and one human [A/PR/8/34 (H1N1)]—were used as positive controls for each gene segment to demonstrate that the primers could detect a wide range of influenza strains. No other influenza strains have ever been used in this laboratory. We were able to amplify and sequence all nine fragments of influenza virus RNA from 1918 case 1. Replicate RT-PCR reactions with different tissue blocks from the same case gave identical sequence results. None of the other cases selected by histologic criteria and examined by RT-PCR for influenza virus RNA were positive. The sequence (in the messenger sense orientation) between the primers for each gene fragment is shown in Table 1. For all fragments, the sequence generated from 1918 case 1 was distinct from those of the control strains.

Although the length of sequence between primers was small, phylogenetic analyses for each gene segment were possible with the use of the corresponding regions of other previously sequenced influenza viruses. Analyses were carried out with two computer software packages: Molecular Evolutionary Genetics Analysis (MEGA), version 1.01 (*17*), and Phylogenetic Analysis Using Parsimony (PAUP), version 3.1 (*18*).

Four fragments of the hemagglutinin (HA) gene segment were analyzed. An interprimer sequence of 49 base pairs (bp) (HA137) was initially generated by RT-PCR with primers designed to amplify all H1 sequences. A Basic Local Alignment Search Tool (BLAST) search (*19*) of the files of the National Center for Biotechnology Information indicated that the se-

Division of Molecular Pathology, Department of Cellular Pathology, Armed Forces Institute of Pathology, Washington DC 20306–6000, USA.

*To whom correspondence should be addressed. E-mail: taubenbe@email.afip.osd.mil

†These authors contributed equally to this work.

Table 1. Interprimer 1918 influenza sequences. HA, hemagglutinin; NA, neuraminidase; NP, nucleoprotein; M, matrix (M149 and M811 are fragments of the M1 and M2 genes, respectively). The numbering above sequences refers to the complete sequence of each gene in the sense (mRNA) orientation. Primer sequences are not shown. Methods are described in (*24*).

HA137: 137 AGTACTCGAAAAGAATGTGACCGTGACACACTCTGTTAACCTGCTCGAA 185

HA33: 33 ATGGAGGCAAGACTACTGGTCTTGTTATGTGCATTTGCAGCTACAAATGCAGACACAATATGTATAGGCTACCATGCGAATAACTCAACCGACACTGTTGA

CACAGTACTCGAAAAGAATGTGACCGTGACACACTCTGTTAACCTGCTCGAA 185

HA389: 389 GGAGCAATTGAGCTCAGTGTCATCGTTTGAAAAATTCGAAATATTTCTCAAGACAAGCTCGTGGCCCAATCATGAAACAACCAAAGGTGTAACGGCAGCAT

GCTCCTATGCGGGAGCAAGCAGTTTTT 516

HA980: 980 AATAGGAGAGTGCCCAAAATACGTCAGGAGTACCAAATTGAGGATGGCTACAGGACTAAGAAACATTCCATCTATTCAATCCAGGGGTCTATTT 1073

NA53: 53 GTCAATCTGTATGGTAGTCGGAATAATTAGCCTAATATT 91

NP540: 540 TCTGATGCAGGGTTCAACTCTCCCTAGGAGGTCTGGAGCCGCAGGTGCTGCAGTCAAAGGAGTT 603

NP918: 918 TGAAAGAGAGGGATACTCTCTGGTCGGAATAGACCCTTTCAGACTGCTTCAAAACAGCCAAGTATACAGT 987

M149: 149 AAAGACAAGACCAATCCTGTCACCTCTGACTAAGGGGATTTTAGGATTTGTGTTCACGCTCACCGTGCCCAG 220

M811: 811 TTGATATTGTGGATTCTTGATCGTCTTTTTTTCAAATGCATTTATCGTCGCCTTAAATACGGTTTAAAAGAGGGCCTT 889

quence was unique. The HA137 sequence of 1918 case 1 was compared to 16 other H1-subtype HA genes and 27 H2-H15–subtype HA genes. Sequences were analyzed by the MEGA program with a neighbor-joining (N-J) algorithm and Jukes-Cantor (J-C) distance or p-distance (the proportion of sites that are different). N-J (100 bootstrap replications) placed 1918 case 1 within the H1-subtype HA genes at the 100% level. On the basis of these analyses, primers for two larger regions of HA were designed (HA33, interprimer length 163 bp; and HA389, interprimer length 138 bp). To further analyze these sequences, 30 H1-subtype HA sequences, including both mammalian and avian strains, were chosen from the files of GenBank, the European Molecular Biology Laboratory (EMBL), and the DNA Database of Japan (DDBJ). The result of 500 bootstrap replications for HA33 is shown in Fig. 1A. The sequences were divided into mammalian and avian clusters, with 1918 case 1 falling within the mammalian cluster at the 100% level. An analysis by the unweighted pair-group method with arithmetic means (UPGMA) with J-C or p-distance supported the N-J bootstrap results; avian and mammalian sequences formed two clusters and 1918 case 1 fell within the mammalian group. In the PAUP analysis of HA33, an H3-subtype HA sequence was used as the outgroup. A heuristic search algorithm produced 295 trees of 235 steps. Of these, 100% separated the avian and mammalian sequences into two clusters and placed 1918 case 1 in the mammalian group. Phylogenetic analyses of the third HA fragment, HA389, gave identical results.

Some avian influenza strains of subtypes H5 and H7 are highly pathogenic in certain avian species, including domestic chickens. These strains are characterized by the insertion of coding nucleotides for several basic amino acids at the cleavage site between the HA1 and HA2 subunits of hemagglutinin (*20*). This insertion makes the mutant cleavage site accessible to intracellular proteases and is responsible for the increased virulence of these strains (*21*). This mutation has not been described in mammalian influenza isolates nor in any H1-subtype strains. To test the hypothesis that the 1918 strain had a similar mutation, primers spanning the cleavage site were designed (HA980, interprimer length 94 bp). No insertion at the cleavage site was present. The HA980 sequence of 1918 case 1 was compared to 29 other H1-subtype HA genes with the use of the MEGA program with an N-J algorithm and J-C distance; 94 out of 100 bootstrap replications placed the case in the mammalian cluster.

Our phylogenetic analyses of HA fragments show relationships among known human, swine, and avian H1 strains that are similar to those shown in analyses of the full-length gene (*8*), which suggests that these fragments are representative of the gene as a whole. Analyses of our fragments place HA from 1918 case 1 among the H1 swine virus strains (Fig. 1A).

An interprimer sequence of 38 bp was generated from 1918 case 1 for the neuraminidase (NA) gene and was compared with 11 N1-subtype NA genes and with 16 N2-N9–subtype NA genes. Both N-J (500 bootstrap replications) and UPGMA (500 bootstrap replications) placed 1918 case 1 with the N1-subtype NA genes at the 98 to 100% level (Fig. 1B). A BLAST search indicated that the sequence of the fragment matched two early human strains, A/BH/35 (H1N1) and A/WSN/33 (H1N1), but was distinct from the control strains used in the laboratory. The NA sequence fragment from 1918 case 1 (NA53) is consistent with earlier indirect serological analyses indicating that the 1918 strain was an H1N1 virus.

Two sets of RT-PCR products were produced from the nucleoprotein (NP) gene of 1918 case 1, one yielding 64 bp of interprimer sequence (NP540) and one yielding 70 bp of interprimer sequence (NP918). A BLAST search with both sequences found no perfect matches. Phylogenetic analyses were carried out on these fragments separately. Each 1918 case 1 fragment was compared with 33 NP sequences. By means of MEGA, 78% of 500 N-J bootstrap replications divided the sequences into two clades, one consisting of human and swine sequences, the other of avian and equine sequences. 1918 case 1 fell within the human/swine clade (Fig. 1C). A UPGMA analysis gave an identical result. PAUP also placed the case within the human/swine clade in all 896 trees of length 102 steps.

UPGMA analysis of NP918 produced a tree like the one seen with the NP540 sequence: One large clade consisted exclusively of avian and equine sequences,

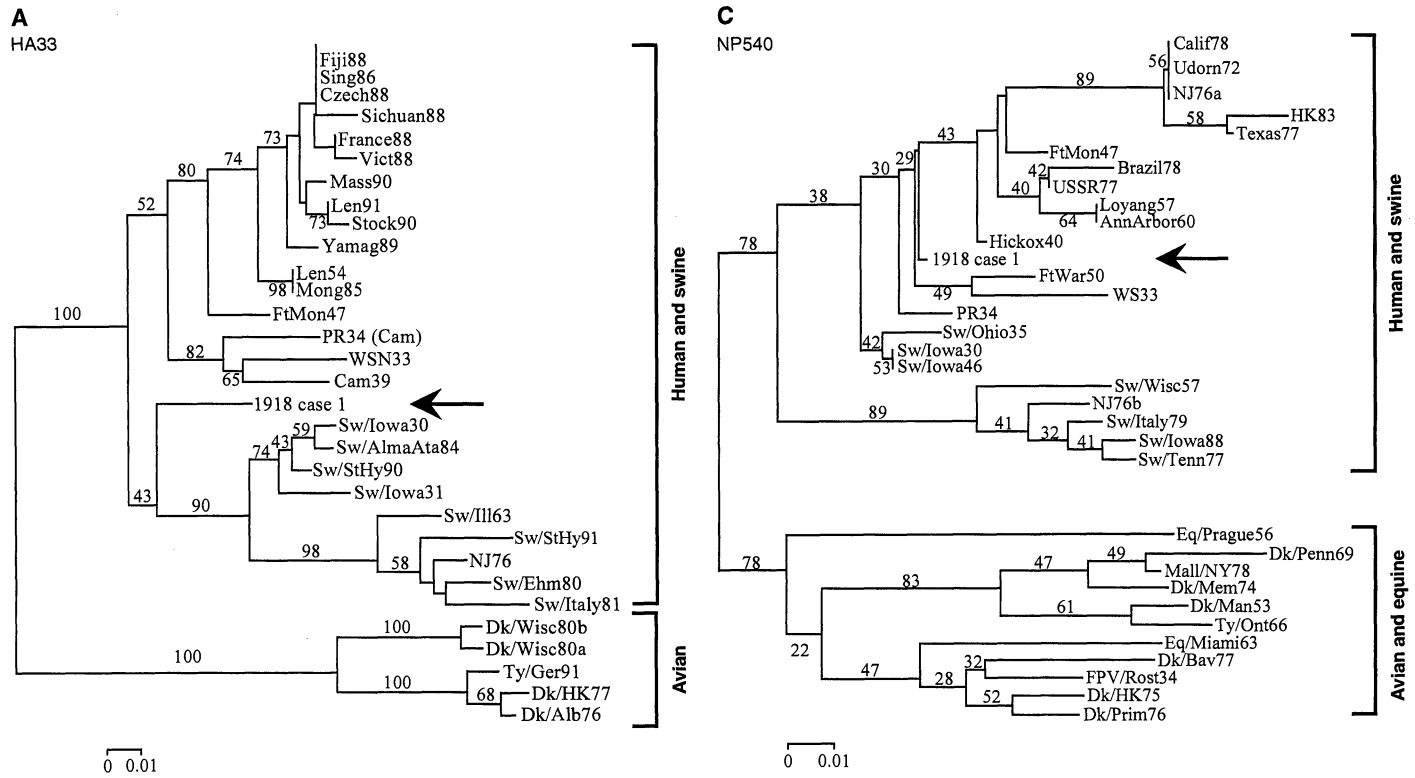

Fig. 1. Influenza phylogenetic trees. All trees are N-J trees using p-distance and were constructed with the MEGA program. Bootstrap values (500 replications) are presented for selected nodes. A branch-length scale bar is shown beneath each tree. (**A**) The hemagglutinin fragment HA33. (**B**) The neuraminidase fragment NA53. (**C**) The nucleoprotein segment NP540. Arrows identify 1918 case 1. Abbreviations used in the analyses are described in (*25*).

whereas another, including 1918 case 1, consisted of only human and swine sequences. PAUP also produced two clades, one entirely human and one entirely avian, as well as a number of unresolved sequences, primarily swine. All 1967 trees of length 136 steps placed 1918 case 1 in the human clade. The N-J analysis of NP918 was less conclusive. As with NP540, two large clades were produced but branch lengths, especially near the root, were very short. In addition, equine sequences were found within the human/swine clade and swine sequences were found within the avian clade. The 1918 case 1 fell very near the root of the tree and within the avian clade.

The influenza virus matrix (M) genes are highly conserved. Although the M2 fragment (M811) was found to be unique, neither the 72-bp region of M1 (M149), nor the 79-bp region of M811 contained enough sequence variation to produce a meaningful phylogenetic tree.

Our data confirm that the 1918 strain was an H1N1 virus, distinct from all subsequently characterized strains. The 1918 case 1 HA sequences appear to be most closely related to early swine influenza strains, corroborating the archaeserological data from the 1930s. Phylogenetic analyses of all influenza genes consistently suggest a common avian ancestor for both human and swine H1N1 virus lineages (*9*, *22*, *23*). It has been suggested that the 1918 strain may itself have been that ancestor (*23*). The fact that our analyses group the 1918 HA, NA, and NP genes with those of mammalian, not avian, viruses may indicate that an ancestral virus had entered the mammalian population at some point before 1918. Because the case originated in Fort Jackson, South Carolina, we propose that the strain be called Influenza A/South Carolina/1/18 (H1N1).

REFERENCES AND NOTES

1. W. H. Frost, *Public Health Rep.* 35, 584 (1920).
2. G. Marks and W. K. Beatty, *Epidemics* (Scribner, New York, 1976), p. 274.
3. M. J. Rosenau and J. M. Last, *Maxcy-Rosenau Preventive Medicine and Public Health* (Appleton-Century-Crofts, New York, 1980), p. 116.
4. F. E. Linder and R. D. Grove, *Vital Statistics Rates in the United States: 1900–1940* (Government Printing Office, Washington, DC, 1947), pp. 254–255.
5. W. Crosby, *America's Forgotten Pandemic: The Influenza of 1918* (Cambridge Univ. Press, Cambridge, 1989), pp. 205–207.
6. R. D. Grove and A. M. Hetzel, *Vital Statistics Rates in the United States: 1940–1960* (Government Printing Office, Washington, DC, 1968), p. 309.
7. W. Smith, C. H. Andrewes, P. P. Laidlaw, *Lancet* **225**, 66 (1933).
8. T. Francis Jr., *Science* **80**, 457 (1934).
9. R. G. Webster, W. J. Bean, O. T. Gorman, T. M. Chambers, Y. Kawaoka, *Microbiol. Rev.* **56**, 152 (1992).
10. R. E. Shope, *J. Exp. Med.* **63**, 669 (1936).
11. R. N. Philip and D. B. Lackman, *Am. J. Hyg.* **75**, 322 (1962).
12. R. E. Shope, *Public Health Rep.* **73**, 165 (1958).
13. A. I. Braude, C. E. Davis, J. Fierer, *Infectious Diseas-*

es and Medical Microbiology (Saunders, Philadelphia, PA, 1986), p. 783.
14. H. Chmel, M. Bendinelli, H. Friedman, *Pulmonary Infections and Immunity* (Plenum, New York, 1994), p. 286.
15. S. B. Wolbach, *Johns Hopkins Hosp. Bull.* **30**, 104 (1919).
16. E. R. LeCount, *J. Am. Med. Assoc.* **72**, 650 (1919).
17. S. Kumar, K. Tamura, M. Nei, *Molecular Evolutionary Genetics Analysis*, version 1.01 (Pennsylvania State University, University Park, PA, 1993).
18. D. L. Swofford, *PAUP: Phylogenetic Analysis Using Parsimony*, version 3.1.1 (Illinois Natural History Survey, Champaign, IL, 1991).
19. S. F. Altschul, W. Gish, W. Miller, E. W. Meyers, D. J. Lipman, *J. Mol. Biol.* **215**, 403 (1990).
20. C. Röhm, T. Horimoto, Y. Kawaoka, J. Süss, R. G. Webster, *Virology* **209**, 664 (1995).
21. T. Horimoto and Y. Kawaoka, *J. Virol.* **68**, 3120 (1994).
22. Y. Kanegae, S. Sugita, K. F. Shortridge, Y. Yoshioka, K. Nerome, *Arch. Virol.* **134**,17 (1994).
23. O. T. Gorman *et al.*, *J. Virol.* **65**, 3704 (1991).
24. RNA lysates from the paraffin-embedded tissues were produced as previously described [A. E. Krafft *et al.*, *J. Wildlife Dis.* **31**, 410 (1995)]. Reverse transcription was carried out at 37°C for 45 min in 20 μl with the use of 300 U of Moloney murine leukemia virus–RT, 1× RT buffer (Gibco-BRL, Grand Island, NY), 5 μM random hexamers, 200 nM deoxynucleotide triphosphates (dNTPs), and 10 mM dithiothreitol. RT reaction (2 μl) was added to a 20-μl PCR reaction containing 50 mM KCl, 10 mM tris, 2.5 mM $MgCl_2$, 1 μM each primer, 100 nM dNTPs, 1 U of Amplitaq Gold (Perkin-Elmer, Foster City, CA), and 2 μCi of [a-^{32}P]dATP (3000 Ci/mmol). Primers (with IUPAC mixed base codes) used in the PCR reactions were as follows (D = A, G, or T; H = A, C, or T; K = G or T; M = A or C; N = A, C, G, or T; R = A or G; S = C or G; W = A or T; Y = C or T); numbers refer to the complete sequence of each gene in the mRNA orientation: HA137 (117 to 208, 92 bp): forward, 5′-TCWACHGACACTGTTGAYACA-3′; reverse, 5′-CAKAGTTTYCCRTTRTGDCTGTC-3′. HA33 (8 to 208, 200 bp): forward, 5′-GCAGGGGAAAATAAAAACAACC-3′; reverse, same as HA137. HA389 (368 to 536, 169 bp): forward, 5′-CATYGACTATGARGARCTGAG-3′; reverse, 5′-CAGCCATAGYAARTTTCTGT-3′. HA980 (960 to 1093, 133 bp): foward, 5′-CAGAATRTACAYCCAGTCAC-3′; reverse, 5′-ATATTTCCRGCAATGGCTCC-3′. NA53 (34 to 113, 80 bp): forward, 5′-CARAARATAATAACMATTGG-3′; reverse, 5′-GWKAYNAKMWTSSCWATYTGCA-3′. NP540 (520 to 623; 103 bp): forward, 5′-ATGGAYCCCAGRATGTGYTC-3′; reverse, 5′-ARYTCCAWNAYCATTGTYCC-3′. NP918 (898 to 1005, 108 bp): forward, 5′-GTRGCCAGYGGRYAYGACTT-3′; reverse, 5′-TYTCRTTYGGTCTRATDAG-3′. M149 (129 to 142, 113 bp): forward, 5′-GAGGYTCTCATGGARTGGCT-3′; reverse, 5′-TACGCTGCAGTCCTCGCTCA-3′. M811 (794 to 906, 113 bp): forward, 5′-TCATTGGGATCTTGCAC-3′; reverse, 5′-WGGYACTCCTTCCGTAG-3′. Cycling conditions were as follows: 9 min at 94°C; 40 cycles of 94°C for 30 s, 50°C (except HA and NA, which were at 45°C) for 30 s, and 72°C for 30 s; then 5 min at 72°C. One-sixth of the reaction was separated on a 7% denaturing polyacrylamide gel, then dried and visualized by autoradiography. Bands were excised, electroeluted, and ethanol-precipitated. One-fourth of the eluted product was added to a 50-μl PCR reaction (50 mM KCl, 10 mM tris, 2.5 mM $MgCl_2$, 200 nM dNTPs, 1 mM each primer, and 1 U Amplitaq Gold) and cycled as above. The reaction product (2 μl) was cloned into the PCR 2.1 vector (Invitrogen, San Diego, CA) according to the manufacturer's instructions. Direct PCR using M13 primers was done on white colonies, and the products were sequenced by cycle sequencing as described [A. H. Reid, R. E. Cunningham, G. Frizzera, T. J. O'Leary, *Am. J. Pathol.* **142**, 395 (1993)].
25. The sequences used in this analysis, obtained from GenBank, EMBL, and DDBJ, were from the following strains: A/Fiji/2/88 (H1N1) [Fiji88], A/Singapore/6/86 (H1N1) [Sing86], A/Czechoslovakia/2/88 (H1N1) [Czech88], A/Yamagata/120/86 (H1N1) [Yamag86], A/Yamagata/32/89 (H1N1) [Yamag89], A/France/15/88 (H1N1) [France88], A/Victoria/43/88 (H1N1) [Vict88], A/Sichuan/4/88 (H1N1) [Sichuan88], A/Massachusetts/1/90 (H1N1) [Mass90], A/Leningrad/109/91 (H1N1) [Len91], A/Stockholm/26/90 (H1N1) [Stock90], A/Leningrad/54/1 (H1N1) [Len54], A/Mongolia/231/85 (H1N1) [Mong85], A/Fort Monmouth/1/47/MA (H1N1) [FtMon47], A/Alma Ata/1417/84 (H1N1) [Sw/AlmaAta84], A/swine/Iowa/1976/31 (H1N1) [Sw/Iowa31], A/swine/Iowa/15/30 (H1N1) [Sw/Iowa30], A/swine/29/37 (H1N1) [Sw/29/37], A/swine/St. Hyacinthe/148/90 (H1N1) [Sw/StHy90], A/swine/Ehime/1/80 (H1N1) [Sw/Ehm80], A/swine/Italy v147/81 (H1N1) [Sw/Italy81], A/swine/St. Hyacinthe/106/91 (H1N1) [Sw/StHy91], A/sw/Illinois/63 (H1N1) [Sw/Ill63], A/NJ/11/76 (X53A) (H1N1) [NJ/1976], A/New Jersey/11/76 (H1N1) [NJ/76], A/Puerto Rico/8/34 (H1N1) [PR34], A/Puerto Rico/8/34 (Cambridge) (H1N1) [PR34(Cam)], A/WSN/33(H1N1) [WSN33], A/swine/Cambridge/39 (H1N1) [Sw/Cam39], A/swine/NI/38 (H1N1) [Sw/NI38], A/turkey/Germany/3/91 (H1N1) [Ty/Germ91], A/DK/HK/196/77 (H1N1) [Dk/HK77], A/DK/WI/259/80 (H1N1) [Dk/Wisc80b], A/duck/Alberta/35/76 (H1N1) [Dk/Alb76], A/DK/WI/1938/80 (H1N1) [Dk/Wisc80a], A/sw/Obihiro/5/92 (H1N1) [Sw/Obih92], A/Hokkaido/2/92 (H1N1) [Hokk92], A/Fort Warren/1/50 (H1N1) [FtWar50], A/Loyang/4/57 (H1N1) [Loyang57], A/FPV/Rostock/34 (H7N1) [FPV/Rost34], A/parrot/Ullster/73 (H7N1) [Parr/Ulls73], A/duck/Ontario/77 (H2N1) [Dk/Ont77], A/tern/Australia/G70C/75 (H11N9) [Tern/Aus75], A/Whale/Maine/1/84 (H13N9) [Whale/Maine84], A/sw/Hong Kong/3/76 (H3N2) [Sw/HK76], A/turkey/Oregon/71 (H7N3) [Ty/Oreg71], A/duck/Alberta/77/77 (H2N3) [Dk/Alb77], A/turkey/Ontario/6118/68 (H8N4) [Ty/Ont68], A/Mallard/Alberta/283/77 (H8N4) [Mall/Alb77], A/duck/England/56 (H11N6) [Dk/Eng56], A/duck/Alberta/28/76 (H4N6) [Dk/Alb76a], A/duck/Alberta/60/76 (H12N5) [Dk/Alb76b], A/shearwater/Australia/72 (H6N5) [Gull/Shear72], A/Ken/1/81 (H3N8) [Eq/Ken81], A/Equine/Alaska/1/91 (H3N8) [Eq/Alask91], A/Cor/16/74 (H7N7) [Eq/Cor74], A/chick/n/Germany/49 (H10N7) [Ck/Germ49], A/California/10/78 (H1N1) [Calif78], A/Udorn/307/72 (H3N2) [Udorn72], A/New Jersey/4/76 (H1N1) [NJ76a], A/Hong Kong/5/83 (H3N2) [HK83], A/Texas/1/77(H3N2) [Texas77], A/Brazil/11/78 (H1N1) [Brazil78], A/USSR/90/77 (H1N1) [USSR77], A/Ann Arbor/6/60 (H2N2) [AnnArbor60], A/Hickox/40 (H1N1) [Hickox40], A/Wilson-Smith/33 (H1N1) [WS33], A/swine/Ohio/23/35 (H1N1) [Sw/Ohio35], A/swine/Iowa/46 (H1N1) [Sw/Iowa46], A/swine/Wisconsin/1/57 (H1N1) [Sw/Wisc57], A/New Jersey/8/76 (H1N1) [NJ76b], A/swine/Italy/2/79 (H1N1) [Sw/Italy79], A/swine/Iowa/17672/88 (H1N1) [Sw/Iowa88], A/swine/Tennessee/24/77 (H1N1) [Sw/Tenn77], A/equine/Prague/1/56 (H7N7) [Eq/Prague56], A/duck/Pennsylvania/1/69 (H6N1) [Dk/Penn69], A/Mallard/NY/6750/78 (H2N2) [Mall/NY78], A/duck/Memphis/928/74 (H3N8) [Dk/Mem74], A/Manitoba/1/53 (H10N7) [Dk/Man53], A/turkey/Ontario/7732/66 (H5N9) [Ty/Ont66], A/equine/Miami/1/63 (H3N8) [Eq/Miami63], A/duck/Bavaria/2/77 (H1N1) [Dk/Bav77], A/duck/Hong Kong/7/75 (H3N2) [Dk/HK75], and A/Anas acuta/Primorje/695/76 (H2N3) [Dk/Prim76].
26. Partially supported by a grant from the American Registry of Pathology and by the intramural funds of the Armed Forces Institute of Pathology. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or Department of Defense. This is a U.S. government work; there are no restrictions on its use.

21 November 1996; accepted 27 January 1997

Modulation of Ras and a-Factor Function by Carboxyl-Terminal Proteolysis

Victor L. Boyartchuk, Matthew N. Ashby, Jasper Rine*

Prenylated proteins contain a covalently linked cholesterol intermediate near their carboxyl-termini. Maturation of most prenylated proteins involves proteolytic removal of the last three amino acids. Two genes in *Saccharomyces cerevisiae*, *RCE1* and *AFC1*, were identified that appear to be responsible for this processing. The Afc1 protein is a zinc protease that participates in the processing of yeast **a**-factor mating pheromone. The Rce1 protein contributes to the processing of both Ras protein and **a**-factor. Deletion of both *AFC1* and *RCE1* resulted in the loss of proteolytic processing of prenylated proteins. Disruption of *RCE1* led to defects in Ras localization and signaling and suppressed the activated phenotype associated with the allele *RAS2^{val19}*.

Eukaryotic cells differ from prokaryotes in their use of specialized compartments and membranes to localize metabolic pathways, structural components, and regulatory processes. Some peripheral membrane proteins depend for proper membrane localization on co- or posttranslational covalent attachment of a lipid moiety to the protein. The most recently discovered such modification involves the covalent attachment of cholesterol biosynthetic intermediates to proteins and is referred to generically as protein prenylation (*1*).

Prenylation is catalyzed by one of three different prenyl-transferase enzymes, which attach either the 15-carbon farnesyl or the 20-carbon geranylgeranyl moieties to a cysteine near the COOH-terminus of the protein. Which enzyme prenylates which substrate is determined by the amino acid sequence at the COOH-terminus of the substrate. A pair of COOH-terminal cysteine residues, together with other determinants, direct geranylgeranylation by a type II pro-

V. L. Boyartchuk and J. Rine, 401 Barker Hall, Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.
M. N. Ashby, Acacia Biosciences Incorporated, 4136 Lakeside Drive, Richmond, CA 94806, USA.
*To whom correspondence should be addressed.