REVISED NOMENCLATURE FOR CORONAVIRUS STRUCTURAL PROTEINS, mRNAs AND GENES

D. Cavanagh, D. Brian, L. Enjuanes, K. Holmes, M. Lai, H. Laude, S. Siddell, W. Spaan, F. Taguchi and P. Talbot

At the Fourth International Symposium on Coronaviruses, July 1989, Cambridge, the Coronavirus Study Group (Vertebrate Virus Subcommittee, International Committee on Taxonomy of Viruses) recommended a simplified nomenclature for coronavirus proteins, mRNAs and genes. This was considered necessary because of the confusion being caused by the use of different terms, acronyms and numbering system. Some papers in this book already contain the new nomenclature while others do not. We present here the "old" and the "new" systems to aid the reader. The Study Group believed that, because of a lack of information, it was inappropriate to make recommendations regarding non-structural proteins. For an introduction to the proteins, mRNAs and genes of coronaviruses, the reader is referred to the review by Spaan et al. (1988) and to Background Papers in this volume by Holmes; Rottier; Masters and Sturman and Lai.

Acronyms for the Structural Proteins

Protein	Former acronym	Recommended acronym
Spike glycoprotein Peplomer glycoprotein	S,E2	S
N-terminal S cleavage product	S1,E2B	S1
C-terminal S cleavage product	S2,E2A	S2
Haemagglutinin-esterase glycoprotein	E3,H,HA	HE
Integral membrane glycoprotein	IMP,M,E1	M
Nucleocapsid protein	N, NC	N

mRNAs

mRNAs are to be referred to by NUMBERS 1,2,3...., starting with the

Coronaviruses and Their Diseases
Edited by D. Cavanagh and T.D.K. Brown
Plenum Press, New York, 1990

genome-sized mRNA. Consequently, the mRNAs of infectious bronchitis virus (IBV), previously denoted as F,E,D,C,B and A should be referred to as 1,2,3,4,5 and 6, respectively. When a protein has a name and an acronym e.g. S, HE, N, M, the corresponding mRNA may be referred to by number, by acronym or both e.g for the IBV mRNA encoding the spike glycoprotein one can use mRNA2, S mRNA or mRNA2 (S), as appropriate.

When "new" mRNAs are discovered, the use of numbers continues. For example, mRNA2 of murine hepatitis virus (MHV) JHM strain encodes a 30K protein. The more recently discovered mRNA which encodes HE should be referred to as "mRNA2-1", since mRNA3 has previously been used to denote the mRNA encoding S of MHV.

Genes/Open-reading-frames (ORFs)

Genes/ORFs are to be referred to by LETTERS. When the corresponding protein has a name, the acronym (UPPER CASE) should be used e.g. S,HE. Otherwise the gene/ORF should be referred to by the number of the corresponding mRNA plus a letter (lower case) when there is more than one ORF. For example, mRNA3 of IBV has three ORFs, 3a, 3b and 3c. Some strains of MHV-A59 do not have a separate message for the HE protein, indeed they do not make HE. Instead mRNA2 encodes, at its 5′ end, a 30K protein (ORF 2a) and also has an untranslated ORF (ORF2b) encoding part of the HE protein.

Reference

1. W. Spaan, D. Cavanagh and M.C. Horzinek. Coronaviruses: Structure and Genome Expression. J.gen.Virol. **69:** 2939 (1988).