THE HIGHLY-CITED SARS RESEARCH LITERATURE

By

Dr. Ronald N. Kostoff
The MITRE Corporation
7515 Colshire Drive
McLean, VA 22102

(The research was funded under MITRE Project 05MSR128-A9: Scientometric Methods and Analysis. The views in this paper are solely those of the author, and do not represent the views of the MITRE Corporation.)

KEYWORDS: Citation Analysis; Literature Survey; Literature Review; SARS; Severe Acute Respiratory Syndrome; Pandemic

ABSTRACT

A chronically weak area in research papers, reports, and reviews is the complete identification of important background reference documents that formed the building blocks for the research. A method for systematically determining these important references is presented. Citation-Assisted Background (CAB) is based on the assumption that important documents tend to be highly cited. Application of CAB to the field of Severe Acute Respiratory Syndrome (SARS) research is presented. While CAB is a highly systematic approach for identifying highly-cited references, it is not a substitute for the judgment of the researchers, and serves as a supplement.

INTRODUCTION

Research is a method of systematically exploring the unknown to acquire knowledge and understanding. Efficient research requires awareness of important prior research and technology that could impact the research topic of interest, and builds upon these past advances to create discovery and new advances. This prior research could originate mainly within the discipline, requiring awareness of a modest number of references. Or, like the newer literature-based discovery approaches [e.g., Swanson, 1986; Kostoff, 2006], it could cover a wide variety of sometimes disparate disciplines, requiring awareness of (or access to) a much larger number of references. The importance of this awareness of prior art is recognized throughout the research community. It is expressed in diverse ways, including requirements for background sections in journal research articles, invited literature surveys in targeted research areas, and required descriptions of prior art in patent applications.

For the most part, development of background material for any of the above applications is relatively slow and labor intensive, and limited in scope. Background material development usually involves some combination of manually sifting through outputs of massive computer searches, manually tracking references through multiple generations, and searching one's own records for personal references. The few studies that have been done on the adequacy of background material in documents show that only a modest fraction of relevant material is included (MacRoberts and MacRoberts, 1989, 1996; Liu, 1993; Calne and Calne, 1992; Shadish et al, 1995; Moravcsik and Murugesan, 1975).

In particular, an analysis of Medline papers on the haemodynamic response to orotracheal intubation showed that recognized deficiencies in research method were not acknowledged. The authors recommended that, when submitting work for publication, investigators should provide evidence of how they searched for previous work (Smith and Goodman, 1997).

Another specific example was provided by MacRoberts and MacRoberts (1997). Their earlier work in a journal on genetics indicated only 30% of influences evident in text were reflected in a paper's references. Later, they studied the text of an issue of Sida to extract influences of previous work evident therein. Influences they judged present in the text appeared in the references only 29% of the time.

Why is this result important? Modern approaches to radical discovery link similar concepts from sometimes very disparate literatures to generate new and

unexpected findings [Smalheiser, 2005; Kostoff, 2006]. Disparate literatures may have very different terminologies, and identifying concepts through text matching may be very limited. Linking similar concepts through shared references may be far more promising in these cases, but requires that the papers be referenced adequately.

Typically missing from standard background section or review article development, as well as in the specific examples cited above, is a systematic approach for identifying the important documents and events that provided the groundwork for the research topic of interest. The present paper presents such a systematic approach for identifying the important documents, called Citation-Assisted Background (CAB). The next section describes the CAB concept, and provides an outline of its operation, with an application to the area of SARS research.

CONCEPT DESCRIPTION (APPROACH)

The CAB concept (Kostoff and Shlesinger, 2004) identifies the highly-cited background documents for a research area using citation analysis. CAB rests on the assumption that a document viewed as a significant building block for a specific research area will typically have been referenced positively by a substantial number of people who are active researchers in that specific area. Implementation of the CAB concept then requires the following steps:

- The research area of interest must be defined clearly
- The documents that define the area of interest must be identified and retrieved
- The references most frequently used in these documents must be identified and selected
- These critical references must be analyzed, and integrated in a cohesive narrative manner to form a comprehensive background section or separate literature survey

These required steps are achieved in the following manner.

1. The research topic of interest is defined clearly by the researchers who are documenting their study results. For example, consider the topic of SARS. In a recent text mining study of SARS, the topical area was defined to include SARS research, clinical issues, and epidemiology-related issues.
2. The topical definition is sharpened further by the development of a literature retrieval query. In the text mining study mentioned above, the literature retrieval query was "((SARS and (coronavirus or infect* or virus* or viral or epidemic* or epidemiology or antibodies or antibody or vaccine* or influenza or pandemic* or outbreak* or syndrome)) OR "sars patient*" or "sars transmission" OR "SARS-CoV" OR "severe acute respiratory syndrome")". Because of the relatively sharp focus of SARS, the query is modest in size. In this particular case, a reasonable approximation would have been just the last term in the query above. Other text mining queries for broader literatures have required hundreds of query terms (Kostoff et al, 2004, 2007).
3. The query is entered into a database search engine, and documents relevant to the topic are retrieved. In the SARS text mining study mentioned above, 2874 documents were retrieved from the Web version of the Science Citation Index (SCI) for the years 2003-early 2008. The SCI was used because it is the only major research database to contain references in a readily extractable format.

4. These documents are combined to create a separate database, and all the references contained in these documents are extracted. Identical references are combined, the number of occurrences of each reference is tabulated, and a table of references and their occurrence frequencies is constructed. In the SARS text mining study, ~45,000 useful separate references were extracted and tabulated. Table 1 contains the ten highest frequency (most cited) references extracted from the SARS database.

TABLE 1

CITED AUTHOR	TIME	JOURNAL	VOL	PAGE	# CIT
KSIAZEK TG ***A novel coronavirus associated with SARS***	2003	NEW ENGL J MED	V348	P1953	903
DROSTEN C ***Identification of a novel coronavirus in patients with SARS***	2003	NEW ENGL J MED	V348	P1967	838
ROTA PA ***Characterization of a novel coronavirus associated with SARS***	2003	SCIENCE	V300	P1394	796
PEIRIS JSM ***Coronavirus as a possible cause of SARS***	2003	LANCET	V361	P1319	750
MARRA MA ***The genome sequence of the SARS-associated coronavirus***	2003	SCIENCE	V300	P1399	719
LEE N ***Features of a major outbreak of SARS in Hong Kong***	2003	NEW ENGL J MED	V348	P1986	551
PEIRIS JSM ***Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia***	2003	LANCET	V361	P1767	475
LI WH ***Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus***	2003	NATURE	V426	P450	410
POUTANEN SM	2003	NEW ENGL J MED	V348	P1995	399

Identification of SARS in Canada

TSANG KW 2003 NEW ENGL J MED V348 P1977 354
Clinical features of a cluster of cases of SARS in Hong Kong

Two frequencies are computed for each reference, but only the first is shown in Table 1. The frequency shown in the rightmost column is the number of times each reference was cited by the 2874 records in the retrieved database only. This number reflects the importance of a given reference to the specific discipline of SARS. The second frequency number (not shown) is the total number of citations the reference received from all sources, and reflects the importance of a given reference to all the fields of science that cited the reference. This number is obtained from the citation field or citation window in the SCI. In CAB, only the first frequency is used, since it is topic-specific. Using the first discipline-specific frequency number obviates the need to normalize citation frequencies for different disciplines (as a consequence of different levels of activity in different disciplines), as would be the case if total citation frequencies were used to determine the ordering of the references.

Before presenting a specific implementation algorithm for the SARS application, a few caveats will be discussed. First, listing and selection of the most highly cited references are dependent on the comprehensiveness and balance of the total records retrieved. Any imbalances (from skewed databases or incorrect queries) can influence the weightings of particular references, and result in some references exceeding the selection threshold where not warranted, and others falling below the threshold where not warranted.

Second, it is important that the query used for record retrieval be extensive (Khan and Khor, 2004; Harter and Hert, 1997; Kantor, 1994), as was shown for the SARS application. The query needs to be checked for precision and recall, which becomes complicated when assumptions of binary relevance and binary retrieval are relaxed (Della Mea and Mizzaro, 2004). There are myriad issues to be considered when evaluating queries and their impact on precision and recall. A 2004 systems analytic approach to analyzing the information retrieval process concluded that, for completeness, the interaction of the environment and the information retrieval system must be considered in query development (Kagolovsky and Moehr, 2004). The author's experiences (with the five studies done so far with CAB, including the study reported in this paper) have shown that

modest query changes may substitute some papers at the citation selection threshold, but the truly important papers have citations of such magnitude that they are invulnerable to modest query changes. For this reason, the cutoff threshold for citations has been, and should be, set slightly lower, to compensate for query uncertainties.

Third, there may be situations where at least minimal citation representation is desired from each of the major technical thrust areas in the documents retrieved. In this case, the retrieved documents could be clustered into the major technical thrust areas, and the CAB process could be performed additionally on the documents for each cluster. The additional references identified with the cluster-level CAB process, albeit with lower citations than from the aggregated non-clustered CAB process, would then be added to the list obtained with the aggregated CAB process. The author has not found this cluster-level CAB process necessary for any of the disciplines studied with CAB so far.

Fourth, there may be errors in citation counts because of references errors, and the subsequent fragmenting of a reference's occurrence frequency metric into smaller metric values. Care needs to be taken in insuring that a given reference is not divided into multiple large fragments, which are not subsequently combined.

Fifth, the CAB approach is most accurate for recent references, and its accuracy drops as the references recede into the distant past. This derives from the tendency of authors to reference more recent documents and, given the restricted real estate in journals, not reference the original documents. To get better representation, and more accurate citation numbers, for early historical documents, the more recent references need to be retrieved, collected into a database, and have their references analyzed in a similar manner (essentially examining generation of citations).

Sixth, high citation frequencies are not unique to important documents only; different types of references can have high citation frequencies. Documents that contain critical research advances, and were readily accessible in the open literature, tend to be cited highly, and represent the foundation of the CAB approach. Application of CAB to four technical research areas so far (in addition to the present SARS study) shows that this type of document is predominant in the highly cited references list. Books or review articles also appear on the highly cited references list. These documents do not usually represent new advances, but rather are summaries of the state of the art (and its background) at the time the document was written. These types of documents are still quite useful as background material. Finally, documents that receive large numbers of citations

highly critical of the document could be included in the list of highly cited documents. In the five studies so far, the author has not identified such papers in the detailed development of the background.

Additionally, one of the five application studies concerns high speed compressible flow, a discipline in which the author worked decades ago. Using the CAB approach, the author found that all the key historical documents with which he was familiar were identified, and all the historical documents identified appeared to be important. Thus, for that data point at least, the weaknesses identified above (imbalances, undervaluing early historical references, unwanted highly cited documents) did not materialize. To ensure that any critical documents were not missed because of imbalance problems, the threshold was set a little bit lower to be more inclusive.

The converse problem to multiple types of highly cited references, some of which may not be the important documents desired, is influential references that do not have substantial citation frequencies. If the authors of these references did not publish them in widely and readily accessible forums, or if they do not contain appropriate verbiage for optimal query accessibility, then they might not have received large numbers of citations. Additionally, journal or book space tends to be limited, with limited space for references. In this zero-sum game for space, research authors tend to cite relatively recent records at the expense of the earlier historical records. Also, extremely recent but influential references have not had the time to accumulate sufficient citations to be listed above the selection threshold on the citation frequency table. Methods of including these influential records located at the wings of the temporal distribution will be described in the following implementation section. Inclusion of the references that were not widely available when published is more problematical, and tends to rely on the background developers' personal knowledge of these documents, and their influence.

CONCEPT IMPLEMENTATION (RESULTS)

Important Documents

To identify the total candidate references for the background section, a table similar in structure to Table 1, but containing all the references from the retrieved records, is constructed. A threshold frequency for selection can be determined by arbitrary inspection (e.g., a background section consisting of 150 key references is arbitrarily selected). The author has found a dynamic selection process more useful. In this dynamic process, references are selected, analyzed, and grouped based on their order in the citation frequency table until the resulting background is judged sufficiently complete by the background developers.

To ensure that the influential documents at the wings of the temporal distribution are included, the following total process is used. The reference frequency table is ordered by inverse frequency, as above, and a high value of the selection frequency threshold is selected initially. Then, the table is re-ordered chronologically. The early historical documents with citation frequencies substantially larger than those of their contemporaries are selected, as are the extremely recent documents with citation frequencies substantially larger than those of their contemporaries. By contemporaries, it is meant documents published in the same time frame, not limited to the same year. Then, the dynamic selection process defined above is applied to the early historical references, the intermediate time references (those falling under the high frequency threshold), and the extremely recent references.

Table 2 contains the final references selected by the process. The fourth reference and first bolded paper listed, Kermack's 1927 paper, had many more citations (27) than any papers published previously, and was next exceeded by Reed's 1938 paper. In turn, Reed's paper was only exceeded 54 years later by Yaeger's 1992 paper in Nature. The bolded and italicized papers especially stood out in relation to their contemporaries, while the underlined papers stood out more than the threshold for this list. This is a graphic example of how we interpret a paper's having substantially more citations than its contemporaries.

TABLE 2 – DOCUMENTS SELECTED FOR LITERATURE REVIEW

CITED AUTHOR	**YEAR**	**SOURCE**	**VOL**	**PAGE**	**# CIT**
SPEARMAN C	1908	BRIT J PSYCHOL	V2	P227	3
ROSS R	1911	PREVENTION MALARIA			5
HARRIS JW	1919	J AMER MED ASSOC	V72	P978	3
KERMACK WO	***1927***	***P ROY SOC LOND A MAT***	***V115***	***P700***	***18***

KARBER G	1931	ARCH EXP PATHOL PH	V162	P480	10
REED LJ	***1938***	***AM J HYG***	***V27***	***P493***	***42***
TYRRELL DAJ	1965	BRIT MED J	V1	P1467	10
HAMRE D	1966	P SOC EXP BIOL MED	V121	P190	16
SCHECHTER I	1967	BIOCHEM BIOPH RES CO	V27	P157	15
MCINTOSH K	1967	P NATL ACAD SCI USA	V57	P933	12
FELLER W	1968	INTRO PROBABILITY TH	V1		10
LAEMMLI UK	1970	NATURE	V227	P680	12
MCINTOSH K	1970	AM J EPIDEMIOL	V91	P585	9
SIDWELL RW	1972	SCIENCE	V177	P705	11
MCINTOSH K	***1974***	***J INFECT DIS***	***V130***	***P502***	***22***
BAILEY NTJ	1975	MATH THEORY INFECT D			11
MOSER MR	1979	AM J EPIDEMIOL	V110	P1	19
HOROWITZ M	1979	PSYCHOSOM MED	V41	P209	16
STURMAN LS	1980	J VIROL	V33	P449	16
WEISS RC	***1981***	***COMP IMMUNOL MICROB***	***V4***	***P175***	***25***
COLLINS AR	1982	VIROLOGY	V119	P358	19
SIDDELL S	1983	J GEN VIROL	V64	P761	16
TOOZE J	1984	EUR J CELL BIOL	V33	P281	13
FRANA MF	1985	J VIROL	V56	P912	17
STURMAN LS	1985	J VIROL	V56	P904	13
FUERST TR	1986	P NATL ACAD SCI USA	V83	P8122	18
BRIERLEY I	1987	EMBO J	V6	P3779	17
DEGROOT RJ	1987	J MOL BIOL	V196	P963	16
SPAAN W	1988	J GEN VIROL	V69	P2939	17
BARIC RS	1988	J VIROL	V62	P4280	16
GORBALENYA AE	1989	NUCLEIC ACIDS RES	V17	P4847	30
BRIERLEY I	1989	CELL	V57	P537	24
SAMBROOK J	1989	MOL CLONING LAB MANU			23
VENNEMA H	1990	J VIROL	V64	P1407	31
ANDERSON RM	1991	INFECT DIS HUMANS DY			36
JONES TA	1991	ACTA CRYSTALLOGR A	V47	P110	26
WILLIAMS RK	1991	P NATL ACAD SCI USA	V88	P5533	22
YEAGER CL	***1992***	***NATURE***	***V357***	***P420***	***67***
DELMAS B	1992	NATURE	V357	P417	39
LASKOWSKI RA	1993	J APPL CRYSTALLOGR	V26	P283	37
THOMPSON JD	1994	NUCLEIC ACIDS RES	V22	P4673	45
KUBO H	1994	J VIROL	V68	P5403	38
VENNEMA H	1996	EMBO J	V15	P2020	57
LAI MMC	***1997***	***ADV VIRUS RES***	***V48***	***P1***	***103***
OTWINOWSKI Z	1997	METHOD ENZYMOL	V276	P307	38
ZIEBUHR J	1997	J VIROL	V71	P3992	31
BRUNGER AT	1998	ACTA CRYSTALLOGR D 5	V54	P905	39
SAWICKI SG	1998	ADV EXP MED BIOL	V440	P215	35
SANCHEZ CM	1999	J VIROL	V73	P7607	38
ZIEBUHR J	***2000***	***J GEN VIROL 4***	***V81***	***P853***	***99***
KUO LL	2000	J VIROL	V74	P1393	49
TIPNIS SR	2000	J BIOL CHEM	V275	P33238	46
DONOGHUE M	2000	CIRC RES	V87	E1	37
CORSE E	2000	J VIROL	V74	P4319	36
GALLAGHER TM	2001	VIROLOGY	V279	P371	95

LAI MMC	2001	FIELDS VIROLOGY		P1163	80
ANAND K	2002	EMBO J	V21	P3213	89
DEHAAN CAM	2002	VIROLOGY	V296	P177	47
CRACKOWER MA	2002	NATURE	V417	P822	45
GOSERT R	2002	J VIROL	V76	P3697	40
HEGYI A	2002	J GEN VIROL 3	V83	P595	39
KSIAZEK TG	***2003***	***NEW ENGL J MED***	***V348***	***P1953***	***903***
DROSTEN C	***2003***	***NEW ENGL J MED***	***V348***	***P1967***	***838***
ROTA PA	***2003***	***SCIENCE***	***V300***	***P1394***	***796***
PEIRIS JSM	***2003***	***LANCET***	***V361***	***P1319***	***750***
MARRA MA	***2003***	***SCIENCE***	***V300***	***P1399***	***719***
LEE N	2003	NEW ENGL J MED	V348	P1986	551
PEIRIS JSM	2003	LANCET	V361	P1767	475
LI WH	2003	NATURE	V426	P450	410
POUTANEN SM	2003	NEW ENGL J MED	V348	P1995	399
TSANG KW	2003	NEW ENGL J MED	V348	P1977	354
GUAN Y	2003	SCIENCE	V302	P276	352
FOUCHIER RAM	2003	NATURE	V423	P240	293
SNIJDER EJ	2003	J MOL BIOL	V331	P991	275
BOOTH CM	2003	JAMA-J AM MED ASSOC	V289	P2801	267
KUIKEN T	2003	LANCET	V362	P263	261
DONNELLY CA	2003	LANCET	V361	P1761	218
NICHOLLS JM	2003	LANCET	V361	P1773	215
ANAND K	2003	SCIENCE	V300	P1763	211
RUAN YJ	2003	LANCET	V361	P1779	186
THIEL V	2003	J GEN VIROL 9	V84	P2305	181
SETO WH	2003	LANCET	V361	P1519	160
RILEY S	2003	SCIENCE	V300	P1961	157
YANG HT	2003	P NATL ACAD SCI USA	V100	P13190	147
LIPSITCH M	2003	SCIENCE	V300	P1966	145
MARTINA BEE	2003	NATURE	V425	P915	136
PEIRIS JSM	2003	NEW ENGL J MED	V349	P2431	135
SO LKY	2003	LANCET	V361	P1615	115
XIAO XD	2003	BIOCHEM BIOPH RES CO	V312	P1159	109
BOSCH BJ	2003	J VIROL	V77	P8801	104
CINATL J	2003	LANCET	V362	P293	92
WONG SK	***2004***	***J BIOL CHEM***	***V279***	***P3197***	***152***
YANG ZY	***2004***	***NATURE***	***V428***	***P561***	***151***
SUBBARAO K	2004	J VIROL	V78	P3572	139
HE JF	2004	SCIENCE	V303	P1666	133
BISHT H	2004	P NATL ACAD SCI USA	V101	P6641	122
SUI JH	2004	P NATL ACAD SCI USA	V101	P2536	120
VANDERHOEK L	2004	NAT MED	V10	P368	119
BUKREYEV A	2004	LANCET	V363	P2122	94
LIU SW	2004	LANCET	V363	P938	91
JEFFERS SA	2004	P NATL ACAD SCI USA	V101	P15748	88
HAMMING I	2004	J PATHOL	V203	P631	83
HAAGMANS BL	2004	NAT MED	V10	P290	82
BABCOCK GJ	2004	J VIROL	V78	P4552	81
YANG ZY	2004	J VIROL	V78	P5642	80
BUCHHOLZ UJ	2004	P NATL ACAD SCI USA	V101	P9804	79

SIMMONS G	2004	P NATL ACAD SCI USA	V101	P4240	78
BOSCH BJ	2004	P NATL ACAD SCI USA	V101	P8455	73
YU ITS	2004	NEW ENGL J MED	V350	P1731	73
FAN KQ	2004	J BIOL CHEM	V279	P1637	72
LAU SKP	***2005***	***P NATL ACAD SCI USA***	***V102***	***P14040***	***120***
LI WD	***2005***	***SCIENCE***	***V310***	***P676***	***114***
WOO PCY	2005	J VIROL	V79	P884	97
SONG HD	2005	P NATL ACAD SCI USA	V102	P2430	82
LI F	2005	SCIENCE	V309	P1864	72
LI WH	2005	EMBO J	V24	P1634	63
SPIEGEL M	2005	J VIROL	V79	P2079	29
HUANG IC	2006	J BIOL CHEM	V281	P3198	29
LI WH	2006	J VIROL	V80	P4211	24
TANG XC	2006	J VIROL	V80	P7481	20
CHAN VSF	2006	NAT GENET	V38	P38	19
RATIA K	2006	P NATL ACAD SCI USA	V103	P5717	19
SPRUTH M	2006	VACCINE	V24	P652	19
KAMITANI W	2006	P NATL ACAD SCI USA	V103	P12885	18
ROBERTS A	2006	J INFECT DIS	V193	P685	17
WOO PCY	2006	VIROLOGY	V351	P180	17
GORBALENYA AE	2006	VIRUS RES	V117	P17	16
KOPECKYBROMBERG SA	2006	J VIROL	V80	P785	16
SURJIT M	2006	J BIOL CHEM	V281	P10669	16
KOPECKYBROMBERG SA	2007	J VIROL	V81	P548	28
ROBERTS A	2007	PLOS PATHOG	V3	P23	25
MCCRAY PB	2007	J VIROL	V81	P813	18
TSENG CTK	2007	J VIROL	V81	P1162	17
STERTZ S	2007	VIROLOGY	V361	P304	13
FRIEMAN M	2007	J VIROL	V81	P9812	11
ROCKX B	2007	J VIROL	V81	P7410	11
VIJAYKRISHNA D	2007	J VIROL	V81	P4012	10
XUE XY	2007	J MOL BIOL	V366	P965	9
ROBERTS A	2008	VIRUS RES	V133	P20	6
BOSCH BJ	2008	J VIROL	V82	P8887	5
FRIEMAN M	2008	VIRUS RES	V133	P101	5
NARAYANANJ K	2008	J VIROL	V82	P4471	5
XUE XY	2008	J VIROL	V82	P2515	5

These results were examined by the author. All papers in the table were judged to be relevant for a background section, or review paper. Because of space considerations, not all papers listed will be included in the historical narrative shown in the next section.

The analysis and discussion above have focused on the contents of the background; i.e., which documents should be included. In some cases, including the present study, the abstracts of the important references have been retrieved and clustered, to produce a structure for the background. Thus, the CAB approach can be used to

determine both the content and structure of the background section. Again, CAB does not exclude content and structure determinations by the experts. CAB can be viewed as the starting point for content and structure determination, upon which the experts can build with their own insights and experience.

While the CAB approach is systematic, it is not automatic. Judgment is required to determine when adequate numbers of references have been selected for the background, and further judgment is required to analyze, group, and link the references to form a cohesive background section. Additionally, the highly influential references that were not highly cited as a consequence of insufficient dissemination should be included by the background developers, if they know of such documents. CAB is not meant to replace individual judgment or specification of background material. CAB is meant to augment individual judgment and reference selection, as reflected in its name of Citation-Assisted.

Literature Review

The following is a review of the important background SARS literature. There are three major components: epidemiology, clinical aspects (detection and treatment), and laboratory research. The latter section is structured on a clustering of key phrases. Within each section, the references and key events are tracked chronologically.

EPIDEMIOLOGY

The epidemiology section is divided into five sub-sections, based on conceptual sequences in the etiology of the epidemic. It tracks the events that led to identifying the sources of the virus, and the subsequent cross-species transmission of the virus/disease. Then, it transitions into the tracking and modeling of the global epidemic, and ends with a discussion of how the viral spread could be controlled within the confines of treatment facilities and in the larger society.

Sources of Virus

THOMPSON (1994) described a new program CLUSTAL W that improved the sensitivity of progressive multiple sequence alignment for DNA or proteins through sequence weighting, position-specific gap penalties and weight matrix choice. It allows evolutionary relationships to be seen via viewing Cladograms or Phylograms. A decade later, after the SARS epidemic had peaked then receded, LAU et al (2005) identified severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. At the same time, LI et al (2005) demonstrated that bats are natural reservoirs of SARS-like coronaviruses. A year later, TANG et al (2006) described prevalence and genetic diversity of coronaviruses in bats from China, showing that bats may play an integral role in the ecology and evolution of coronaviruses. In parallel, WOO et al (2006) described molecular diversity of coronaviruses in bats, revealing six novel coronaviruses from six different bat species, in addition to the two previously described coronaviruses. These important papers, and the supportive research, solidified the belief that bats were the natural reservoirs of the SARS coronavirus. The question that follows immediately is: how does the virus transmit from bats to human beings?

Cross-Species Transmission

The initial seminal paper on cross-species transmission occurred slightly after the turn of the twentieth century. ROSS (1911) proved the role of Anopheles

mosquitoes in the transmission of malaria parasites in humans, an early demonstration of cross-species infection. His initial work on the topic (for which he received the Nobel Prize in Medicine in 1902) was done and reported in the late 1890s. Fifty years later, MCINTOSH (1967) demonstrated that viruses recovered from upper respiratory washings in adults exhibited an unusual morphology closely resembling that of avian infectious bronchitis virus, as well as two strains recovered from man: 229E and B814. The next major event from the perspective of the SARS literature occurred 35 years later when GUAN (2003) described the isolation and characterization of viruses related to the SARS coronavirus from animals (especially bats) in southern China. At the same time, FOUCHIER (2003) demonstrated Koch's postulates fulfilled for SARS virus, where SARS-CoV-infected macaques developed a disease comparable to SARS in humans; the virus was re-isolated from these animals and they developed SARS-CoV-specific antibodies. In parallel, MARTINA (2003) showed that ferrets (Mustela furo) and domestic cats (Felis domesticus) are susceptible to infection by SARS coronavirus (SCV) and that they can efficiently transmit the virus to previously uninfected animals that are housed with them. The observation that these two distantly related carnivores can so easily be infected with the virus indicated that the reservoir for this pathogen may involve a range of animal species.

Two years later, SONG (2005) described cross-host evolution of SARS coronavirus in palm civets and humans, and concluded that major genetic variations in some critical genes, particularly the Spike gene, seemed essential for the transition from animal-to-human transmission to human-to-human transmission. The following year, LI (2006) described animal origins of the SARS coronavirus, including insights from ACE2-S-protein interactions. Further, VIJAYKRISHNA (2007) provided evolutionary insights into the ecology of coronaviruses, suggesting that bats are likely the natural hosts for all presently known coronavirus lineages and that all coronaviruses recognized in other species were derived from viruses residing in bats.

Tracking Epidemic

An early influential paper by MOSER (1979) described an outbreak of influenza on-board a commercial airliner. Twenty-five years later, DONNELLY (2003) identified epidemiological determinants of spread of causal agent of SARS in Hong Kong. At the same time, PEIRIS (2003) reviewed the cause, epidemiology, and clinical features of SARS, and was the first to observe the cytopathic effects of the virus. One year later, HE (2004) described molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. He showed that the

Spike protein exhibited the strongest initial responses to positive selection pressures, followed by subsequent purifying selection and eventual stabilization. Two years later, CHAN (2006) showed that homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection.

Modeling Epidemic

Seminal epidemic modeling studies have a long history. SPEARMAN (1908) generalized a method for treating 'sensitivity data' (experimental data for which the measurement at any point in the scale destroys the sample). A decade later, HARRIS (1919) performed a statistical study of 130 cases of Influenza occurring in pregnant women, including effects of mortality and miscarriage. After another decade, KERMACK (1927) generated a mathematical theory of epidemics, focused on the formal treatment of the course of epidemics in closed populations, including the influence of population density, infectivity of affected individuals, and probability of death or recovery after infection. Four decades later, FELLER (1968) produced an introductory book on probability theory, with emphasis on discrete probability. Almost another decade passed before BAILEY's (1975) classical text on the mathematical theory of infectious diseases and its applications was published.

ANDERSON's (1991) seminal work (Infectious Diseases of Humans: Dynamics and Control) covers infectious diseases -- viral, bacterial, protozoan and helminth -- in terms of the dynamics of their interaction with host populations, combining mathematical models with extensive use of epidemiological and other data. During the SARS epidemic, RILEY (2003) showed the impact of public health interventions on transmission dynamics of the etiological agent of SARS in Hong Kong. At the same time, LIPSITCH (2003) described the transmission dynamics and control of SARS, using detailed epidemiologic data from Singapore and epidemic curves from other settings.

Controlling the Spread of SARS

Both the stress associated with SARS and the infection had to be controlled. The seminal stress control paper was by HOROWITZ (1979), which contained the Impact of Event Scale: a measure of subjective stress related to a specific event. The important paper by SETO (2003), published during the SARS epidemic, showed the effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of SARS. One year later, YU (2004) demonstrated evidence of airborne transmission of the SARS virus. There were also many lesser

cited papers at this time addressing the control of infection spread in healthcare facilities as well as the value of quarantine measures on broader scale control of epidemic.

CLINICAL

There were many papers written during the 2003-2004 timeframe on clinical aspects of SARS, focusing on the detection and treatment problem. The most highly cited of these follow.

Detecting SARS

In the development of the structure of this background important literature survey, one technique for generating categories was factor matrix. Approximately a thousand of the highest frequency research/laboratory-related phrases were selected from the total SARS retrieved corpus, and a factor matrix was generated. There were 8-10 distinguishable themes (factors), and most of them were used in the laboratory research section (next major section). However, two of the themes applied to the present section. They were diffuse alveolar damage, which was characteristic of SARS patients, and real-time PCR, an established technique used in the disease detection phase. Thus, these factors were not research themes, but rather applications of known techniques used for detection.

An early seminal paper by TYRRELL (1965) showed that an increase in virus isolation from patients with common colds was achieved through the use of human embryonic tracheal and nasal organ cultures, resulting in cultivation of a novel type of common-cold virus in organ cultures. Along the same lines, one year later HAMRE (1966) showed the isolation of the 229E coronavirus from the human respiratory tract. MCINTOSH's (1970) sero-epidemiologic study of infection by coronavirus strains 229E, OC38, OC43, and mouse hepatitis virus (MHV) strain A-59 showed that, in hospitalized children, infection with 229E was rare; infection with OC38, OC43, and MHV occurred less often in hospitalized children with lower respiratory tract disease than in a control group with non-respiratory tract disease. A few years later, MCINTOSH (1974) described coronavirus infection in acute lower respiratory tract disease of infants.

Three decades later, during the SARS epidemic, many highly cited papers were produced. Only the top levels of these were selected as the most important papers. KSIAZEK (2003) identified a novel coronavirus associated with SARS, and showed that it could be grown in cell culture in Vero/African green monkey kidney

cells. DROSTEN (2003) identified a novel coronavirus (only distantly related to known coronaviruses) in patients with SARS. ROTA (2003) characterized a novel coronavirus associated with SARS, and showed the genome of SARS-CoV is 29,727 nucleotides in length and has 11 open reading frames, and downstream of ORF1a and ORF1b are ORFs that encode the four main structural proteins. He concluded that its genome organization is similar to that of other coronaviruses. Phylogenetic analyses and sequence comparisons showed that SARS-CoV is not closely related to any of the previously characterized coronaviruses. PEIRIS (2003) demonstrated coronavirus as a possible cause of SARS by isolating a virus belonging to the family Coronaviridae from two patients, and used serological and reverse-transcriptase PCR specific for this virus to show that 45 of 50 patients with SARS, but no controls, had evidence of infection with this virus. MARRA (2003) generated the genome sequence of the SARS-associated coronavirus, and showed that this coronavirus is only moderately related to other known coronaviruses, including two human coronaviruses, HCoV-OC43 and HCoV-229E. POUTANEN (2003) described identification of SARS in Canada. TSANG (2003) described a cluster of cases of SARS in Hong Kong, suggesting that a virus may have been the cause of the illness. KUIKEN (2003) showed newly discovered coronavirus as the primary cause of SARS. NICHOLLS (2003) assessed lung pathology of fatal SARS. VANDERHOEK (2004) reported the identification of a fourth human coronavirus, HCoV-NL63, whose complete genome sequence indicated that this virus was not a recombinant, but rather a new group 1 coronavirus. WOO (2005) characterized and generated a complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia.

Treating SARS

LEE (2003) described detection and treatment of a major outbreak of SARS in Hong Kong. PEIRIS (2003) produced a prospective study evaluating clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia. BOOTH (2003) described clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area. SO (2003) described development of a standard treatment protocol for SARS.

LABORATORY

As mentioned previously, the categories in this section were generated with help of a factor matrix, with the factors being the category themes. To aid in describing each theme more precisely, the key phrases for each factor are presented

immediately following each category's theme. They can be viewed as the definition for the theme.

Heptad Repeat

heptad; repeat; HR1; HR2; fusion; bundle; entry; glycoprotein; HR; Spike; membranes; S; membrane; conformation; transmembrane; protein; binding

DEGROOT (1987) provided evidence for a coiled structure in the Spike proteins of coronaviruses. At the turn of the 21st century, GALLAGHER (2001) addressed coronavirus Spike proteins in viral entry and pathogenesis. Shortly thereafter, XIAO (2003) described the cloning, expression, and characterization of the full-length and various soluble fragments of the SARS-CoV (Tor2 isolate) S glycoprotein. In parallel, BOSCH (2003) performed structural and functional characterization of the fusion core complex, and showed the coronavirus Spike protein is a class I virus fusion protein. LIU (2004) then described the interaction between heptad repeat 1 and 2 regions in Spike protein of SARS-associated coronavirus, with implications for virus fusogenic mechanism and identification of fusion inhibitors. He showed that one peptide derived from the HR2 region, CP-1, might be modifiable to increase its anti-SARS-CoV activity and could be further developed as an antiviral agent for treatment or prophylaxis of SARS-CoV infection. Simultaneously, BABCOCK (2004) showed that amino acids 270 to 510 of the SARS coronavirus Spike protein are required for interaction with receptor.

YANG (2004) then demonstrated that pH-dependent entry of SARS coronavirus is mediated by the Spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. Both cell-mediated infection and direct infection were inhibited by anti-S antisera, indicating that strategies directed toward this gene product were likely to confer a therapeutic benefit for antiviral drugs or the development of a SARS vaccine. SIMMONS (2004) characterized SARS-associated coronavirus (SARS-CoV) Spike glycoprotein-mediated viral entry. He described important tools that could be used to study SARS-CoV S glycoprotein structure and function, including approaches that can be used to identify inhibitors of the entry of SARS-CoV into target cells. BOSCH (2004) demonstrated SARS coronavirus (SARS-CoV) infection inhibition using Spike protein heptad repeat-derived peptides. Two years later, KOPECKY-BROMBERG (2006) showed that 7a protein of SARS coronavirus inhibited cellular protein synthesis and activated p38 mitogen-activated protein kinase. BOSCH (2008) then demonstrated that cathepsin L activates the membrane fusion function of the spike protein, and that cleavage was

mapped to the same region where (in coronaviruses carrying furin-activated spikes) the receptor binding subunit of the protein is separated from the membrane-anchored fusion subunit.

Immunization

mice; immunized; neutralizing; vaccine; antibodies; immune; epitopes; antibody; recombinant; immunogenicity; immunity; humoral; epitope; BALB/c; immunization; neutralization; vaccinated; IgG; antigen; vaccines; SARS-CoV; serum; sera; T; ELISA; CTL; RBD; receptor-binding; Spike; S; protein

There is a long history of important papers (as reflected in the present retrieved SARS database) underlying modern advances in immunization against infectious diseases. KARBER (1931) described a method for calculating the tissue culture infectious dose 50% endpoint titers (log10TCID50). REED (1938) described a simple method for determining the LD50 value in experimental biology. Over four decades later, WEISS (1981) compared antibody-mediated enhancement of disease in feline infectious peritonitis with dengue hemorrhagic fever. COLLINS (1982) showed how monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell--cell fusion.

FUERST (1986) described eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. VENNEMA (1990) described early death after feline infectious peritonitis virus challenge because of recombinant vaccinia virus immunization. WILLIAMS (1991) showed that the MHV receptor was the first member of the human carcino-embryonic antigen family of glycoproteins to be identified as a virus receptor. KUBO (1994) demonstrated localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus Spike protein, indicating indicated that a domain composed of 330 aa at the N terminus of the S1 protein is responsible for binding to the MHV-specific receptor.

A decade later, YANG (2004) showed that a DNA vaccine encoding the Spike (S) glycoprotein of the SARS-CoV induced T cell and neutralizing antibody responses, as well as protective immunity, in a mouse model. SUBBARAO (2004) demonstrated that antibodies, acting alone, can prevent replication of the SARS coronavirus in the lung, a promising observation for the development of vaccines, immunotherapy, and immunoprophylaxis regimens. BISHT (2004) showed that

SARS coronavirus Spike protein expressed by attenuated vaccinia virus protectively immunized mice. SUI (2004) described a potent neutralization of SARS coronavirus by a human mAb to S1 protein that blocked receptor association, suggesting that the 80R human monoclonal antibody may be a useful viral entry inhibitor for the emergency prophylaxis and treatment of SARS, and that the ACE2-binding site of S1 could be an attractive target for subunit vaccine and drug development.

BUKREYEV (2004) demonstrated mucosal immunization of African green monkeys (Cercopithecus aethiops) with an attenuated parainfluenza virus expressing the SARS coronavirus Spike protein for the prevention of SARS. LI (2005) presented the crystal structure of SARS coronavirus Spike receptor-binding domain (RBD) complexed with receptor, and this structure of the RBD suggested ways to make truncated disulfide-stabilized RBD variants for use in the design of coronavirus vaccines. SPRUTH (2006) then showed that a double-inactivated whole virus candidate SARS coronavirus vaccine stimulated neutralising and protective antibody responses. ROBERTS (2006) demonstrated that therapy with a SARS-associated coronavirus-neutralizing human monoclonal antibody reduced disease severity and viral burden in golden Syrian hamsters. ROBERTS (2007) further showed that a mouse-adapted SARS-coronavirus caused disease and mortality in BALB/c mice, and proposed this virus could be of value as a stringent challenge in evaluation of the efficacy of vaccines and antivirals. ROCKX (2007) described the construction of a panel of isogenic SARS coronavirus (SARS-CoV) strains bearing variant spike glycoproteins that are representative of zoonotic strains found in palm civets and raccoon dogs, as well as isolates spanning the early, middle, and late phases of the SARS-CoV epidemic. All viruses replicated efficiently, but none produced clinical disease or death in young animals. In contrast, severe clinical disease, diffuse alveolar damage, hyaline membrane formation, alveolitis, and death were noted in 12-month-old mice inoculated with the palm civet HC/SZ/61/03 strain or early-human-phase GZ02 variants but not with related middle- and late-phase epidemic or raccoon dog strains. ROBERTS (2008) summarized findings of SARS-CoV infections in several animal models each of which support viral replication in lungs, and concluded that vaccines and monoclonal antibodies specific to SARS-CoV spike protein are highly efficacious in prophylaxis.

Structural Proteins

membrane; protein; proteins; expressed; nucleocapsid; N; localization; endoplasmic; structural; amino; reading; reticulum; encodes; Western

STURMAN (1980) demonstrated isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid, resulting in a model for the arrangement of the three major structural proteins in the coronavirus A59 virion in relation to the viral envelope and RNA. BRIERLEY (1987), in the first non-retroviral example of ribosomal frame-shifting in higher eukaryotes, described an efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. SPAAN (1988) described progress in coronavirology, based on accumulated sequence data and focused on the viral nucleic acids and proteins and their function. JONES (1991) described improved methods for building protein models in electron density maps and the location of errors in these models. LASKOWSKI (1993) described PROCHECK - a program to check the stereochemical quality of protein structures.

VENNEMA (1996) demonstrated nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes, showing that the nucleocapsid-independent formation of apparently bona fide viral envelopes represented a novel mode of virus assembly. CORSE (2000) showed that infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles.

BUCHHOLZ (2004) investigated the contributions of the structural proteins of SARS coronavirus (CoV) to protective immunity by expressing them individually and in combinations from a recombinant parainfluenza virus (PIV) type 3 vector called BHPIV3. He identified S among the structural proteins as the only significant SARS-CoV neutralization antigen and protective antigen and showed that a single mucosal immunization is highly protective in an experimental animal that supports efficient replication of SARS-CoV. SURJIT (2006) showed that the nucleocapsid protein of SARS-coronavirus inhibited the activity of cyclin-cyclin-dependent kinase complex and blocked S phase progression in mammalian cells. KOPECKY-BROMBERG (2007) demonstrated that SARS coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins functioned as interferon antagonists.

Intracellular Replication

cells; expression; replication; cell; inhibited; activation; IFN; interferon; phosphorylation; mRNA; signaling; Vero; apoptosis; induction; inhibition; innate

SIDDELL (1983) summarized progress made in understanding the basic molecular features of coronavirus replication. Shortly thereafter, TOOZE (1984) described replication of coronavirus MHV-A59 in sac-cells, and determination of the first site of budding of progeny virions. BARIC (1988) described interactions between coronavirus nucleocapsid protein and viral RNAs. The implications for viral transcription were that the MHV N protein was associated with MHV-specific RNAs and RNA intermediates and may play an important functional role during MHV transcription and replication. SAMBROOK (1989), in his highly cited molecular cloning manual, showing why cloning works. A decade later, LAI (1997) presented a comprehensive exposition on the molecular biology of coronaviruses. SAWICKI (1998) proposed a new model for coronavirus transcription, which predicted that subgenome-length negative strands would be derived directly by transcription using the genome RNA as a template, and the subgenome-length templates would contain the common 5' leader sequence and serve as templates for the production of subgenomic mRNAs. LAI (2001) addressed the Coronaviridae viruses and their replication. GOSERT (2002) concluded that RNA replication of mouse hepatitis virus takes place at double-membrane vesicles, whereby double membrane vesicles carry the MHV RNA replication complex and are the site of MHV RNA synthesis.

Based on assessment of the antiviral potential of recombinant interferons against two clinical isolates of SARS-CoV-FFM-1, CINATL (2003) concluded interferon beta could be useful alone or in combination with other antiviral drugs for the treatment of SARS. HAAGMANS (2004) showed that pegylated interferon-alpha protected type 1 pneumocytes against SARS, and could be considered a candidate drug for SARS therapy coronavirus infection in macaques. SPIEGEL et al (2005) demonstrated that SARS-CoV escapes IFN-mediated growth inhibition by preventing the induction of IFN-beta. Specifically, their data suggested that nuclear transport of IRF-3 is an immediate-early reaction to virus infection and may precede its hyperphosphorylation, homodimer formation, and binding to CBP. In order to escape activation of the IFN system, SARS-CoV appeared to block a step after the early nuclear transport of IRF-3. KAMITANI (2006) showed that SARS coronavirus nsp1 protein suppressed host gene expression by promoting host mRNA degradation. STERTZ (2007) analyzed the replication and budding sites of severe acute respiratory syndrome coronavirus (SARS-CoV) at early time points of infection, and found that SARS-CoV establishes replication complexes at ER-derived membranes. In parallel, FRIEMAN (2007) demonstrated that SARS-COV ORF6 protein is localized to the endoplasmic reticulum (ER)/Golgi membrane in infected cells, where it binds to and disrupts nuclear import complex formation by tethering karyopherin alpha 2 and karyopherin beta 1 to the membrane. He further

showed that retention of import factors at the ER/Golgi membrane leads to a loss of STAT1 transport into the nucleus in response to interferon signaling, thus blocking the expression of STAT1-activated genes that establish an antiviral state. FRIEMAN (2008) further concluded that interaction between specific viral genes and the innate immune system functions as a key determinant in regulating virulence and disease outcomes, and that the SARS-CoV uses specific strategies to evade and antagonize the sensing and signaling arms of the interferon pathway. NARAYANANJ (2008) reinforced this conclusion by demonstrating that SARS-CoV nsp1 suppressed host innate immune functions, including type I IFN expression, in infected cells and suggested that SARS-CoV nsp1 most probably plays a critical role in SARS-CoV virulence.

Protease Inhibitors

protease; substrate; catalytic; inhibitors; 3CL(pro); activity; active; drug; proteinase; dimer; crystal; cleavage; structure; M-pro; enzymatic; pocket; inhibitor; target; residues; proteases; structures; site; enzyme; attractive; compounds; polyproteins; substrates; polyprotein; docking; monomer; energy; 3CLpro; binding

Four decades ago, SCHECHTER (1967) discussed the size of the active site in proteases. In a widely cited paper, LAEMMLI (1970) showed that many hitherto unknown proteins have been found in bacteriophage T4 and some of these have been identified with specific gene products. Four major components of the head are cleaved during the process of assembly, apparently after the precursor proteins have assembled into some large intermediate structure. Two decades later, FRANA (1985) described proteolytic cleavage of the E2 glycoprotein of murine coronavirus, and focused on host-dependent differences in proteolytic cleavage and cell fusion. At the same time, STURMAN (1985) described similar proteolytic cleavage of the E2 glycoprotein of murine coronavirus, but focused on activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments.

A decade later, OTWINOWSKI (1997) presented the HKL macromolecular crystallography package for processing of x-ray diffraction data collected in oscillation mode. At the same time, ZIEBUHR (1997) described the biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase, showing that the residues Cys-3109 and His-3006 are indispensable for catalytic activity, and that His-3127 is involved in substrate recognition. They

confirm the requirement of the carboxyl-terminal extension found in coronavirus 3C-like proteinases for enzymatic activity. BRUNGER (1998) described a new software suite, called Crystallography & NMR System (CNS), developed for macromolecular structure determination by X-ray crystallography or solution nuclear magnetic resonance (NMR) spectroscopy.

Three years later, ZIEBUHR (2000) discussed virus-encoded proteinases and proteolytic processing in the Nidovirales. ANAND (2002) showed that structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. At the same time, HEGYI (2002) demonstrated conservation of substrate specificities among coronavirus main proteases by showing that the differential cleavage kinetics of sites within pp1a/pp1ab are a conserved feature of coronavirus main proteases and lead to prediction of similar processing kinetics for the replicase polyproteins of all coronaviruses.

Shortly thereafter, ANAND (2003) determined crystal structures for human coronavirus (strain 229E) Mpro and for an inhibitor complex of porcine coronavirus [transmissible gastroenteritis virus (TGEV)] Mpro. He constructed a homology model for SARS coronavirus (SARS-CoV) Mpro, and showed that coronavirus main proteinase (3CLpro) structure could serve as the basis for design of anti-SARS drugs. YANG (2003) then described the crystal structures of SARS virus main protease and its complex with an inhibitor. FAN (2004) examined biosynthesis, purification, and substrate specificity of SARS coronavirus 3C-like proteinase, providing a basic understanding of the enzyme catalysis and a full substrate specificity spectrum for SARS 3C-like proteinase, which are helpful for structural-based inhibitor design against SARS and other coronavirus.

RATIA (2006) described the 1.85-angstrom crystal structure of the catalytic core of SARS-CoV PLpro and showed that the overall architecture adopts a fold closely resembling that of known de-ubiquitinating enzymes. Key features distinguish PLpro from characterized de-ubiquitinating enzymes, including an intact zinc-binding motif, an unobstructed catalytically competent active site, and a ubiquitin-like N-terminal domain. More recently, XUE (2007) showed production of authentic SARS-CoV M-pro with enhanced activity, and application as a novel tag-cleavage endopeptidase for protein overproduction. XUE (2008) showed further that a Michael acceptor inhibitor (named N3) was co-crystallized with IBV M-Pro and was found to demonstrate in vitro inactivation of IBV M-Pro and potent antiviral activity against IBV in chicken embryos, thereby providing a feasible animal model for designing wide-spectrum inhibitors against CoV-associated

diseases. The structure-based optimization of N3 yielded two more efficacious lead compounds, N27 and H16, with potent inhibition against SARS-CoV M-Pro.

Genome Sequence

genome; coronaviruses; phylogenetic; sequence; viruses; sequences; genomes; RNA; virus; species; genetic; gene; mutations; reading

Over thirty years ago, SIDWELL (1972) described broad-spectrum antiviral activity of Virazole (1-beta-D-ribofuranosyl-1,2,4-triazole-3-carboxamide), a synthetic nucleoside active in tissue culture against at least 16 DNA and RNA viruses, including some influenza viruses. Almost two decades later, GORBALENYA (1989) predicted putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis for the coronavirus genome. At the same time, BRIERLEY (1989) showed by creation of complementary nucleotide changes that the RNA downstream of the sequence UUUAAAC (which is the likely site of ribosomal slippage), folds into a tertiary structure termed a pseudoknot, the formation of which is essential for efficient frameshifting. A decade later, SANCHEZ (1999) showed that targeted recombination demonstrates that the Spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. KUO (2000) described the retargeting of coronavirus by substitution of the Spike glycoprotein ectodomain, and subsequent crossing of the host cell species barrier. Specifically, targeted RNA recombination was used to construct a mutant of the coronavirus mouse hepatitis virus (MHV) in which the ectodomain of the Spike glycoprotein (S) was replaced with the highly divergent ectodomain of the S protein of feline infectious peritonitis virus. The resulting chimeric virus, designated fMHV, acquired the ability to infect feline cells and simultaneously lost the ability to infect murine cells in tissue culture. DEHAAN (2002) then showed that the group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host.

The following year, SNIJDER (2003) showed unique and conserved features of genome and proteome of SARS-coronavirus (an early split-off from the coronavirus group 2 lineage), using published genome sequences. RUAN (2003) performed comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and identified common mutations associated with putative origins of infection. THIEL (2003) determined the sequence of SARS coronavirus

(SARS-CoV), isolate Frankfurt 1, and characterized key RNA elements and protein functions involved in viral genome expression, including important regulatory mechanisms, such as the (discontinuous) synthesis of eight subgenomic mRNAs, ribosomal frameshifting and post-translational proteolytic processing, were addressed. GORBALENYA (2006) conducted a review focusing on the monophyletic group of animal RNA viruses united in the order Nidovirales, including the distantly related coronaviruses, toroviruses, and roniviruses, which possess the largest known RNA genomes.

Angiotensin-Converting Enzyme

ACE2; angiotensin-converting; receptor; angiotensin; ACE; renin-angiotensin; converting; homologue; heart; enzyme

YEAGER (1992) demonstrated that human aminopeptidase N, a cell-surface metalloprotease on intestinal, lung and kidney epithelial cells, is a receptor for human coronavirus strain HCV-229E, but not for HCV-OC43. Along the same lines, DELMAS (1992) demonstrated that aminopeptidase N, an ectoenzyme abundantly expressed at the apical membrane of the enterocytes, serves as a receptor for the entero-pathogenic coronavirus TGEV. A decade later, TIPNIS (2000) identified a novel human zinc metalloprotease that had considerable homology to human angiotensin-converting enzyme. It demonstrated cloning and functional expression as a captopril-insensitive carboxypeptidase. DONOGHUE (2000) showed that a novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. CRACKOWER (2002) demonstrated that angiotensin-converting enzyme 2 is an essential regulator of heart function.

LI (2003) demonstrated that angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. WONG (2004) showed that a 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2, thereby identifying an independently folded receptor-binding domain of the SARS-CoV S protein. JEFFERS (2004) showed that CD209L (L-SIGN) is a receptor for SARS, and that large S glycoprotein of SARS-CoV may use both ACE2 and CD209L in virus infection and pathogenesis. HAMMING (2004) showed the tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. ACE2 is abundantly present in humans in the epithelia of the lung and small intestine, which might provide possible routes of entry for the SARS-CoV. LI (2005) demonstrated receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. HUANG (2006) showed that SARS

coronavirus, but not human coronavirus NL63, utilized the enzymatic activity of the cysteine protease cathepsin L to infect ACE2-expressing cells, indicating that two coronaviruses that utilize a common receptor nonetheless enter cells through distinct mechanisms. Very recently, MCCRAY (2007) showed lethal infection of K18-hACE2 mice infected with SARS coronavirus. At the same time, TSENG (2007) demonstrated SARS coronavirus infection of mice transgenic for the human Angiotensin-converting enzyme 2 virus receptor.

DISCUSSION AND CONCLUSIONS

The CAB approach has been applied to the SARS literature, and found to be thorough by researchers who have published on the topic since the pandemic. For the first time since the CAB approach was developed, the references were assigned to categories generated by clustering of the total SARS core literature. The combination of the clustering and the CAB approach provided a systematic method for conducting seminal literature surveys.

The seminal literature had three main components: epidemiology, clinical, and laboratory. The epidemiology roots of statistical modeling and cross-species transmission blended to show how the SARS transmission occurred, and how it was finally controlled. The clinical roots showed the seminal approaches in virus detection, especially coronaviruses, and how they matured into the detection of the SARS coronavirus. There was little of note on treatments, understandable because of recent reviews that concluded none of the treatments worked. The laboratory roots traced research into determination of viral structures, viral entry, viral replication, viral inhibition, with emphasis on vaccine development. With the advent of SARS, expanded efforts were made in each of these areas, vaccine development being the ultimate target of many of these studies.

A more detailed analysis of recent papers showed that those who succumbed to SARS tended to exhibit immune system deficiencies upon presentation. Missing in the seminal literature were myriad approaches for strengthening the immune system without, or in addition to, vaccines. However, it should be remembered the seminal literature is a small fraction of the total literature. Examination of the total biomedical literature shows there are many potential approaches to strengthening the immune system, and hopefully providing protection against similar viral pandemics in the future. A follow-on study will identify these potential immune-strengthening approaches.

REFERENCES

Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. EMBO Journal. 21:13. 3213-3224. 2002.

Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CL(pro)) structure: Basis for design of anti-SARS drugs. Science. 300:5626. 1763-1767. 2003.

Anderson RM, and May RM. Infectious Disease of Humans: Dynamics and Control. Oxford, U.K.: Oxford University Press. 1991.

Babcock GJ, Esshaki DJ, Thomas WD, Ambrosino DM. Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor. Journal of Virology. 78:9. 4552-4560. 2004.

Bailey NTJ. The mathematical theory of infectious diseases and its applications. 2nd ed. London : Griffin. 1975.

Baric RS, Nelson GW, Fleming JO, Deans RJ, Keck JG, Casteel N and Stohlman SA. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. J Virol. 62(11). 4280-4287. 1988.

Bisht H, Roberts A, Vogel L, Bukreyev A, Collins PL, Murphy BR, Subbarao K, Moss B. Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice. Proceedings of the National Academy of Sciences of the United States of America. 101:17. 6641-6646. 2004.

Booth CM, Matukas LM, Tomlinson GA, Rachlis AR, Rose DB, Dwosh HA, Walmsley SL, Mazzulli T, Avendano M, Derkach P, Ephtimios IE, Kitai I, Mederski BD, Shadowitz SB, Gold WL, Hawryluck LA, Rea E, Chenkin JS, Cescon DW, Poutanen SM, Detsky AS. Clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area. JAMA-Journal of the American Medical Association. 289:21. 2801-2809. 2003.

Bosch BJ, Bartelink W, Rottier PJM. Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather

than adjacent to the fusion peptide. Journal of Virology. 82:17. 8887-8890. 2008.

Bosch BJ, Martina BEE, van der Zee R, Lepault J, Haijema BJ, Versluis C, Heck AJR, de Groot R, Osterhaus ADME, Rottier PJM. Severe Acute Respiratory Syndrome coroavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proceedings of the National Academy of Sciences of the United States of America. 101:22. 8455-8460. 2004.

Bosch BJ, van der Zee R, de Haan CAM, Rottier PJM. The coronavirus spike protein is a class I virus fusion protein: Structural and functional characterization of the fusion core complex. Journal of Virology. 77:16. 8801-8811. 2003.

Brierley I, Boursnell ME, Binns MM, Bilimoria B, Blok VC, Brown TD, Inglis SC. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. The EMBO journal. 6(12), 3779-3785. 1987.

Brierley I, Digard P, Inglis SC. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell. 57(4). 537-47. 1989.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallographica Section D-Biological Crystallography 54. 905-921. 1998.

Buchholz UJ, Bukreyev A, Yang LJ, Lamirande EW, Murphy BR, Subbarao K, Collins PL. Contributions of the structural proteins of Severe Acute Respiratory Syndrome coronavirus to protective immunity. Proceedings of the National Academy of Sciences of the United States of America. 101:26. 9804-9809. 2004.

Bukreyev A, Lamirande EW, Buchholz UJ, Vogel LN, Elkins WR, St Claire M, Murphy BR, Subbarao K, Collins PL. Mucosal immunisation of African green monkeys (Cercopithecus aethiops) with an attenuated parainfluenza virus expressing the SARS coronavirus spike protein for the prevention of SARS. Lancet. 363:9427. 2122-2127. 2004.

Calne DB, Calne R. 1992. Citation of original research. Lancet. 340 (8813): 244-244.

Chan VSF, Chan KYK, Chen YX, Poon LLM, Cheung ANY, Zheng BJ, Chan KH, Mak W, Ngan HYS, Xu XN, Screaton G, Tam PKH, Austyn JM, Chan LC, Yip SP, Peiris M, Khoo US, Lin CLS. Homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection. Nature Genetics. 38:1. 38-46. 2006.

Cinatl J, Morgenstern B, Bauer G, Chandra P, Rabenau H, Doerr HW. Treatment of SARS with human interferons. Lancet. 362:9380. 293-294. 2003.

Collins AR, Knobler RL, Powell H. and Buchmeier, MJ. Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell-cell fusion. Virology. 119. 358-371. 1982.

Corse E, Machamer CE. Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. Journal of Virology. 74:9. 4319-4326. 2000.

Crackower MA, Sarao R, Oudit GY, Yagil C, Kozieradzki I, Scanga SE, Oliveira-dos-Santos AJ, da Costa J, Zhang LY, Pei Y, Scholey J, Ferrario CM, Manoukian AS, Chappell MC, Backx PH, Yagil Y, Penninger JM. Angiotensin-converting enzyme 2 is an essential regulator of heart function. Nature. 417:6891. 822-828. 2002.

de Groot RJ, Luytjes W, Horzinek MC, van der Zeijst BA, Spaan WJ, Lenstra JA. Evidence for a coiled-coil structure in the spike proteins of coronaviruses. Journal of molecular biology. 196(4). 963-966. 1987.

de Haan CAM, Masters PS, Shen XL, Weiss S, Rottier PJM. The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. Virology. 296:1. 177-189. 2002.

Della Mea V, Mizzaro S. 2004. Measuring retrieval effectiveness: A new proposal and a first experimental validation. Journal of the American Society for Information Science and Technology. 55 (6): 530-543.

Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjöström H, Norén O, Laude H. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. Nature. 357(6377). 417-420. 1992.

Donnelly CA, Ghani AC, Leung GM, Hedley AJ, Fraser C, Riley S, Abu-Raddad LJ, Ho LM, Thach TQ, Chau P, Chan KP, Lam TH, Tse LY, Tsang T, Liu SH, Kong JHB, Lau EMC, Ferguson NM, Anderson RM. Epidemiological determinants of spread of causal agent of Severe Acute Respiratory Syndrome in Hong Kong. Lancet. 361:9371. 1761-1766. 2003.

Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R, Breitbart RE, Acton S. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circulation Research. 87:5. E1-E9. 2000.

Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RAM, Berger A, Burguiere AM, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra JC, Muller S, Rickerts V, Sturmer M, Vieth S, Klenk HD, Osterhaus ADME, Schmitz H, Doerr HW. Identification of a novel coronavirus in patients with Severe Acute Respiratory Syndrome. New England Journal of Medicine. 348:20. 1967-1976. 2003.

Fan KQ, Wei P, Feng Q, Chen SD, Huang CK, Ma L, Lai B, Pei JF, Liu Y, Chen JG, Lai LH. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. Journal of Biological Chemistry. 279:3. 1637-1642. 2004.

Feller W. An Introduction to Probability Theory and Its Applications, Vol. I, third ed. Wiley, New York. 1968.

Fouchier RAM, Kuiken T, Schutten M, van Amerongen G, van Doornum J, van den Hoogen BG, Peiris M, Lim W, Stohr K, Osterhaus ADME. Aetiology - Koch's postulates fulfilled for SARS virus. Nature. 423:6937. 240-240. 2003.

Frana MF, Behnke JN, Sturman LS, and Holmes KV. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: host-dependent differences in proteolytic cleavage and cell fusion. J. Virol. 56. 912-920. 1985.

Frieman M, Heise M, Baric R. SARS coronavirus and innate immunity. Virus Research. 133:1. 101-112. 2008.

Frieman M, Yount B, Heise M, Kopecky-Bromberg SA, Palese P, Baric RS. Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function

by sequestering nuclear import factors on the rough endoplasmic Reticulum/Golgi membrane. Journal of Virology. 81:18. 9812-9824. 2007.

Fuerst TR, Niles EG, Studier FW, and Moss B. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. PNAS. 83. 8122-8126. 1986.

Gallagher TM, Buchmeier MJ. Coronavirus spike proteins in viral entry and pathogenesis. Virology. 279:2. 371-374. 2001.

Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. Nidovirales: Evolving the largest RNA virus genome. Virus Research. 117:1. 17-37. 2006.

Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. Nucleic Acids Res. 17(12). 4847–4861. 1989.

Gosert R, Kanjanahaluethai A, Egger D, Bienz K, Baker SC. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. Journal of Virology. 76:8. 3697-3708. 2002.

Guan Y, Zheng BJ, He YQ, Liu XL, Zhuang ZX, Cheung CL, Luo SW, Li PH, Zhang LJ, Guan YJ, Butt KM, Wong KL, Chan KW, Lim W, Shortridge KF, Yuen KY, Peiris JSM, Poon LLM. Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. Science. 302:5643. 276-278. 2003.

H Vennema H, R J de Groot RJ, D A Harbour DA, M Dalderup M, T Gruffydd-Jones T, M C Horzinek MC, W J Spaan WJ. Early death after feline infectious peritonitis virus challenge due to recombinant vaccinia virus immunization. Journal of virology. 64(3). 1407-1409. 1990.

Haagmans BL, Kuiken T, Martina BE, Fouchier RAM, Rimmelzwaan GF, van Amerongen G, van Riel D, de Jong T, Itamura S, Chan KH, Tashiro M, Osterhaus ADME. Pegylated interferon-alpha protects type 1 pneumocytes against SARS coronavirus infection in macaques. Nature Medicine. 10:3. 290-293. 2004.

Hamming I, Timens W, Bulthuis MLC, Lely AT, Navis GJ, van Goor H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first

step in understanding SARS pathogenesis. Journal of Pathology. 203:2. 631-637. 2004.

Hamre D and Procknow JJ. A new virus isolated from the human respiratory tract. Experimental Biology and Medicine. 121. 190-193. 1966.

Harris JW. Influenza occurring in pregnant women: A statistical study of thirteen hundred and fifty cases. J Am Med Assoc. 72(14). 978-980. 1919.

Harter SP, Hert CA. 1997. Evaluation of information retrieval systems: Approaches, issues, and methods. Annual Review of Information Science and Technology. 32: 3-94.

He JF, Peng GW, Min J, Yu DW, Liang WJ, Zhang SY, Xu RH, Zheng HY, Wu XW, Xu J, Wang ZH, Fang L, Zhang X, Li H, Yan XG, Lu JH, Hu ZH, Huang JC, Wan ZY, Hou JL, Lin JY, Song HD, Wang SY, Zhou XJ, Zhang GW, Gu BW, Zheng HJ, Zhang XL, He M, Zheng K, Wang BF, Fu G, Wang XN, Chen SJ, Chen Z, Hao P, Tang H, Ren SX, Zhong Y, Guo ZM, Liu Q, Miao YG, Kong XY, He WZ, Li YX, Wu CI, Zhao GP, Chiu RWK, Chim SSC, Tong YK, Chan PKS, Tam JS, Lo YMD. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science. 303:5664. 1666-1669. 2004.

Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. Journal of General Virology. 83. 595-599. 2002.

Horowitz M, Wilner N, and Alvarez W. Impact of Event Scale: a measure of subjective stress. Psychosom Med. 41. 209-218. 1979.

Huang IC, Bosch BJ, Li F, Li WH, Lee KH, Ghiran S, Vasilieva N, Dermody TS, Harrison SC, Dormitzer PR, Farzan M, Rottier PJM, Choe H. SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. Journal of Biological Chemistry. 281:6. 3198-3203. 2006.

Jeffers SA, Tusell SM, Gillim-Ross L, Hemmila EM, Achenbach JE, Babcock GJ, Thomas WD, Thackray LB, Young MD, Mason RJ, Ambrosino DM, Wentworth DE, DeMartini JC, Holmes KV. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. Proceedings of the National Academy of Sciences of the United States of America. 101:44. 15748-15753. 2004.

Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A 47. 110-119. 1991.

Kagolovsky Y, Moehr JR. 2004. A new look at information retrieval evaluation: Proposal for solutions. Journal of Medical Systems. 28 (1): 103-116.

Kamitani W, Narayanan K, Huang C, Lokugamage K. Ikegami T, Ito N, Kubo H, Makino S. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proceedings of the National Academy of Sciences of the United States of America. 103 (34). 12885-12890. 2006.

Kantor PB. Information-retrieval techniques. 1994. Annual Review of Information Science and Technology. 29: 53-90.

Karber G. Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. Arch. Exp. Path. Pharmak. 162. 480-483. 1931.

Kermack WO and McKendrick AG. A contribution to the mathematical theory of epidemics. Proc. Roy. Soc. Lond. A 115, 700-721. 1927.

Khan MS, Khor S. 2004. Enhanced Web document retrieval using automatic query expansion. Journal of the American Society for Information Science and Technology. 55 (1): 29-40.

Kopecky-Bromberg SA, Martinez-Sobrido L, Frieman M, Baric RA, Palese P. Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. Journal of Virology. 81:2. 548-557. 2007.

Kopecky-Bromberg SA, Martinez-Sobrido L, Palese P. 7a protein of severe acute respiratory syndrome coronavirus inhibits cellular protein synthesis and activates p38 mitogen-activated protein kinase. Journal of Virology. 80:2. 785-793. 2006.

Kostoff, R. N., Eberhart, H. J., and Toothman, D. R. 1998. Database Tomography for Technical Intelligence: A Roadmap of the Near-Earth Space Science and Technology Literature. Information Processing and Management. 34:1.

Kostoff, R. N., Green, K. A., Toothman, D. R., and Humenik, J. 2000. Database tomography applied to an aircraft science and technology investment strategy. Journal of Aircraft. 37:4. 727-730.

Kostoff, R. N., Shlesinger, M., and Tshiteya, R. 2004. Nonlinear dynamics roadmaps using bibliometrics and database tomography. International Journal of Bifurcation and Chaos. 14:1. 61-92.

Kostoff, R.N., and Shlesinger, M. F. 2005. CAB-citation-assisted background. Scientometrics. 62:2. 199-212.

Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong SX, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling AE, Humphrey CD, Shieh WJ, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang JY, Cox N, Hughes JM, LeDuc JW, Bellini WJ, Anderson LJ. A novel coronavirus associated with severe acute respiratory syndrome. New England Journal of Medicine. 348:20. 1953-1966. 2003.

Kubo H, Yamada YK, and Taguchi F. Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. J. Virol. 68. 5403-5410. 1994.

Kuiken T, Fouchier RAM, Schutten M, Rimmelzwaan GF, van Amerongen G, van Riel D, Laman JD, de Jong T, van Doornum G, Lim W, Ling AE, Chan PKS, Tam JS, Zambon MC, Gopal R, Drosten C, van der Werf S, Escriou N, Manuguerra JC, Stohr K, Peiris JSM, Osterhaus ADME. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. Lancet. 362:9380. 263-270. 2003.

Kuo LL, Godeke GJ, Raamsman MJB, Masters PS, Rottier PJM. Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: Crossing the host cell species barrier. Journal of Virology. 74:3. 1393-1406. 2000.

Laemmli U. Cleavage of structural proteins during the assembly of bacteriophage T4. Nature. 227. 680−685. 1970.

Lai MMC and Holmes KV. Coronaviruses. In: Fields Virology, (D.M. Knipe and P.M. Howley, eds), pp 1163-1185. Lippincott, Williams & Wilkins, Philadelphia, Pa. 2001.

Lai MMC, Cavanagh D. The molecular biology of coronaviruses. Advances in Virus Research. Book Series: Advances in Virus Research. 48. 1-100. 1997.

Laskowski RA, McArthur MW, Moss DS, and Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26. 283-291. 1993.

Lau SKP, Woo PCY, Li KSM, Huang Y, Tsoi HW, Wong BHL, Wong SSY, Leung SY, Chan KH, Yuen KY. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proceedings of the National Academy of Sciences of the United States of America. 102:39. 14040-14045. 2005.

Lee N, Hui D, Wu A, Chan P, Cameron P, Joynt GM, Ahuja A, Yung MY, Leung CB, To KF, Lui SF, Szeto CC, Chung S, Sung JJY. A major outbreak of severe acute respiratory syndrome in Hong Kong. New England Journal of Medicine. 348:20. 1986-1994. 2003.

Li F, Li WH, Farzan M, Harrison SC. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science. 309:5742. 1864-1868. 2005.

Li WD, Shi ZL, Yu M, Ren WZ, Smith C, Epstein JH, Wang HZ, Crameri G, Hu ZH, Zhang HJ, Zhang JH, McEachern J, Field H, Daszak P, Eaton BT, Zhang SY, Wang LF. Bats are natural reservoirs of SARS-like coronaviruses. Science. 310:5748. 676-679. 2005.

Li WH, Moore MJ, Vasilieva N, Sui JH, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. 426:6965. 450-454. 2003.

Li WH, Wong SK, Li F, Kuhn JH, Huang IC, Choe H, Farzan M. Animal origins of the severe acute respiratory syndrome coronavirus: Insight from ACE2-S-protein interactions. Journal of Virology. 80:9. 4211-4219. 2006.

Li WH, Zhang CS, Sui JH, Kuhn JH, Moore MJ, Luo SW, Wong SK, Huang IC, Xu KM, Vasilieva N, Murakami A, He YQ, Marasco WA, Guan Y, Choe HY, Farzan M. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO Journal. 24:8. 1634-1643. 2005.

Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, Gopalakrishna G, Chew SK, Tan CC, Samore MH, Fisman D, Murray M. Transmission dynamics and control of severe acute respiratory syndrome. Science. 300:5627. 1966-1970. 2003.

Liu SW, Xiao GF, Chen YB, He YX, Niu JK, Escalante CR, Xiong HB, Farmar J, Debnath AK, Tien P, Jiang SB. Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors. Lancet. 363:9413. 938-947. 2004.

Liu, M.X . 1993. Progress In Documentation - The Complexities of Citation Practice – A Review of Citation Studies. Journal of Documentation. 49 (4): 370-408.

MacRoberts M, MacRoberts B. 1996. Problems of citation analysis. Scientometrics. 36(3): 435-444.

Macroberts, M.H, Macroberts B.R. 1989. Problems of Citation Analysis - A Critical-Review. Journal of the American Society for Information Science. 40 (5): 342-349.

MacRoberts, M.H, MacRoberts, B.R. 1997. Citation content analysis of a botany journal. Journal of the American Society for Information Science. 48 (3): 274-275.

Marra MA, Jones SJM, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YSN, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Coughlin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JE, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Krajden M, Petric M, Skowronski DM, Upton C, Roper RL. The genome sequence of the SARS-associated coronavirus. Science. 300:5624. 1399-1404. 2003.

Martina BEE, Haagmans BL, Kuiken T, Fouchier RAM, Rimmelzwaan GF, van Amerongen G, Peiris JSM, Lim W, Osterhaus ADME. SARS virus infection of cats and ferrets. Nature. 425:6961. 915-915. 2003.

McCray PB, Pewe L, Wohlford-Lenane C, Hickey M, Manzel L, Shi L, Netland J, Jia HP, Halabi C, Sigmund CD, Meyerholz DK, Kirby P, Look DC, Perlman S. Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. Journal of Virology. 81:2. 813-821. 2007.

McIntosh K, Chao RK, Krause HE, Wasil R, Mocega HE, Mufson MA. Coronavirus infection in acute lower respiratory tract disease of infants. J Infect Dis. 130. 502-507. 1974.

McIntosh K, Dees JH, Becker WB, Kapikian AZ, and Chanock RM. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. PNAS. 57. 933-940. 1967.

Mcintosh K, Kapikian AZ, Turner HC, Hartley JW, Parrott RH, and Chanock RM. Seroepidemiologic studies of coronavirus infection in adults and children. Am. J. Epidemiol. 91: 585-592.

Moravcsik MJ, Murugesan P. 1975. Some results on function and quality of citations. Social Studies of Science. 5 (1): 86-92.

Moser MR, Bender TR, Margolis HS, Noble GR, Kendal AP, and Ritter DG. An outbreak of influenza aboard a commercial airliner. Am. J. Epidemiol. 110. 1-6. 1979.

Narayananj K, Huang C, Lokugamage K, Kamitani W, Ikegami T, Tseng CTK, Makino S. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. Journal of Virology. 82:9. 4471-4479. 2008.

Nicholls JM, Poon LLM, Lee KC, Ng WF, Lai ST, Leung CY, Chu CM, Hui PK, Mak KL, Lim W, Yan KW, Chan KH, Tsang NC, Guan Y, Yuen KY, Peiris JSM. Lung pathology of fatal severe acute respiratory syndrome. Lancet. 361:9371. 1773-1778. 2003.

Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A. Book Series: Methods in Enzymology. 276. 307-326. 1997.

Peiris JSM, Chu CM, Cheng VCC, Chan KS, Hung IFN, Poon LLM, Law KI, Tang BSF, Hon TYW, Chan CS, Chan KH, Ng JSC, Zheng BJ, Ng WL, Lai RWM, Guan Y, Yuen KY. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. Lancet. 361:9371. 1767-1772. 2003.

Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, Lim W, Nicholls J, Yee WKS, Yan WW, Cheung MT, Cheng VCC, Chan KH, Tsang DNC, Yung RWH, Ng TK, Yuen KY. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet. 361:9366. 1319-1325. 2003.

Peiris JSM, Yuen KY, Osterhaus ADME, Stohr K. Current concepts: The severe acute respiratory syndrome. New England Journal of Medicine. 34:25. 2431-2441. 2003.

Poutanen SM, Low DE, Henry B, Finkelstein S, Rose D, Green K, Tellier R, Draker R, Adachi D, Ayers M, Chan AK, Skowronski DM, Salit I, Simor AE, Slutsky AS, Doyle PW, Krajden M, Petric M, Brunham RC, McGeer AJ. Identification of severe acute respiratory syndrome in Canada. New England Journal of Medicine. 348:20. 1995-2005. 2003.

Ratia K, Saikatendu KS, Santarsiero BD, Barretto N, Baker SC, Stevens RC, Mesecar AD. Severe acute respiratory syndrome coronavirus papain-like protease: Structure of a viral deubiquitinating enzyme. Proceedings of the National Academy of Sciences of the United States of America. 103:15. 5717-5722. 2006.

Reed LJ and Muench H. A simple method of estimating fifty per cent endpoints. Am. J. Epidemiol. 27. 493-497. 1938.

Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, Leung GM, Ho LM, Lam TH, Thach TQ, Chau P, Chan KP, Leung PY, Tsang T, Ho W, Lee KH, Lau EMC, Ferguson NM, Anderson RM. Transmission dynamics of the etiological agent of SARS in Hong Kong: Impact of public health interventions. Science. 300:5627. 1961-1966. 2003.

Roberts A, Deming D, Paddock CD, Cheng A, Yount B, Vogel L, Herman BD, Sheahan T, Heise M, Genrich GL, Zaki SR, Baric R, Subbarao K. A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. PLOS Pathogens. 3:1. 23-37. 2007.

Roberts A, Lamirande EW, Vogel L, Jackson JP, Paddock CD, Guarner J, Zaki SR, Sheahan T, Baric R, Subbarao K. Animal models and vaccines for SARS-CoV infection. Virus Research. 133:1. 20-32. 2008.

Roberts A, Thomas WD, Guarner J, Lamirande EW, Babcock GJ, Greenough TC, Vogel L, Hayes N, Sullivan JL, Zaki S, Subbarao K, Ambrosino DM. Therapy with a severe acute respiratory syndrome-associated coronavirus-neutralizing human monoclonal antibody reduces disease severity and viral burden in golden Syrian hamsters. Journal of Infectious Diseases. 193:5. 685-692. 2006.

Rockx B, Sheahan T, Donaldson E, Harkema J, Sims A, Heise M, Pickles R, Cameron M, Kelvin D, Baric R. Synthetic reconstruction of zoonotic and early human severe acute respiratory syndrome coronavirus isolates that produce fatal disease in aged mice. Journal of Virology. 81:14. 7410-7423. 2007.

Ross R. The prevention of Malaria, second ed. Murray, London. 1911.

Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong SX, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TCT, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus ADME, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science. 300:5624. 1394-1399. 2003.

Ruan YJ, Wei CL, Ee LA, Vega VB, Thoreau H, Yun STS, Chia JM, Ng P, Chiu KP, Lim L, Tao Z, Peng CK, Ean LOL, Lee NM, Sin LY, Ng LFP, Chee RE, Stanton LW, Long PM, Liu ET. Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. Lancet. 361:9371. 1779-1785. 2003.

Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: a Laboratory Manual. Second Edition. Cold Spring Harbor Laboratory Press. 1989.

Sanchez CM, Izeta A, Sanchez-Morgado JM, Alonso S, Sola I, Balasch M, Plana-Duran J, Enjuanes L. Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. Journal of Virology. 73:9. 7607-7618. 1999.

Sawicki SG, Sawicki DL. A new model for coronavirus transcription. Coronaviruses And Arteriviruses. Book Series: Advances in Experimental Medicine and Biology. 440. 215-219. 1998.

Schechter I. and Berger A. On the size of the active site of proteases. I. Papain. Biochem. Biophys. Res. Commun. 27. 157-162. 1967.

Seto WH, Tsang D, Yung RWH, Ching TY, Ng TK, Ho M, Ho LM, Peiris JSM. Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). Lancet. 361:9368. 1519-1520. 2003.

Shadish WR, Tolliver D, Gray M, Sengupta SK. 1995. Author judgments about works they cite - 3 studies from psychology journals. Social Studies of Science. 25 (3): 477-498.

Siddell S, Wege H, and Ter Meulen V. The biology of coronaviruses. J. Gen. Virol. 64. 761−776. 1983.

Sidwell RW, Huffman JH, Khare GP, Allen LB, Witkowski JTR, and Robins K. Broad-spectrum antiviral activity of Virazole: 1-f8- D-Ribofuranosyl- 1,2,4-triazole- 3-carboxamide. Science. 177:4050. 705-706. 1972.

Simmons G, Reeves JD, Rennekamp AJ, Amberg SM, Piefer AJ, Bates P. Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. Proceedings of the National Academy of Sciences of the United States of America. 101:12. 4240-4245. 2004.

Smith, A.J, Goodman, N.W. 1997. The hypertensive response to intubation. Do researchers acknowledge previous work? Canadian Journal of Anaesthesia-Journal Canadien D Anesthesie. 44 (1): 9-13.

Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LLM, Guan Y, Rozanov M, Spaan WJM, Gorbalenya AE. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. Journal of Molecular Biology. 331:5. 991-1004. 2003.

So LKY, Lau ACW, Yam LYC, Cheung TMT, Poon E, Yung RWH, Yuen KY. Development of a standard treatment protocol for severe acute respiratory syndrome. Lancet. 361:9369. 1615-1617. 2003.

Song HD, Tu CC, Zhang GW, Wang SY, Zheng K, Lei LC, Chen QX, Gao YW, Zhou HQ, Xiang H, Zheng HJ, Chern SWW, Cheng F, Pan CM, Xuan H, Chen SJ, Luo HM, Zhou DH, Liu YF, He JF, Qin PZ, Li LH, Ren YQ, Liang WJ, Yu YD, Anderson L, Wang M, Xu RH, Wu XW, Zheng HY, Chen JD, Liang GD, Gao Y, Liao M, Fang L, Jiang LY, Li H, Chen F, Di B, He LJ, Lin JY, Tong SX, Kong XG, Du L, Hao P, Tang H, Bernini A, Yu XJ, Spiga O, Guo ZM, Pan HY, He WZ, Manuguerra JC, Fontanet A, Danchin A, Niccolai N, Li YX, Wu CI, Zhao GP. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proceedings of the National Academy of Sciences of the United States of America. 102:7. 2430-2435. 2005.

Spaan W, Cavanagh D, and Horzinek MC. Coronaviruses: Structure and Genome Expression. J Gen Virol. 69. 2939-2952. 1988.

Spearman C. The method of 'right and wrong cases' ('constant stimuli') without Gauss' formulae. British Journal of Psychology. 2. 227-242. 1908.

Spiegel M, Pichlmair A, Martinez-Sobrido L, Cros J, Garcia-Sastre A, Haller O, Weber F. Inhibition of beta interferon induction by severe acute respiratory syndrome coronavirus suggests a two-step model for activation of interferon regulatory factor 3. Journal of Virology. 79:4. 2079-2086. 2005.

Spruth M, Kistner O, Savidis-Dacho H, Hitter E, Crowe B, Gerencer M, Bruhl P, Grillberger L, Reiter M, Tauer C, Mundt W, Barrett PN. A double-inactivated whole virus candidate SARS coronavirus vaccine. Vaccine. 24:5. 652-661. 2006.

Stertz S, Reichelt M, Spiegel M, Kuri T, Martinez-Sobrido L, Garcia-Sastre A, Weber F, Kochs G. The intracellular sites of early replication and budding of SARS-coronavirus. Virology. 361:2. 304-315. 2007.

Sturman LS, Holmes KV, and Behnke J. Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. J Virol. 33(1). 449-462. 1980.

Sturman LS, Ricard CS, and Holmes KV. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. J. Virol. 56. 904-911. 1985.

Subbarao K, McAuliffe J, Vogel L, Fahle G, Fischer S, Tatti K, Packard M, Shieh WJ, Zaki S, Murphy B. Prior infection and passive transfer of neutralizing antibody prevent replication of severe acute respiratory syndrome coronavirus in the respiratory tract of mice. Journal of Virology. 78:7. 3572-3577. 2004.

Sui JH, Li WH, Murakami A, Tamin A, Matthews LJ, Wong SK, Moore MJ, Tallarico ASC, Olurinde M, Choe H, Anderson LJ, Bellini WJ, Farzan M, Marasco WA. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. Proceedings of the National Academy of Sciences of the United States of America. 101:8. 2536-2541. 2004.

Surjit M, Liu BP, Chow VTK, Lal SK. The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin-cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. Journal of Biological Chemistry. 281:16. 10669-10681. 2006.

Tang XC, Zhang JX, Zhang SY, Wang P, Fan XH, Li LF, Li G, Dong BQ, Liu W, Cheung CL, Xu KM, Song WJ, Vijaykrishna D, Poon LLM, Peiris JSM, Smith GJD, Chen H, Guan Y. Prevalence and genetic diversity of coronaviruses in bats from China. Journal of Virology. 80:15. 7481-7490. 2006.

Thiel V, Ivanov KA, Putics A, Hertzig T, Schelle B, Bayer S, Weissbrich B, Snijder EJ, Rabenau H, Doerr HW, Gorbalenya AE, Ziebuhr J. Mechanisms and enzymes involved in SARS coronavirus genome expression. Journal of General Virology. 84. 2305-2315. 2003.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22). 4673-4680. 1994.

Tipnis SR, Hooper NM, Hyde R, Karran E, Christie G, Turner AJ. A human homolog of angiotensin-converting enzyme - Cloning and functional expression as

a captopril-insensitive carboxypeptidase. Journal of Biological Chemistry. 275:43. 33238-33243. 2000.

Tooze J, Tooze S, and Warren G. Replication of coronavirus MHV-A59 in sac- cells: determination of the first site of budding of progeny virions. Eur. J. Cell Biol. 33. 281−293. 1984.

Tsang KW, Ho PL, Ooi GC, Yee WK, Wang T, Chan-Yeung M, Lam WK, Seto WH, Yam LY, Cheung TM, Wong PC, Lam B, Ip MS, Chan J, Yuen KY, Lai KN. A cluster of cases of severe acute respiratory syndrome in Hong Kong. New England Journal of Medicine. 348:20. 1977-1985. 2003.

Tseng CTK, Huang C, Newman P, Wang N, Narayanan K, Watts DM, Makino S, Packard MM, Zaki SR, Chan TS, Peters CJ. Severe acute respiratory syndrome coronavirus infection of mice transgenic for the human angiotensin-converting enzyme 2 virus receptor. Journal of Virology. 81:3. 1162-1173. 2007.

Tyrrell DAJ and Bynoe ML. Cultivation of a novel type of common-cold virus in organ cultures. Br Med J. 1. 1467-1470. 1965.

van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJM, Wolthers KC, Wertheim-van Dillen PME, Kaandorp J, Spaargaren J, Berkhout B. Identification of a new human coronavirus. Nature Medicine. 10:4. 368-373. 2004.

Vennema H, Godeke GJ, Rossen JW, Voorhout WF, Horzinek MC, Opstelten DJ, Rottier PJ. Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. EMBO J. 15(8). 2020–2028. 1996.

Vijaykrishna D, Smith GJD, Zhang JX, Peiris JSM, Chen H, Guan Y. Evolutionary insights into the ecology of coronaviruses. Journal of Virology. 81:8. 4012-4020. 2007.

Weiss RO and Scott FW. Antibody mediated enhancement of disease in feline infectious peritonitis: comparisons with dengue haemorrhagic fever. Comparative Immunology, Microbiology and Infectious Diseases. 4. 175-189. 1981.

Williams RK, Jiang GS, and Holmes KV. Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. PNAS. 88:5533-5536. 1991.

Wong SK, Li WH, Moore MJ, Choe H, Farzan M. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. Journal of Biological Chemistry. 279:5. 3197-3201. 2004.

Woo PCY, Lau SKP, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BHL, Wong HL, Poon RWS, Cai JJ, Luk WK, Poon LLM, Wong SSY, Guan Y, Peiris JSM, Yuen KY. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. Journal of Virology. 79:2. 884-895. 2005.

Woo PCY, Lau SKP, Li KSM, Poon RWS, Wong BHL, Tsoi HW, Yip BCK, Huang Y, Chan KH, Yuen KY. Molecular diversity of coronaviruses in bats. Virology. 351:1. 180-187. 2006.

Xiao XD, Chakraborti S, Dimitrov AS, Gramatikoff K, Dimitrov DS. The SARS-CoV S glycoprotein: expression and functional characterization. Biochemical and Biophysical Research Communications. 312:4. 1159-1164. 2003.

Xue XY, Yang HT, Shen W, Zhao Q, Li J, Yang KL, Chen C, Jin YH, Bartlam M, Rao ZH. Production of authentic SARS-CoV M-pro with enhanced activity: Application as a novel tag-cleavage endopeptidase for protein overproduction. Journal of Molecular Biology. 366:3. 965-975. 2007.

Xue XY, Yu HW, Yang HT, Xue F, Wu ZX, Shen W, Li J, Zhou Z, Ding Y, Zhao Q, Zhang XJC, Liao M, Bartlam M, Rao Z. Structures of two coronavirus main proteases: Implications for substrate binding and antiviral drug design. Journal of Virology. 82:5. 2515-2527. 2008.

Yang HT, Yang MJ, Ding Y, Liu YW, Lou ZY, Zhou Z, Sun L, Mo LJ, Ye S, Pang H, Gao GF, Anand K, Bartlam M, Hilgenfeld R, Rao ZH. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proceedings of the National Academy of Sciences of the United States of America. 100:23. 13190-13195. 2003.

Yang ZY, Huang Y, Ganesh L, Leung K, Kong WP, Schwartz O, Subbarao K, Nabel GJ. pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. Journal of Virology. 78:11. 5642-5650. 2004.

Yang ZY, Kong WP, Huang Y, Roberts A, Murphy BR, Subbarao K, Nabel GJ. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. Nature. 428:6982. 561-564. 2004.

Yeager CL, Ashmun RA, Williams RK, Cardellichio CB, Shapiro LH, Look AT, Holmes KV. Human aminopeptidase N is a receptor for human coronavirus 229E. Nature. 357(6377). 420-422. 1992.

Yu ITS, Li YG, Wong TW, Tam W, Chan AT, Lee JHW, Leung DYC, Ho T. Evidence of airborne transmission of the severe acute respiratory syndrome virus. New England Journal of Medicine. 350:17. 1731-1739. 2004.

Ziebuhr J, Heusipp G, Siddell SG. Biosynthesis, purification, and characterization of the human coronavirus 229E 3C-like proteinase. Journal of Virology 71:5. 3992-3997. 1997.

Ziebuhr J, Snijder EJ, Gorbalenya AE. Virus-encoded proteinases and proteolytic processing in the Nidovirales. Journal of General Virology. 81. 853-879. 2000.