Available online at www.sciencedirect.com
ScienceDirect
Virology 359 (2007) 436–446
VIROLOGY
www.elsevier.com/locate/yviro

Mapping antigenic diversity and strain specificity of mumps virus: A bioinformatics approach

Urmila Kulkarni-Kale [a,*], Janaki Ojha [a], G. Sunitha Manjari [a], Deepti D. Deobagkar [b], Asha D. Mallya [c], Rajeev M. Dhere [c], Subhash V. Kapre [c]

[a] *Bioinformatics Centre, University of Pune, Pune 411007, India*
[b] *Department of Zoology, University of Pune, Pune 411007, India*
[c] *Serum Institute of India Ltd., 212/2, Hadapsar, Pune 411028, India*

Received 13 July 2006; returned to author for revision 18 August 2006; accepted 15 September 2006
Available online 1 November 2006

Abstract

Mumps is an acute infectious disease caused by mumps virus, a member of the family *Paramyxoviridae*. With the implementation of vaccination programs, mumps infection is under control. However, due to resurgence of mumps epidemics, there is a renewed interest in understanding the antigenic diversity of mumps virus. Hemagglutinin–neuraminidase (HN) is the major surface antigen and is known to elicit neutralizing antibodies. Mutational analysis of HN of wild-type and vaccine strains revealed that the hypervariable positions are distributed over the entire length with no detectable pattern. In the absence of experimentally derived 3D structure data, the structure of HN protein of mumps virus was predicted using homology modeling. Mutations mapped on the predicted structures were found to cluster on one of the surfaces. A predicted conformational epitope encompasses experimentally characterized epitopes suggesting that it is a major site for neutralization. These analyses provide rationale for strain specificity, antigenic diversity and varying efficacy of mumps vaccines.

Keywords: Mumps virus; Hemagglutinin–neuraminidase; HN; Wild and vaccine strain; Homology model; Sequential epitope; Conformational epitope; Bioinformatics; Molecular immunology

Introduction

Mumps is an acute infectious viral disease characterized by enlargement of the parotid and salivary glands. Other complications of mumps include permanent deafness, orchitis, pancreatitis and aseptic meningitis (AM). The virus usually spreads through respiratory droplets and humans are the only natural host although non-human primates, rodents and other species can be experimentally infected (Wolinsky, 1996).

Mumps virus is a member of the family *Paramyxoviridae*, genus *Rubulavirus* (Rima et al., 1995). The virus is enveloped and its 15.3-kb genome is non-segmented, single-stranded, negative-sense RNA. It codes for seven proteins viz., nucleocapsid (N), phosphoprotein (P), matrix (M), fusion (F), small hydrophobic (SH), hemagglutinin–neuraminidase (HN) and large (L) proteins. Besides these, the phosphoprotein gene (P) also codes for two more proteins, V and I (Paterson and Lamb, 1990). HN is the major antigenic protein, known to elicit neutralizing antibodies. The *SH* gene is known to be the most variable part of the genome and phylogenetic reconstruction studies of different isolates of mumps virus using *SH* genes revealed the existence of 12 genotypes, designated A to L and 3 unnamed potential new genotypes (Jin et al., 2005; Tecle et al., 2001).

Different genotypes have been shown to co-circulate (Afzal et al., 1997a, 1997b; Wu et al., 1998; Tecle et al., 2001, 2002; Takahashi et al., 2000) and their distribution may vary among closely related regions within a country (Takahashi et al., 2000; Tecle et al., 2001). Strains of mumps virus are also known to exhibit varying degrees of neurovirulence (Merz and Wolinsky, 1981; Saito et al., 1996; Rubin et al., 1998, 2000; Rafiefard et al., 2005; Sauder et al., 2006). However, the relative degree of

* Corresponding author. Fax: +91 20 2569 0087.
E-mail addresses: urmila@bioinfo.ernet.in, urmila.kulkarni.kale@gmail.com (U. Kulkarni-Kale).

0042-6822/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2006.09.040

their neurovirulence could not be determined mainly due to lack of animal models. Preliminary results obtained using neonatal rat (Rubin et al., 2000, 2005) and Marmoset monkey models (Saika et al., 2004) for vaccine strains, though promising, require further investigations.

Live attenuated mumps vaccines are available as monovalent mumps, bivalent measles–mumps (MM) and trivalent measles–mumps–rubella (MMR) vaccines. The strains viz., Jeryl Lynn (JL), Leningrad-3 (L-3), L-Zagreb, Rubini and Urabe have been in use all over the world since early 1980s and have been reviewed extensively (Galazka et al., 1999; Furesz, 2002; Folb et al., 2004; Ivancic et al., 2005). It has been reported that the JL vaccine is a mixture of two strains, JL-2 (also called minor) and JL-5 (also called major) (Afzal et al., 1993; Amexis et al., 2002). Similarly, the Urabe vaccine strain has also been reported as a mixture of strains (Brown et al., 1996). The Rubini strain does not appear to provide long-term protection (Galazka et al., 1999; Utz et al., 2004; Ong et al., 2005). While offering varying degrees of protection (Ong et al., 2005), most of the vaccine strains are known to cause adverse reactions such as AM (Flynn and Mahon, 2003; Folb et al., 2004; Nagai et al., 2006). Another challenge is co-circulation of wild-type virus SBL-1 while mumps vaccine is in use (Orvell et al., 1997; Tecle et al., 1998).

Mumps infection is under control due to the implementation of vaccination programs. However, there is a renewed interest to understand the antigenic diversity of mumps virus because of recent outbreaks of mumps epidemics (CDC, 2006). The knowledge thus gained will play a decisive role in mumps vaccine development (Amexis et al., 2001; Furesz, 2002). The contribution of specific humoral response to mumps virus as a defense factor has not been definitively explained (Pipkin et al., 1999; Kacprzak-Bergman et al., 2001). However, HN protein is a major target for humoral immune response in mumps virus infection as it elicits neutralizing antibodies (Tanaka et al., 1992; Cusi et al., 2001). The regions viz., 265–288, 329–340 and 352–360 of HN have been reported to evoke immune response and are also responsible for virulence (Kovamees et al., 1990; Orvell et al., 1997; Cusi et al., 2001). Furthermore, the region 329–340 is shown to have the ability to induce neutralizing antibodies not only to the attenuated virus strains but also to wild-type strains (Cusi et al., 2001).

Nucleotide and protein sequences of HN from a number of mumps virus strains/isolates are available in the public domain repositories. Extensive analyses of these sequences helped in identification of strain-specific variations and characterization of mutants (Yates et al., 1996). Attempts have also been made to correlate mutations with properties such as antigenicity and neurovirulence (Sauder et al., 2006). It has been reported that mutations in HN are distributed over the entire length and no pattern could be detected using sequence data alone (Ivancic et al., 2005).

It is known that the patterns, which are apparently hidden at sequence level, become evident when mapped onto structure. Experimentally derived 3D structure data of HN protein of mumps virus are not available. Therefore, an attempt has been made to predict its 3D structure using knowledge-based homology modeling approach. The structures of HN protein of one wild-type (SBL-1) and three vaccine strains, namely JL2, JL5 and L-Zagreb were predicted. The observed variations of amino acids among the groups of vaccine and wild-type strains were then mapped on the predicted structure. Sequential and conformational epitopes were predicted and analyzed in the context of observed mutations (Kulkarni-Kale et al., 2005; Kolaskar and Kulkarni-Kale, 1999; Kolaskar and Tongaonkar, 1990).

Results

Compilation of HN sequences

A total of sixty-four nucleotide sequences of HN from various strains/isolates of mumps virus were retrieved from GenBank. It must be mentioned that curation of data is an important step in any Bioinformatics analyses. Although information is available, it is scattered either in different entries of the same database or among various databases. For example, only a few sequence entries for HN in GenBank are annotated with respect to genotype data. Since the genotyping of mumps virus is carried out using *SH* gene, the sequence entries of SH for a given strain/isolate were searched to retrieve the genotype information. However, only 49 out of 64 strains could be annotated with respect to genotype information using this approach.

Multiple sequence alignment

The nucleotide and protein sequences of HN were aligned using ClustalW (Chenna et al., 2003). As expected, though there exists sequence similarity among the strains of mumps virus, a few strain-specific variations are observed. Multiple sequence alignment of HN protein showed 90% identity (marked with * in the MSA) and 96% similarity (identity + favorable substitutions that are marked with : and . in the MSA) among vaccine strains (see Annexure V). These values varied up to 74% and 91% respectively in wild-type strains (MSA data not shown). The known motifs viz., leucine-zipper, neuraminidase (240-NRKSCS-245) and receptor-binding site of hemagglutinin (405-GAEGRV-410) are conserved among all 64 entries (Jorgensen et al., 1987; Mirza et al., 1994; Lim et al., 2003). Similarly, there are 9 potential N-linked glycosylation sites with signature sequence N-X-S or N-X-T (Apweiler et al., 1999) at positions: 12, 127, 284, 329, 400, 448, 464, 507 and 514. Of these, glycosylation sites at positions 127, 284, 448, 507, and 514 are conserved in both vaccine and wild-type strains. Mutation of N to D/S at position 12 results in loss of a potential cytosolic glycosylation site in a few strains. This site was found to be missing in the isolates of mumps virus from a vaccinated population in Singapore (Lim et al., 2003). The 3rd amino acid position in the glycosylation site 329 contains either, T or S. However, at similar position in glycosylation site 400, I is found in entries with GenBank accession numbers: AF448531, AF448530, AF448527, AF448534 and AF448528, an unfavorable mutation as far as glycosylation is concerned. Such strains may not get glycosylated and this may account for further