

WHOLE GENOME SEQUENCES OF LOCAL VARICELLA-ZOSTER VIRUS (VZV) STRAINS OF BASRAH CITY/IRAQ

Muntadher Almarjan¹, Hayder Abdulhussein Al-Hmudi², Hassan Nima Habib³

^{1,2}College of Science, University of Basrah, Basrah, Iraq

³College of Agriculture, University of Basrah, Basrah, Iraq

ABSTRACT

Studies of genetic diversity of VZV play an important role in further understanding of the epidemiology and evolution of the virus, and may in future serve as a tool for genetic prediction of virus pathogenicity or resistance development. The aim of the study was to determine the whole genome sequence of local VZV strains for phylogenetic analysis to determine the variability in the viral sequence. Three VZV DNA samples (S1, S2 and S3) were sent to whole genome sequencing, all samples have been successfully sequenced. All clinical strains were found to contain a genome that was 124,884 bp, the ORF map of S1 was drawn. The multiple sequence alignments (MSA) of these samples have been done with different sequences to identify any mutations and finding SNP among ORFs of genomes. The sequences showed the presence of many different mutations or SNPs, some of them were silent, and the other caused a change in the amino acid. Phylogenetic tree was created by using the maximum likelihood method for sequences of three full nucleotide clinical samples with 9 reference VZV strains and the evolutionary distances were showed. Analysis of the phylogenetic tree showed 5 main clades, the present results revealed that the S1, S2 and S3 strains showed similarity or clustered with one sequence, Dumas strain (X04370), corresponding to clade 1, forming one of the clade with genetic distance or sequence divergence 0.0000 for S1 and 0.0001 for S2 and S3 strains reflecting a close relationship of these strains. All studied strains were from genotype E1, including the first fully sequenced VZV isolate clade 1 (Dumas); it is the most commonly circulating genotype in North America and Europe.

I. INTRODUCTION

Varicella-zoster virus (VZV) is a pathogenic human alpha-herpesvirus belongs to the herpesvirus family *Herpesviridae* (Garcés-Ayala et al., 2015). VZV together with herpes simplex virus type 1 (HSV1; HHV1) and type 2 (HSV2; HHV2) were grouped into the subfamily of Alphaherpesvirinae. Moreover, by its genome organization, VZV was classified into the genus *Varicellovirus*, whereas HSV was classified into the genus *Simplexvirus* (Science et al., 2014).

VZV causes chickenpox (varicella) as the primary disease and shingles (zoster) as a recurrent manifestation of infection, both being generally benign and self-limiting. Chickenpox is one of the major childhood diseases that are prevalent in temperate zones with significant mortality (Gilden et al., 2010). Studies of genetic diversity of VZV play an important role in further understanding of the epidemiology and evolution of the virus, and may in future serve as a tool for genetic prediction of virus pathogenicity or resistance development (Ku et al., 2002). The VZV genome is a linear double-stranded DNA molecule of ~125,000 bp that encodes at least 71 unique ORFs and related promoter sequences (Quinlivan & Breuer, 2005). Based on sequence variations, VZV has been classified into five clades (Schmidt-Chanasit and Sauerbrei, 2011). Earlier studies indicated that classification of VZV DNA was based on restriction fragment length polymorphism (RFLP) analysis. The RFLP markers usually considered in VZV clade identification, vaccine selection and epidemiological studies constitute polymorphisms of open reading frames (ORF) 38 (*PstI*), 54 (*BglI*) and 62 (*SmaI*) (Loparev et al., 2000; Sauerbrei et al., 2003). VZV genotyping is usually performed on the basis of sequencing to screen for single nucleotide polymorphisms (SNP) in different ORFs of VZV genome (Faga et al., 2001). The aim of the study was to determine the whole genome sequence of VZV for phylogenetic analysis to determine the variability in the viral sequence.

II. MATERIALS AND METHODS

Three patients diagnosed by a specialist physician who attended Al-Basrah teaching hospital and Al-Fayhaa hospital in Basrah governorate. Viral DNA was extracted from the clinical samples according to the instructions of WizPrep™ the viral DNA/RNA Mini Kit V2 (Wizbiosolutions Inc. South Korea). Three purified viral DNA samples (S1, S2 and S3) were sent to Apical Scientific Laboratories/ Malaysia for sequencing by using next generation sequencing (NGS) for determining of WGSs using a long PCR amplicon-based strategy. Briefly, The DNA was extracted from blood, skin lesion swabs and saliva, and the template quality played a major role in the robustness of the PCR amplification. The genome was divided into 8 segments; sets of primers were used to amplify 80 segments. The WGSs were assembled by using Dumas strain (X04370) as a template. The obtained DNA sequences were compared in order to locate SNPs with WGSs (Table, 1) belonging to different clades were downloaded from the GenBank database of National Center for Biotechnology Information (NCBI). The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model (Tamura and Nei, 1993). The tree with the highest log likelihood (-1517024.39) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 12 nucleotide sequences. There were a total of 125451 positions in the final dataset. Evolutionary analyses were conducted in MEGA X (Tamura et al., 2010).

Table (1) Reference strains used in genome alignment and phylogenetic analysis

GenBank accession no.	VZV strain	Clade	Reference
X04370	Dumas	1	Davison and Scott (1986)
DQ008355	VariVax	2	Tillieux <i>et al.</i> (2008)
AB097933	Oka, substrain pOka	2	Gomi <i>et al.</i> (2002)
AB097932	Oka, substrain vOka	2	Gomi <i>et al.</i> (2002)
AJ871403	HJ0	3	Norberg <i>et al.</i> (2006)
DQ479955	11	3	Peters <i>et al.</i> (2006)
DQ452050	DR	4	Norberg <i>et al.</i> (2006)
DQ479960	8	4	Peters <i>et al.</i> (2006)
DQ457052	CA123	5	Norberg <i>et al.</i> (2006)

I. RESULTS

Whole genome sequencing

Three VZV DNA samples (S1, S2 and S3) were sent to whole genome sequencing, all samples have been successfully sequenced. The first sample S1 was for a 57-year-old female patient with shingles infection, while the second sample S2, it was a 2-year-old female child with chickenpox infection, the third sample S3 was for a 36-year-old male patient with chickenpox infection. By using SNP located in ORF 38 and ORF 54, all studied samples were identified as wild type *PstI*⁺/*BglI*⁺ within the genotype B.

Genome structure

According to assembled of the clinical strains by using Dumas strain (X04370) as a template, the current study showed that all samples have a length 124,884 bp. There are 3 genes in IRS (ORFs 62-64) were inversely repeated in TRS (ORFs 69-71). The structure of the genome in the studied samples was typical of VZV in that the genome could be divided into TRL, UL, IRL, IRS, US and TRS. The ORF map of S1 is presented in figure (1).

Analyses of genomic variation

The multiple sequence alignments (MSA) of these samples have been done (Figure, 2; Figure, 3) with different sequences in table (1) and with Dumas strain (X04370) only to identify any mutations and finding SNP among ORFS of genomes (Table, 2).

The sequences showed the presence of many different mutations or SNPs, some of them were silent, and the other caused a change in the amino acid, and the most important of these mutations were tabulated according to the

presence of these mutations in the Dumas strain (Table, 2). Interestingly, most clinical strains have substitutions leads to a change in amino acids, the ORF0 (nucleotide position 30) was mutated from G-to-T substitution leads to a change in amino acids Aspartic acid to Glutamic acid. The same reading mutation was found in some strains. Most of the mutations were among 0, 62 and 63 ORFs.

On the other hand, all clinical strains have substitutions often completely different from others reference strains. Several strains shared clinical strains the same mutations resulting in a change in the amino acid such as HJ0, vOka and 11 strains. Some reference strains (CA123, DR and VariVax) contain missing nucleotides in ORF62 among S2 samples because of the small size of the genome.

Phylogenetic analysis

Phylogenetic tree was created by using the maximum likelihood method for sequences of three full nucleotide clinical samples with 9 reference VZV strains (Table, 1), and the evolutionary distances were showed. As shown in figure (4), the evolutionary relationships between the studied samples and the reference strains were shown, and the genetic distance was also calculated among the 12 VZV strains. Analysis of the phylogenetic tree showed 5 main clades as suggested by a nomenclature meeting of VZV at 2008 (Breuer *et al.*, 2010). The present results revealed that the S1, S2 and S3 strains showed similarity or clustered with one sequence, Dumas strain (X04370), corresponding to clade 1, forming one of the clade with genetic distance or sequence divergence 0.0000 for S1 and 0.0001 for S2 and S3 strains reflecting a close relationship of these strains. This indicates that all the studied samples belong to clade 1 suggested by a nomenclature meeting of VZV at 2008. All studied strains were of the E1 genotype, which included Dumas strain (clade 1); it is the most common genotype in North America and Europe. However, E1 viruses belong to genotype C by the scattered SNP method (Breuer *et al.*, 2010) and to genotype A by Iowa–Canada Glycoprotein method (Faga *et al.*, 2001). According to clades that suggested by a nomenclature meeting of VZV at 2008 (Breuer *et al.*, 2010), present study showed that clade 2 forming a distinctive clade included vOka, pOka and VariVax strains, clade 3 included HJ0 and 11 strains, also DR and 8 strains forming a clade 4, while only CA123 strain forming a clade 5.

	Dumas	8	11	CA123	DR	VariVax	HJ0	pOkA	vOkA	Type	Amino acids		
S 1	T	G	G	G	G	G	G	G	G	30	Missense	Aspartic acid to Glutamic acid	0
	G	A	A	A	C	A	C	T	C	267	Silent		0
	A	C	C	C	C	C	C	A	C	268	Missense	Proline to Threonine	0
	C	A	A	T	A	A	T	T	C	58447	Missense	Lysine to Glutamine	30
	A	T	C	T	A	C	A	A	G	96324	Missense	Histidine to Glutamine	55
	T	C	A	A	A	C	C	G	A	105710	Missense	Alanine to Valine	62
	T	C	A	C	G	A	G	C	C	110818	Missense	Arginine to Serine	63
	C	G	T	G	C	C	C	A	C	110819	Missense	Arginine to Serine	63
	G	A	T	A	A	G	G	C	C	110820	Missense	Arginine to Serine	63
	C	G	G	G	G	C	T	G	C	110821	Aspartic acid	Aspartic acid to Arginine	63
G	A	C	T	T	A	G	T	G	110822	Aspartic acid	Aspartic acid to Arginine	63	
S 2	G	G	A	A	A	C	A	C	T	267	Silent		0
	T	A	A	A	T	A	T	T	T	1572	Silent		2
	C	T	G	C	C	G	C	A	C	69360	Silent		38
	C	G	T	G	G	A	T	C	C	76198	Missense	Valine to Leucine	41
	G	A	G	G	T	G	T	T	A	108793	Missense	Asparagine to Aspartic acid	62
	T	C	A	C	G	A	G	C	C	110818	Missense	Arginine to Serine	63
	C	G	T	G	C	C	C	A	C	110819	Missense	Arginine to Serine	63
	G	A	T	A	A	G	G	C	C	110820	Missense	Arginine to Serine	63
	C	G	G	G	G	C	T	G	C	110821	Missense	Aspartic acid to Arginine	63
	G	A	C	T	T	A	G	T	G	110822	Missense	Aspartic acid to Arginine	63
	C	T	G	G	A	G	G	T	T	120381	Silent		61
	T	A	A	G	C	G	G	T	G	122559	Silent		62
	G	C	T	G	T	T	C	A	G	122560	Missense	Arginine to Aspartic acid	62
	A	G	A	G	C	C	A	G	A	122561	Missense	Arginine to Aspartic acid	62
T	C	C	G	*	*	*	G	A	124846		Proline to Serine	62	
	T	G	G	G	G	G	G	G	G	30	Missense	Aspartic acid to Glutamic acid	0
	T	A	A	A	G	A	G	C	G	761	Missense	Glutamine to Leucine	1
	C	G	G	G	G	G	C	A	T	762	Missense	Glutamine to Leucine	1
	T	A	A	A	T	A	T	T	T	1572	Silent		2
	C	A	C	C	A	C	G	T	C	32443	Missense	Isoleucine to Leucine	21
	G	A	G	A	T	T	T	T	T	33075	Silent		21
	C	A	G	A	G	C	G	T	G	34302	Silent		22

S 3	T	C	A	C	G	A	G	C	C	T	11081 8	Missense	Arginine to Serine	63
	C	G	T	G	C	C	C	A	C	T	11081 9	Missense	Arginine to Serine	63
	G	A	T	A	A	G	G	C	C	G	11082 0	Missense	Arginine to Serine	63
	C	G	G	G	G	C	T	G	C	T	11082 1	Missense	Aspartic acid to Arginine	63
	G	A	C	T	T	A	G	T	G	A	11082 2	Missense	Aspartic acid to Arginine	63
	A	T	G	T	T	G	A	C	A	A	11295 2	Missense	Valine to Glutamic acid	66
	G	C	G	T	G	T	T	G	A	A	11295 3	Missense	Valine to Glutamic acid	66
	T	A	A	T	C	G	A	C	C	C	11543 3	Missense	Lysine to Isoleucine	67
	C	T	G	G	A	G	G	T	T	T	12038 1	Silent		68

* missing nucleotides because of the small size of the genome

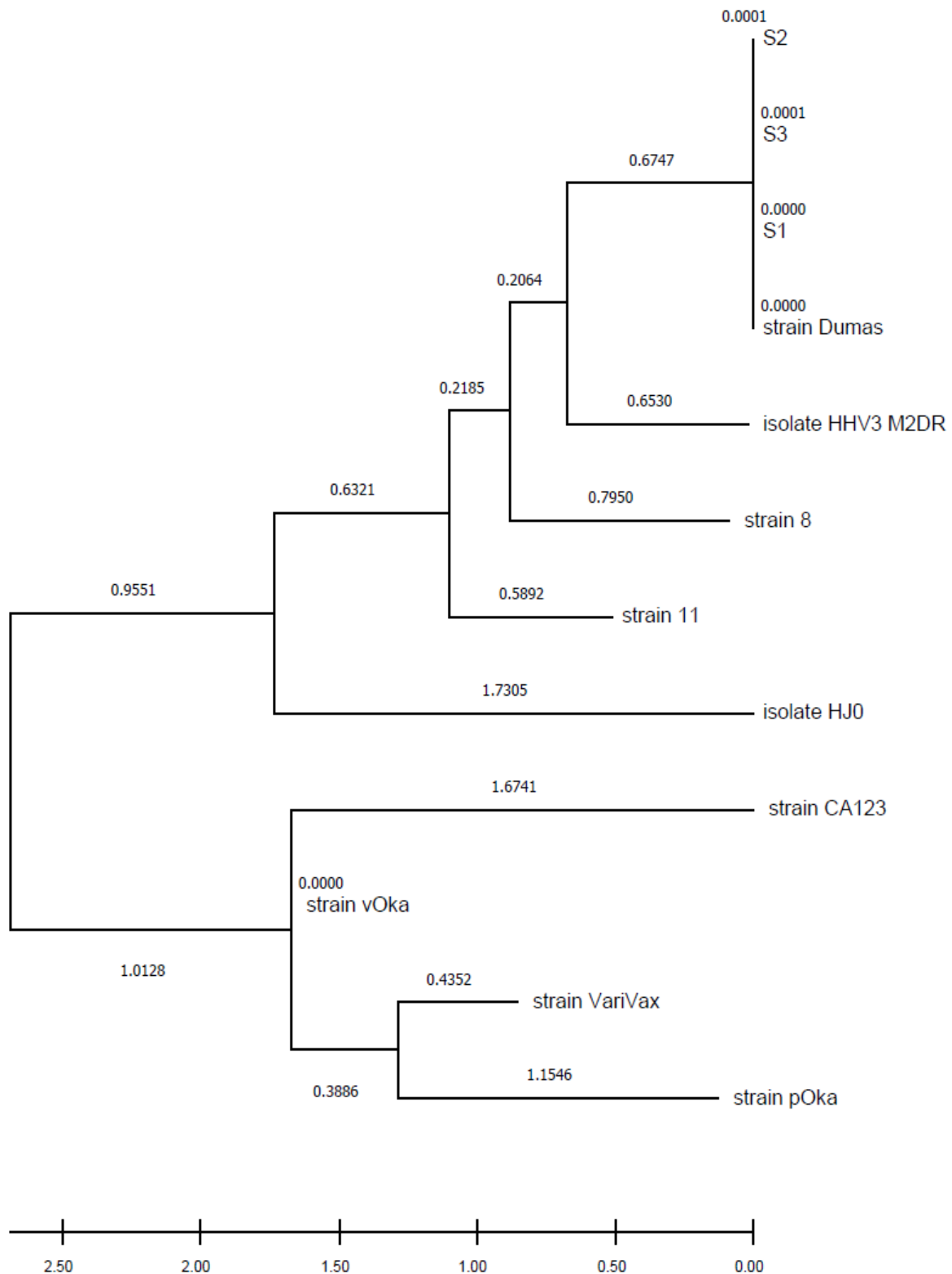


Figure (4) Phylogenetic analysis of 12 VZV strains

II. DISCUSSION

The VZV genome is highly conserved, averaging about one SNP for every 1000 bp of sequence (Loparev *et al.*, 2004; Norberg *et al.*, 2006; Peters *et al.*, 2006). Recent studies suggest that recombination has likely played a role in the emergence of currently circulating genotypes of VZV (Norberg *et al.*, 2006; Peters *et al.*, 2006; Quinlivan *et al.*, 2002). Those three types of genomic variation (excluding repeat regions and ORIs) are observed among VZV strains: 1) genotype-specific mutations, 2) mutations shared by some but not all genotypes, and 3) strain-specific individual mutations. Examples of genotype-specific SNP are observed at positions 35725 (E1), 68101

(E2),101464 (M1), and 11317 (M2); of shared markers at positions 625, 789, 790, 791, 2596, and 6850; and of strain-specific markers at position 100123, which has thus far been observed only in the E1 strain. Several methods have relied on SNP analysis to resolve VZV genotypes (Carr *et al.*, 2004; Loparev *et al.*, 2004, 2007b; Muir *et al.*, 2004; Parker *et al.*, 2006; Sauerbrei *et al.*, 2004). The availability of technically practical, reliable methods for genotyping VZV strains will serve a critical function in countries with broad varicella vaccination policies, since tracking individual strains and identifying probable sources of infection are essential to effectively monitor vaccine impact. In addition, the study of global VZV genotype patterns will likely lead to a better understanding of global transmission patterns both before and after vaccination and of the evolutionary trends of this nearly ubiquitous virus (Loparev *et al.*, 2007).

For the first time in Iraq and the region, complete genome analysis of local strains was studied. However, sequence according to this method is difficult for all samples due to the difficulty of propagating the virus in cell cultures and the low load of the viral DNA in the blood. The sequences showed the presence of many different mutations or SNPs, some of them were silent, and the other caused a change in the amino acid. The evolutionary relationships between the studied samples and the reference strains were shown, and the genetic distance was also calculated among the 12 VZV strains. The present results revealed that the S1, S2 and S3 strains showed similarity or clustered with one sequence, Dumas strain (X04370), corresponding to clade 1, forming one of the clades with genetic distance or sequence divergence 0.0000 for S1 and 0.0001 for S2 and S3 strains reflecting a close relationship of these strains. This indicates that all the studied samples belong to clade 1 proposed by the VZV Nomenclature Meeting 2008. All studied strains were from genotype E1, including the first fully sequenced VZV isolate clade 1 (Dumas); it is the most commonly circulating genotype in North America and Europe. By using SNP located in ORF 38 and ORF 54, all studied samples were identified as wild type *PstI*⁺/*BglI*⁺.

This full-genome sequence-based typing analysis of VZV is to our knowledge the broadest study of differences between clades/genotypes undertaken in the alphaherpesviruses. Most of the previous sequence-based schemes have focused on a limited number of genes, primarily the glycoprotein genes, IE62, and ORFs 1, 21, 22, 50, and 54; examining previously understudied regions has allowed us to locate a number of clade-specific SNPs that can be used in typing strains. The ORF 22 method is unable to distinguish between the European/North American clades A and D, with both classified as genotype E, but is advantageous in ease of use, requiring sequencing of a single 447-bp amplicon to genotype a strain. The scattered SNP method requires sequencing of 4 different ORFs to assign a genotype, which can be somewhat unclear (Loparev *et al.*, 2009). Dumas strain should be regarded as the reference for genotype E1, the HJO strain for genotype E2, the pOka strain for genotype J, the DR strain for genotype M1, and the CA123 strain for genotype M2. Each of these reference viruses is representative of the sequence variations characteristic of the five respective genotypes (Loparev *et al.*, 2007). Most VZV strains collected in the United States had the greatest overall similarity to the reference European strain Dumas. In a like manner, all Japanese strains exhibited maximum similarity to other Japanese isolates and differed the most from Dumas and most U.S. isolates. This was true for both *PstI*⁻ and *PstI*⁺ Japanese strains (Loparev *et al.*, 2004). Strains bearing the Japanese *PstI*⁺/*BglI*⁺ genotype are also common in countries with a history of European colonization, and this is the predominant genotype in equatorial Africa, India, Bangladesh, China, and Western Australia (Loparev *et al.*, 2000). Furthermore, in eastern London, in which a large number of Indian and Bangladeshi immigrants reside, the percentage of *BglI*⁺ VZV strains increased from 10% in the 1980s to more than 30% in the 1990s (Hawrami *et al.*, 1997). *BglI*⁺ strains are also commonly isolated in tropical climates (e.g., equatorial Africa, India, Bangladesh, southern China, Central America, and northern Australia). In contrast, *BglI*⁻ viruses are common in temperate climates (Loparev *et al.*, 2007; Norberg *et al.*, 2006). Recombination between E1 and J strains could also play a role in the appearance of genotype E2 strains. Alternatively, E2 strains might have emerged in Europe (Loparev *et al.*, 2007). All 308 E1, E2 and M4 strains were *PstI*⁺ (ORF38) *BglI*⁻ (ORF54), and the genotype M1, M2 strains identified in this study were *PstI*⁺/*BglI*⁺. Among the European isolates, 319 were *PstI*⁺/*BglI*⁻ (E1, E2 and M4), and 23 were *PstI*⁺/*BglI*⁺ (M1, M2, M3 and *PstI*⁺ J strains). No *PstI*⁻/*BglI*⁺ (the pOka J genotype carried by the vaccine) strains were identified. No *BglI*⁻/*PstI*⁻ isolates were identified in this study (Loparev *et al.*, 2009).

Analysis of the phylogenetic tree of in present study showed 5 main clades as suggested by a nomenclature meeting of VZV in 2008 (Breuer *et al.*, 2010). The detection of all five genotypes predicted by whole-genome phylogenetic analysis required the expansion of ORF22-based sequencing to include sequence data that reliably discriminate genotype E1 and E2 strains. Using ORF22-based sequencing as a foundation for the expanded technique made sense, since the protocol has successfully discriminated most of the major circulating genotypes on nearly 1,000 clinical isolates obtained from countries on every inhabited continent (Dayan *et al.*, 2004;

Loparev *et al.*, 2004; Sergeev *et al.*, 2006). In previous genotypic studies of VZV from Southeast Asian countries, genotype M was the most frequent genotype (Krause *et al.*, 2000). In Bangladesh, genotype M1 was found, and in Nepal and India, genotypes M1 and M2, respectively, were found, in early studies with three isolates from south China, the isolates were genotype M2, and three isolates from north China were genotype E) (Loparev *et al.*, 2004; Loparev *et al.* 2007). In a study of the Middle Eastern region of China, the VZV isolates were genotype J. Therefore, the genotypes of VZV isolates in China differed according to the area where VZV was collected (Liu *et al.*, 2009). However, in a genotypic analysis of 130 VZV samples in the United States, 3% (4/130) were revealed to be genotype J (Sauerbrei *et al.*, 2007).

The current study concludes that the genotypes that spread in Basrah city are the E genotypes, European type which is carried most often PstI+BglI⁻ marker and the other with PstI+BglI⁺ marker, whose complete genome analysis has shown its similarity with the E genotypes that are spread in countries with a history of European colonization.

REFERENCES

- 1 Breuer, J., Grose, C., Norberg, P., Tipples, G., & Schmid, D. S. (2010). A proposal for a common nomenclature for viral clades that form the species varicella-zoster virus: Summary of VZV Nomenclature Meeting 2008, Barts and the London School of Medicine and Dentistry, 24-25 July 2008. *Journal of General Virology*, 91(4), 821–828. <https://doi.org/10.1099/vir.0.017814-0>
- 2 Faga, B., Maury, W., Bruckner, D. A., & Grose, C. (2001). Identification and mapping of single nucleotide polymorphisms in the varicella-zoster virus genome. *Virology*, 280(1), 1–6. <https://doi.org/10.1006/viro.2000.0775>
- 3 Garcés-Ayala, F., Rodríguez-Castillo, A., Ortiz-Alcántara, J. M., Gonzalez-Durán, E., Segura-Candelas, J. M., Pérez-Agüeros, S. I., Escobar-Escamilla, N., Méndez-Tenorio, A., Diaz-Quiñonez, J. A., & Ramirez-González, J. E. (2015). Full-genome sequence of a novel varicella-zoster virus clade isolated in Mexico. *Genome Announcements*, 3(4). <https://doi.org/10.1128/genomeA.00752-15>
- 4 Golden, D., Cohrs, R. J., Mahalingam, R., & Nagel, M. A. (2010). Neurological Disease Produced by Varicella Zoster Virus Reactivation Without Rash (pp. 243–253). Springer, Berlin, Heidelberg. https://doi.org/10.1007/82_2009_3
- 5 Ku, C.-C., Padilla, J. A., Grose, C., Butcher, E. C., & Arvin, A. M. (2002). Tropism of Varicella-Zoster Virus for Human Tonsillar CD4 + T Lymphocytes That Express Activation, Memory, and Skin Homing Markers . *Journal of Virology*, 76(22), 11425–11433. <https://doi.org/10.1128/jvi.76.22.11425-11433.2002>
- 6 Nandal, N., & Nandal, N. (2019). BSCQUAL: A Measuring Instrument of Service Quality for the B-Schools . *International Journal of Psychosocial Rehabilitation*, Vol. 23, Issue 04, 1574-1589
- 7 Loparev, V., Martro, E., Rubtcova, E., Rodrigo, C., Piette, J. C., Caumes, E., Vernant, J. P., Schmid, D. S., & Fillet, A. M. (2007). Toward universal varicella-zoster virus (VZV) genotyping: Diversity of VZV strains from France and Spain. *Journal of Clinical Microbiology*, 45(2), 559–563. <https://doi.org/10.1128/JCM.01738-06>
- 8 Loparev, V. N., Gonzalez, A., Deleon-Carnes, M., Tipples, G., Fickenscher, H., Torfason, E. G., & Schmid, D. S. (2004). Global Identification of Three Major Genotypes of Varicella-Zoster Virus: Longitudinal Clustering and Strategies for Genotyping. *JOURNAL OF VIROLOGY*, 78(15), 8349–8358. <https://doi.org/10.1128/JVI.78.15.8349-8358.2004>
- 9 Loparev, V. N., Rubtcova, E. N., Bostik, V., Tzaneva, V., Sauerbrei, A., Robo, A., Sattler-Dornbacher, E., Hanovcova, I., Stepanova, V., Splino, M., Eremin, V., Koskiniemi, M., Vankova, O. E., & Schmid, D. S. (2009). Distribution of varicella-zoster virus (VZV) wild-type genotypes in northern and southern Europe: Evidence for high conservation of circulating genotypes. *Virology*, 383(2), 216–225. <https://doi.org/10.1016/j.virol.2008.10.026>
- 10 Quinlivan, M., & Breuer, J. (2005). Molecular and therapeutic aspects of varicella-zoster virus infection. In *Expert Reviews in Molecular Medicine* (Vol. 7, Issue 15, pp. 1–24). Cambridge University Press. <https://doi.org/10.1017/S146239940500966X>
- 11 Science, M., Macgregor, D., Richardson, S. E., Mahant, S., Tran, D., & Bitnun, A. (2014). Central nervous system complications of varicella-zoster virus. *Journal of Pediatrics*, 165(4), 779–785. <https://doi.org/10.1016/j.jpeds.2014.06.014>
- 12 Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (n.d.). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. <https://doi.org/10.1093/molbev/mst197>