# DNA MICROARRAYS FOR VZV STRAIN ASSESSMENT

BY

**JOANNE S. HIEBERT**

A Thesis submitted to
the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Medical Microbiology
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

# DNA MICROARRAYS FOR VZV STRAIN ASSESSMENT

BY

JOANNE S. HIEBERT

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

of

MASTER OF SCIENCE

## Table of Contents

**Acknowledgements**

There are many people that I would like to thank. First off I would like to thank Grant Pierce and Elena Dibrov for providing me with my first real taste of the work that goes on in a biology lab. Working with them made me excited to continue my studies as a graduate student and convinced me that I had found the right field of study.

I would also like to thank Graham Tipples for accepting me as a student and providing me with an exciting project. He gave me the opportunity to work on a project that involved learning and applying cutting-edge technology.

I of course also need to thank the members of my examining committee, Charles Grose, Heinz Feldmann and Stephanie Booth as well as the Department of Medical Microbiology at the University of Manitoba.

My time as a graduate student would not have been the same without the people that I worked with everyday including the members of the Viral Exanthemata lab (Michael Gray, David Safronetz, Jennifer Ball and Tamara Kruk) and the Host Genetics and Prions Diseases lab. They (Rhonda, Cathy, Debbie to name a few), provided me with an education not taught in schools and made it fun to go to the lab everyday.

These acknowledgements would not be complete without also thanking Rhonda Mogk and Debra Parchaliuk. Not only did they help me immensely with technical issues, they acted as my personal counselor on more than one occasion. For that I am very grateful. Thank you!

I would also like to take this opportunity to thank my family. I would especially like to thank my Dad and my brother James for their support. Neither one told me that I was crazy when I told them that I was going to do another degree after I had already spent 5 years in University although they were probably thinking it! I wish you both happiness. I owe a great big thank you to Derek Bjarnason. Without you, I wouldn't have been able to finish with a smile on my face. You helped me to see what's really important and not sweat the small stuff. I appreciate all that you have done for me!

This thesis was supported financially through the University of Manitoba Graduate Fellowship and Health Canada Biotechnology Initiative "DNA Microarray Project".

## List of Figures

## List of Tables

## LIST OF ABBREVIATIONS

| | |
|---|---|
| Ω | ohms |
| A | adenine |
| aa-dUTP | 5-(3-aminoallyl)-2'deoxyuridine-5'-triphosphate |
| aa-UTP | 5-(3-aminoallyl)-2'uridine-5'-triphosphate |
| Ala | alanine |
| Arg | arginine |
| Asn | asparagine |
| Asp | aspartate |
| ATP | adenosine triphosphate |
| bp | base pair |
| °C | degrees Celsius |
| C | cytosine |
| cDNA | complementary deoxyribonucleic acid |
| CIP | calf intestinal phosphatase |
| CMC | carboxymethylcellulose |
| CPE | cytopathic effect |
| CTP | cytidine triphosphate |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytidine triphosphate |
| DEPC | diethylpyrocarbonate; Solutions were made RNase free by adding 0.15% DEPC, shaking vigourously, incubating overnight in a fume hood and then autoclaving to deactivate the DEPC |
| dGTP | deoxyguanosine triphosphate |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| dNTPs | mixture of all four deoxynucleotides |
| DTT | dithiothreitol |
| dTTP | deoxythymidine triphosphate |
| E | early |
| *E. coli* | *Escherichia coli* |
| EDTA | ethylene diaminetetracetic acid |
| μF | microfarad ($10^{-6}$ F) |
| FBS | fetal bovine serum |
| μg | microgram ($10^{-6}$ g) |
| ΔG | Activation energy. In the context of nucleic acid hybrids, it refers to the stability of the hybrid and gives a measure of the energy that is needed to destabilize the hybrid. |
| g | gram |
| G | guanine |
| xg | times gravity |
| GFP | green fluorescent protein |
| Gln | glutamine |
| Glu | glutamate |
| Gly | glycine |

| | |
|---|---|
| GTP | guanosine triphosphate |
| His | histidine |
| HSV | herpes simplex virus |
| IE | immediate early |
| Ile | isoleucine |
| kb | kilobases |
| µl | microlitre ($10^{-6}$ l) |
| l | Litre |
| L | late |
| LB | Luria Bertani (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl) |
| Leu | leucine |
| µm | micrometre |
| µM | micromolar ($10^{-6}$ M) |
| M | molar |
| Met | methionine |
| mm | millimetre ($10^{-3}$ m) |
| mM | millimolar ($10^{-3}$ M) |
| mRNA | messenger ribonucleic acid |
| ng | nanogram ($10^{-9}$ g) |
| nl | nanolitre ($10^{-9}$ l) |
| nm | nanometre ($10^{-9}$ m) |
| NML | National Microbiology Laboratory |
| ORF | open reading frame |
| PBS | phosphate buffered saline (50 mM potassium phosphate, 150 mM NaCl, pH 7.2) |
| PCR | polymerase chain reaction |
| PEG | polyethylene glycol |
| PFU | plaque forming unit |
| Pro | proline |
| PSGC | PBS-sucrose-glutamate-serum buffer (1x PBS, 50 g/l sucrose, 1 g/l sodium glutamate, 10% (vol/vol) heat-inactivated FBS) |
| RFLP | restriction fragment length polymorphism |
| RNA | ribonucleic acid |
| rRNA | ribosomal ribonucleic acid |
| rpm | revolutions per minute |
| RT | reverse transcription |
| S | Svedberg unit |
| SDS | sodium dodecyl sulfate |
| Ser | serine |
| SNP | single nucleotide polymorphism |
| SSC (20x) | Saline Sodium Citrate (3 M NaCl, 0.3 M $Na_3C_6H_2O_7 \cdot 2H_2O$, pH 7.0) |
| T | thymine |
| TAE (1x) | Tris-acetate buffer, 40 mM Tris-acetate, 1 mM EDTA |
| Thr | threonine |
| Tris | tris(hydroxymethyl)-aminomethane |
| UTP | uridine triphosphate |

| | |
|---|---|
| UV | ultraviolet |
| V | volts |
| Val | valine |
| VZV | varicella-zoster virus |
| w/v | weight per volume |

## ABSTRACT

Varicella-zoster (VZV) causes varicella (chickenpox) upon primary infection, after which the virus establishes latency. Reactivation of the latent virus causes zoster (shingles). A live attenuated vaccine has been developed and is in use in Canada for the immunization of children. The mechanism of attenuation is unknown although differences in both nucleotide sequence and gene expression have been noted. The sequence differences appear to be mainly differences in repeat number or single nucleotide polymorphisms (SNPs) that are scattered throughout the 125 kb genome. DNA microarrays allow for the analysis of hundreds if not thousands of specific nucleic acid species. Microarrays are commonly used to analyze gene expression through the use of gene specific probes but methods exist in which they can be used to determine the identity of SNPs based on the ability of short oligonucleotides to discriminate between single base mismatches by hybridization. We hypothesized that microarrays would provide a useful tool with which to study both the gene expression and sequence of VZV.

Two separate DNA microarrays were developed using oligonucleotide probes. The first was developed to analyze expression of all 70 open reading frames (ORFs) within the VZV genome and the second to genotype VZV at 62 positions documented to be polymorphic. These 62 SNPs were mainly located in two ORFs (54 and 62) with the exception of two which are found in ORF 38 (*Pst*I restriction site) and ORF 68 (gE 3B3 mutation). The VZV gene expression array proved to be successful at discriminating between cellular and viral transcripts as shown by hybridization to infected and mock-infected cDNA. However, difficulties in culturing enough virus prevented its use in determining the transcription profile of VZV. The VZV SNP array was a success in its

ability to accurately genotype VZV. Eleven strains of VZV were screened with the array and the results compared to those obtained from sequencing. The accuracy of the SNP array was determined to be 90.7% on average. Both arrays promise to be a useful tool in the determination of the mechanism of vaccine attenuation along with a myriad of other potential uses.

# 1.0 INTRODUCTION

## 1.1 Taxonomy and Virology

Reviewed by Roizman and Pellet (2001) and Cohen and Straus (2001).

Varicella-zoster virus (VZV) is a member of the *Herpesviridae* family based on its structure. All herpesviruses consist of a linear, double-stranded DNA genome within an icosadeltahedral capsid, which is surrounded by the tegument and finally an envelope containing viral glycoproteins. The VZV genome is approximately 125 kb in length and encodes at least 70 unique open reading frames (ORFs) (Fig 1). VZV is subclassified as a member of the *Alphaherpesvirinae* subfamily along with the prototype herpes simplex viruses (HSV) by its ability to spread rapidly in culture, lyse host cells and establish latency within sensory ganglia and its relatively short reproductive cycle. It also shares genetic similarity with the prototype HSV-1 in that all but 6 (ORFs 1, 2, 13, 32, 57 and S/L) of its 70 genes have homologs in HSV-1 (Table 1).

## 1.2 Viral Replication

Reviewed by Cohen and Straus (2001) and Sadzot-Delvaux *et al.* (1999).

VZV attaches to host cells through interactions between viral glycoproteins and yet unknown cell surface receptors after which the viral envelope is thought to fuse with the cellular membrane releasing the nucleocapsid and tegument proteins within the cell. The nucleocapsid then migrates to the nucleus where the viral genome is uncoated and released into the nucleus. Three tegument proteins (ORFs 4, 10 and 62) may accompany the nucleocapsid to the nucleus and are thought to initiate viral transcription. It is believed that VZV gene expression follows a tightly coordinated temporal pattern as is

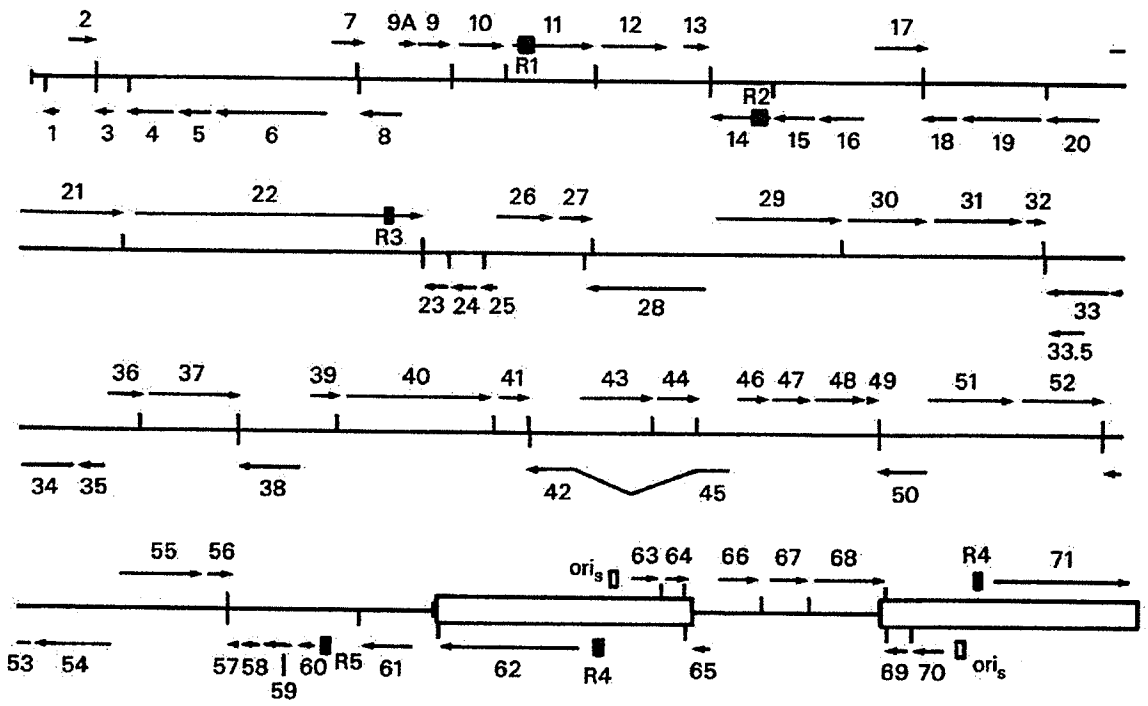



**Figure 1.** VZV ORFs. The genome contains 69 unique ORFs, including ORF9A and ORF33.5 (an in-frame truncation of ORF33). ORFs 69, 70 and 71 are inverted repeats of ORFs 62, 63 and 64. The genome consists of unique regions (horizontal lines), terminal and internal repeats (open boxes) flanking the unique short region, and terminal and internal long repeats flanking the unique long region (short vertical lines). Numbers indicate ORFs and the length and direction of the transcripts are shown by arrows. Potential polyadenylation sites are indicated by light vertical lines. Repeats (R1 to R5) are indicated as filled rectangles. The origins of DNA replication ($ori_s$) are shown. ORF S/L, which begins at the right end of the circularized genome and extends across the terminal repeat into the unique long region at the left end of the genome, is not shown. (Cohen and Straus 2001, adapted from Davison and Scott 1986.)

**Table 1.** VZV genes, functions and HSV-1 homologs. (Cohen and Straus 2001.)

| VZV gene | VZV gene function | Dispensable for growth in cell culture | Protein expressed in latency | HSV-1 homolog |
|---|---|---|---|---|
| 1 | Membrane protein | Yes | | None |
| 2 | | | | None |
| 3 | | | | UL55 |
| 4 | Transactivator, tegument protein | | Yes | UL54 (ICP 27) |
| 5 | gK | No | | UL53 |
| 6 | | | | UL52 (helicase-primase complex) |
| 7 | | | | UL51 |
| 8 | Deoxyuridine triphosphatase | Yes | | UL50 |
| 9 | | | | UL49 |
| 9A | Syncytia formation, virion protein | Yes | | UL49A |
| 10 | Transactivator, tegument protein | Yes | | UL48 (VP16, α-TIF, Vmw65) |
| 11 | | | | UL47 (VP13-14) |
| 12 | | | | UL46 (VP11/12) |
| 13 | Thymidylate synthetase | Yes | | None |
| 14 | gC | Yes | | UL44 |
| 15 | | | | UL43 |
| 16 | | | | UL42 (polymerase processivity factor) |
| 17 | | | | UL41 (virion shutoff protein) |
| 18 | Ribonucleotide reductase, small subunit | | | UL40 |
| 19 | Ribonucleotide reductase, large subunit | Yes | | UL39 |
| 20 | | | | UL38 (VP19C) |
| 21 | Nucleocapsid protein | | Yes | UL37 |
| 22 | | | | UL36 |
| 23 | | | | UL35 (VP26) |
| 24 | | | | UL34 |
| 25 | | | | UL33 |
| 26 | | | | UL32 |
| 27 | | | | UL31 |
| 28 | DNA polymerase | | | UL30 |
| 29 | ssDNA binding protein | | Yes | UL29 (ICP8) |
| 30 | | | | UL28 |
| 31 | gB, fusogen | | | UL27 |
| 32 | Probable substrate for ORF47 kinase | Yes | | None |
| 33 | Protease | | | UL26 (VP24) |
| 33.5 | Assembly protein | | | UL26.5 (VP22) |
| 34 | | | | UL25 |
| 35 | | | | UL24 |
| 36 | Thymidine kinase | Yes | | UL23 |
| 37 | gH | | | UL22 |
| 38 | | | | UL21 |
| 39 | | | | UL20 |
| 40 | Major nucleocapsid protein | | | UL19 (VP5) |
| 41 | | | | UL18 (VP23) |
| 42/45 | | | | UL15 |
| 43 | | | | UL17 |
| 44 | | | | UL16 |
| 46 | | | | UL14 |
| 47 | Protein kinase, tegument protein | Yes | | UL13 (protein kinase) |
| 48 | | | | UL12 (deoxyribonuclease) |
| 49 | | | | UL11 |
| 50 | | | | UL10 (gM) |
| 51 | Origin binding protein | | | UL9 |
| 52 | | | | UL8 (helicase/primase complex) |
| 53 | | | | UL7 |
| 54 | | | | UL6 |
| 55 | | | | UL5 (helicase/primase complex) |
| 56 | | | | UL4 |
| 57 | Cytoplasmic protein | Yes | | None |
| 58 | | | | UL3 |
| 59 | Uracil-DNA glycosylase | Yes | | UL2 |
| 60 | gL, chaperone for gH | | | UL1 |
| 61 | Transactivator, transrepressor | Yes | | ICP0 (Vmw110) |
| 62, 71 | Transactivator, tegument protein | | Yes | ICP4 (Vmw 175) |
| 63, 70 | Tegument protein | | Yes | US 1 (ICP22, Vmw 68) |
| 64, 69 | | | | US10 |
| 65 | Virion protein | Yes | | US9 |
| 66 | Putative protein kinase | Yes | | US3 |
| 67 | gI | Yes | | US7 |
| 68 | gE | | | US8 |
| S/L | Cytoplasmic protein | Yes | | None |

seen in HSV-1 but it has never been demonstrated. Transcription of the HSV-1 genome occurs in three kinetic classes: immediate early (IE), early (E) and late (L) and this is postulated to occur in VZV as well (Fig 2). The first genes expressed, 4 to 6 hours after infection, are termed α genes and encode IE proteins. These IE proteins travel back into the nucleus where they turn on the transcription of β genes that encode E proteins such as the DNA polymerase and thymidine kinase proteins. The E mRNAs are translated and the subsequent proteins are transported to the nucleus where they are involved in replication of the viral genome. The genome, which circularizes upon entry into the cells, is probably replicated by a rolling circle model to generate concatemers that are cleaved to give linear genomes ready to be packaged. After DNA replication, γ genes are transcribed and the resulting L mRNAs are translated. The L proteins include capsid proteins which migrate to the nucleus where nucleocapsids are assembled and the nascent viral genomes packaged. How the naked nucleocapsids egress from the nucleus, acquire envelopes containing the tegument proteins and studded with viral glycoproteins and are released from the cell remains controversial with differing pathways hypothesized. This process results in lysis of the host cell and can be completed in as little as 8 to 16 hours.

Multi-nucleated giant cell or syncytia formation is a hallmark of VZV infection, evident both in skin biopsies and cell culture (Nikkels *et al.*, 1995, Tyzzer, 1906, and Weigle and Grose, 1983). In cell culture, VZV remains highly cell-associated with little infectious virus released (Taylor-Robinson, 1959 and Weller *et al.*, 1958). Spread of the infection appears to occur primarily by virus-mediated cell-cell fusion rather than release and entry of viral particles both *in vitro* and in human skin *in vivo* (Cole and Grose, 2003). A recent paper postulates that viral glycoproteins embedded in the membrane of

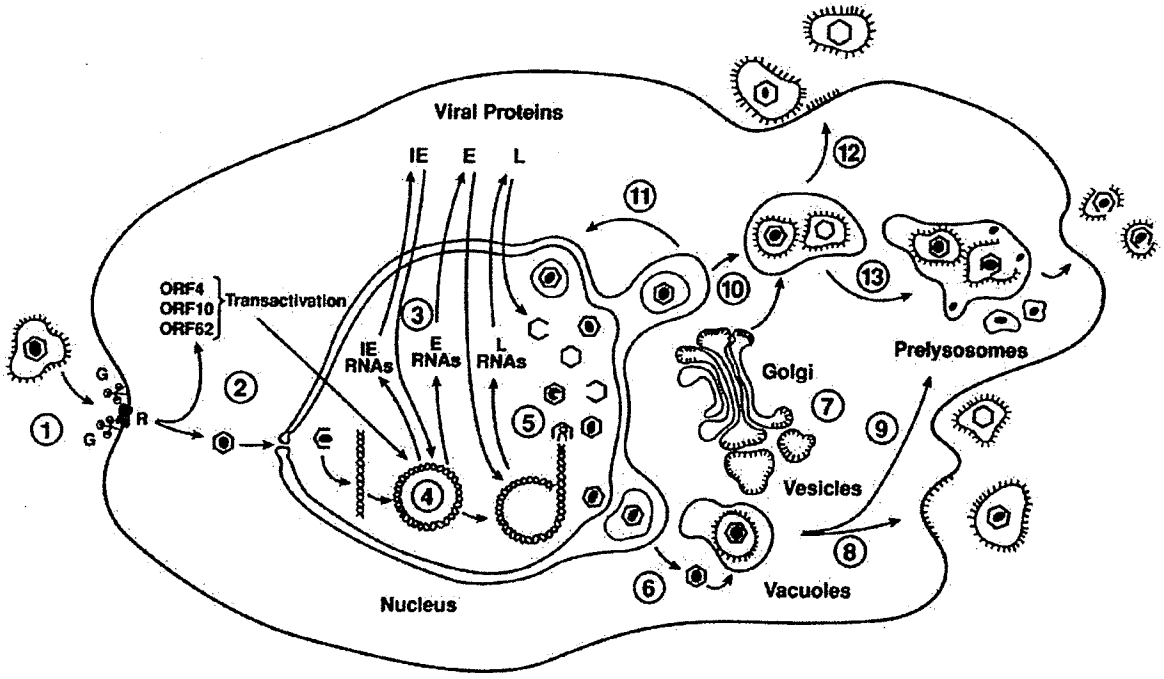



**Figure 2.** Life cycle of VZV. The schematic shown is compiled from multiple sources as well as analogies from HSV. (1) Infectious virions attach to cell surface glycosaminoglycans (G), engage unidentified receptors (R) and nucleocapsids are released into the cell. (2) The nucleocapsids migrate to the nucleus where regulatory tegument proteins encoded by ORFs 4, 10, and 62 initiate (3) a cascade of viral transcription resulting in synthesis of immediate-early (IE), early (E), and late (L) proteins. (4) Viral DNA is replicated via a rolling-circle model. (5) Nucleocapsids are assembled and nascent viral genomes are packaged. Controversy exists regarding the mode of nucleocapsid egress from the nucleus, the site of envelopment, and the pathway leading to release of particles from the infected cell and is reflected in the schematic. (6) In one view, nucleocapsids are transiently enveloped as they bud from the nucleus but then lose their envelopes when they fuse with the rough endoplasmic reticulum. (7) Virion glycoproteins are synthesized and line the inner surfaces of cytoplasmic vesicles while viral tegument proteins line the outer surfaces of the vesicles. (8) The nucleocapsids are drawn into these vesicles and enveloped virions result. These vacuoles may fuse with the cell membrane to release infectious progeny; however (9) many of them fuse with pre-lysosomes resulting in partial degradation before release and an abundance of defective progeny VZV. (10) Another model proposes that nucleocapsids retain their envelopes as they egress the nucleus which are then glycosylated by fusion with Golgi-derived vesicles. Late in infection some particles reenter the nucleus (11), although most are transported to the cell surface (12) for release or (13) to pre-lysosomes. Pathways that avoid the pre-lysosomes (8 and 12) must exist in humans for efficient spread of infection. (Cohen and Straus 2001.)

the infected cell, namely gH and gE, can mediate fusion to a neighbouring cell membrane and this cell may then become infected by passage of infectious particles (nucleocapsids) in an independent process (Cole and Grose, 2003). The authors suggested that cell fusion can occur without subsequent infection which is supported experimentally by the discovery of cells within syncytium that do not contain the major regulatory proteins IE62 and IE63 indicating the absence of infection (Kenyon *et al.*, 2002).

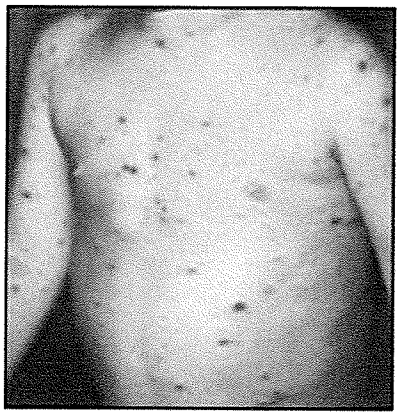
## 1.3 Clinical Disease

### 1.3.1 Varicella

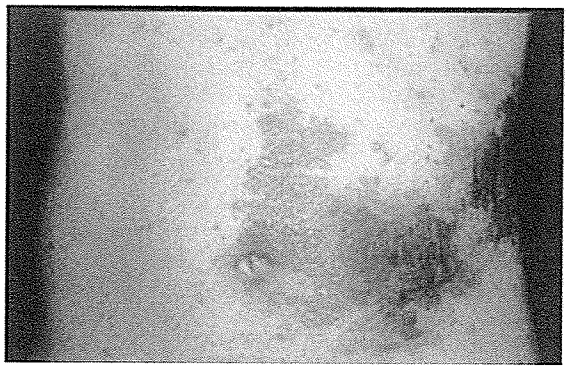
Reviewed by Arvin (2001), Seward *et al.* (2000) and LaRussa (2000).

Primary infection with VZV causes varicella also known as chickenpox. Varicella is a highly contagious largely childhood disease with 90% of children infected by the age of 12 in countries with temperate climates such as Canada (NACI, 2002). It is characterized by itchy lesions on the trunk, face, scalp and extremities (Fig 3a) and can be accompanied by fever, headache, malaise and anorexia (LaRussa, 2000). Healthy children have an average of 300 lesions throughout the course of the disease (Ross *et al.*, 1962) with a range of as few as five to over 1000 (LaRussa, 2000). Severe disease can develop: the highest risk of severe disease and hospitalization is for adults (6-15 times higher risk), followed by infants (1.3-8 times higher) and finally children 1-4 or 5-9 with the lowest risk (Guess *et al.*, 1986, Wharton *et al.*, 1990, Fairley and Miller, 1996 and Wharton, 1996). Varicella is seen to be a mild illness of childhood but serious complications can occur and are more likely to happen when varicella occurs in adolescence or adulthood as demonstrated by case fatality rates which are 10-30 times

A.



B.



C.

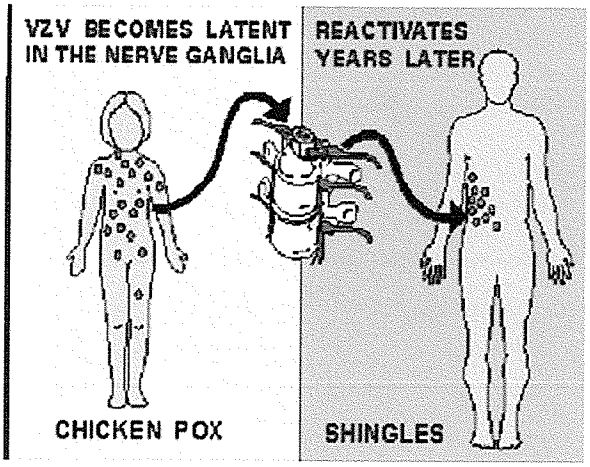



**Figure 3.** Disease caused by VZV. A shows a picture of a child with varicella. (www.vzvfoundation.org/VZVInfections.htm) B is a picture of a patient with zoster (www.vzvfoundation.org/VZVInfections.htm) The cartoon shown in C depicts primary disease (varicella), latency and reactivated disease (zoster). (www.racoon.com/herpes/BH/Hvinfo2/sld011.htm)

higher in adults than in children (NACI, 2002). Complications most frequently involve the skin and soft tissue but can also be neurologic in nature. In healthy children, the most frequent complication of varicella is secondary bacterial infections of skin lesions (*Staphylococcus* or *Streptococcus*) (LaRussa, 2000). Other complications in children can include septicemia, pneumonia, empyema, osteomyelitis, fasciitis and supporative arthritis (LaRussa, 2000). The most common complication in adults is pneumonia (Choo *et al.*, 1995, and Peterson *et al.*, 1996). Neurologic complications include Reye's syndrome, cerebellar ataxia and encephalitis (LaRussa, 2000). In children, the most common neurologic complication is cerebellar ataxia (Applebaum *et al.*, 1953 and Peters *et al.*, 1978) but it is self-limiting and those affected recover completely (LaRussa, 2000). In adults, encephalitis is the most frequent neurologic complication with a mortality rate of 35% (Straus *et al.*, 1988). Varicella in immunocompromised individuals is generally more severe than in healthy persons and more likely to involve multiple organs (LaRussa, 2000).

### 1.3.2 Zoster

Reviewed by Arvin (2001), Schmader (2000) and Oxman (2000).

After primary infection with VZV, the virus establishes latency within sensory nerve ganglia and reactivation of this endogenous virus, typically decades after primary infection, results in zoster, also known as shingles (Fig 3c). Zoster is characterized by localized pain and a vesicular rash generally limited to a single dermatome (Fig 3b). The clinical manifestations of shingles can be divided into three categories: the prodrome, the rash, and pain.

The prodrome precedes the rash by a few days and usually consists of pain and

paresthesia in the affected dermatome and acute neuralgia. The neuralgia is due to replication and spread of the reactivated virus in the ganglion, causing neuronal necrosis and inflammation, as the virus moves down the sensory nerve to the skin. This process can manifest itself in a range of symptoms from itching, tingling, and burning to sharp, stabbing pain. Prodromal pain occurs in the majority of zoster patients over the age of 60 but is rare in healthy individuals under the age of 30.

The localized nature of the rash is the most distinctive feature of zoster and earned it its name from the Greek term for girdle referring to the appearance when dermatomes partially encircling the abdomen are involved (Arvin, 2001) (Fig 3b). It consists of a cluster of vesicular lesions localized unilaterally within the dermatome innervated by the sensory nerve ganglion from which the virus has reactivated. The severity and duration of the rash increases with age with children experiencing rashes that are the least severe and of the shortest duration. The rash usually resolves completely but in some cases it may result in changes in pigmentation, scarring and chronic sensory abnormalities.

In elderly zoster patients, pain is a major feature of the disease and begins a few days before the onset of rash in more than half of the patients (McKendrick *et al.*, 1986). The pain also accompanies the rash in 90% of patients aged 60 years and older but is rare in children and uncommon in younger adults (de Moragas and Kierland, 1957, Burgoon *et al.*, 1957, and Wood, 1991). This pain usually subsides as the rash resolves but may persist for months or years beyond in some patients and is termed post-herpetic neuralgia. Some patients may experience pain without any rash in a syndrome termed *zoster sine herpete* (Gilden *et al.*, 1992, 1994a and 1994b).

The most common complication of zoster is post-herpetic neuralgia (PHN),

debilitating pain that lasts beyond the rash. About 10% of all zoster patients develop PHN that lasts from 4 weeks to more than 10 years (Ragozzino *et al.*, 1982a and 1982b). Age is not only a major risk factor for developing zoster and severity of disease; it is also the most important risk factor for PHN. PHN is rare in younger adults but occurs in more than half of zoster patients over the age of 60 (Oxman, 2000). As well as age, prodromal pain and severe pain during the rash are indicators of the likelihood of PHN. Other complications include anesthesia of the involved dermatome and secondary bacterial infections of the lesions. Disseminated zoster can occur in children and immunocompromised patients and can be mistaken for a second occurrence of varicella (Rogers and Tindall, 1972 and Munoz *et al.*, 1998).

### 1.4 Varicella Vaccine

Reviewed in Arvin (2001) and Takahashi and Plotkin (2000).

A live attenuated vaccine was developed in Japan by Takahashi in 1974 and is termed the Oka vaccine (Takahashi *et al.*, 1974). It was developed by passaging a clinical varicella isolate (parental Oka strain) in tissue culture multiple times in different cell lines including primary human embryonic lung cells (HEL) (11 times) and guinea pig embryo fibroblasts (GPEF) (six times) after which time the virus had become attenuated. Studies in the SCID-hu mouse model have shown that the vaccine strain is not able to replicate in human skin as well as low-passage wild type strains but retains the virus' ability to infect T cells (Moffat *et al.*, 1998). As well, when higher-passage VZV strains (such as VZV-Ellen) were studied in the SCID-hu mouse, they too showed decreased ability to replicate in skin, thus it appears that serial passage in tissue culture is sufficient for attenuating VZV virulence for human skin (Moffat *et al.*, 1998).

To date, the molecular mechanisms of attenuation have not been determined although differences between wild type and vaccine strains have been noted. Some of these include differences in restriction enzyme cleavage with enzymes such as *Hpa*I, *Eco*RI (Hayakawa *et al.*, 1984), *Bam*HI, *Bgl*I (Martin *et al.*, 1982), *Pst*I (Brunell *et al.*, 1987), and *Sma*I (Gomi *et al.*, 2000 and Argaw *et al.*, 2000).

Attempts to determine genes responsible for attenuation have included investigation into glycoprotein C (gC, ORF14) with lower levels of this glycoprotein seen in the vaccine (Kinchington *et al.*, 1990) but no difference was seen between the parental and vaccine strains at the sequence level in this gene. Recently, attention has shifted to viral gene regulators and in particular, IE ORF 62 (IE62). A recent study completed by Gomi *et al.* (2000) revealed that the Oka vaccine strain consists of a mixed population with respect to IE62 sequence with differences in 15 out of the 3929 bases that encode IE62 whereas parental Oka consisted of a single clone. They speculated that due to the high number of mutations found in the vaccine IE62 ORF and this protein's role as a major transactivator early in the viral life cycle, that this gene may play an important role in attenuation.

## 1.5 Genotyping VZV

The Oka varicella vaccine was licensed for use in Canada in 1999 under the market name of Varivax™ and Varivax II™ (Merck). It is currently recommended for susceptible infants over the age of 12 months by Health Canada (NACI, 2002). As a result of the fact that the vaccine is a live attenuated virus there is a need to be able to distinguish vaccine from wild type strains. Recently restriction fragment length polymorphism (RFLP) methods have been used to genotype wild type and vaccine strains

and involves PCR followed by restriction enzyme digestion (LaRussa *et al.*, 1992, Takayama *et al.*, 1996, Argaw *et al.*, 2000, and Gomi *et al.*, 2000). There are three such RFLPs exploited by our lab for genotyping purposes and they include *Pst*I, *Bgl*I and *Sma*I sites in ORFs 38, 54 and 62 respectively (LaRussa *et al.*, 1992 and Argaw *et al.*, 2000). The Oka vaccine strain is *Pst*I negative and both *Bgl*I and *Sma*I positive whereas wild type strains (in North America) are *Pst*I positive, either *Bgl*I negative or positive, and *Sma*I negative. The combination of these three sites allows for differentiation of vaccine from wild type strains as well grouping of wild type isolates by *Bgl*I. The presence or absence of these three restriction sites is due to single nucleotide polymorphisms (SNP) in all three cases: A69349G causes the loss of a *Pst*I site, T95241C results in a gain of a *Bgl*I site, and T106262C causes the gain of a *Sma*I site in Oka vaccine (Hondo *et al.*, 1989, LaRussa *et al.*, 1992, and Argaw *et al.*, 2000 respectively). Recently, real-time PCR methods have been developed for rapid differentiation of vaccine and wild-type strains and employ the *Sma*I (Loparev *et al.*, 2000) and *Pst*I (Tipples *et al.*, 2003) SNPs.

In 1998 Santos *et al.* reported the discovery of a VZV strain, VZV-MSP, that lacked a major B cell epitope termed the 3B3 epitope. The authors determined that this loss of the 3B3 epitope was due to a single point mutation of a G to an A at codon 150 that resulted in an amino acid change from aspartic acid to asparagine. The VZV-MSP strain was shown to display increased virulence in tissue culture and in the SCID-hu mouse model (Santos *et al.*, 2000) as well as enhanced cell-cell fusion (Cole and Grose, 2003). Recently, a zoster isolate in BC was found to have this same mutation (VZV-BC) and had been the cause of a severe case of shingles in an elderly patient (Tipples *et al.*, 2002).

Hundreds of other SNPs have been documented throughout the 125 kb VZV genome. Many of these, other than the four discussed, could be used to distinguish wild type from vaccine as well as subgroup wild type isolates. As well, different phenotypes such as increased virulence (MSP) or attenuation (Oka) may be associated with certain mutations that could shed some insight into their phenotypes.

## 1.6 DNA Microarrays

DNA microarrays are a relatively new technology that allow for the hybridization of target nucleic acid samples to thousands of probes immobilized on a solid support such as a glass slide or a membrane. These probes can be double stranded DNA in the form of cDNAs or PCR products or single stranded oligonucleotide probes of a defined sequence and length. The target nucleic acid sample is either fluorescently (glass microarrays) or radioactively labelled (membrane arrays), hybridized to the array and the degree of hybridization, and thus the amount of a particular nucleic acid species present, is quantified by the amount of fluorescence or radioactivity associated with each specific probe.

### 1.6.1 Gene Expression Arrays

DNA microarrays have frequently been used to analyze gene expression of both prokaryotic and eukaryotic cells (Schena *et al.*, 1995, DeRisi *et al.*, 1996, Schena, 1996, and Ramsay 1997). In the past few years, a number of groups have developed gene expression microarrays for herpesviruses. In 1999, Chambers *et al.* developed the first DNA array for measuring the expression of nearly all known ORFs in human cytomegalovirus (HCMV), the largest of the herpesviruses. In this study, the authors made oligonucleotide probes, 75 bases in length, for the sense strand of the

approximately 200 HCMV ORFs and these probes were arrayed onto glass slides. They infected cells *in vitro* with HCMV, treated the infected cells with either an inhibitor of protein synthesis (cycloheximide) or viral DNA replication (ganciclovir) and harvested RNA to hybridize to the array at appropriate times. By blocking protein synthesis and harvesting early in infection, the authors were able to determine which of the viral ORFs belong to the IE kinetic class. By harvesting at later times in the viral life cycle and blocking viral DNA replication, the authors were able to determine the E class of transcripts and then by harvesting late in infection, without any treatment, the L class of transcripts could be determined. The authors concluded that their HCMV DNA chip was a reliable and robust method of testing HCMV transcripts in the cellular transcript pool and the expression patterns they determined with their chip were in agreement with previously studied viral ORFs. The following year, another group (Bresnahan and Shenk, 2000) used a series of PCR products to probe all 208 ORFs in the HCMV genome in a membrane array. Using their HCMV array, the authors discovered a subset of viral transcripts that were packaged within the viral capsids. Also in 2000, Stingley *et al.* reported on the generation and use of a DNA microarray for herpes simplex virus type 1 (HSV-1). This array consisted of 75-mer oligonucleotide probes for HSV-1 and cellular transcripts spotted onto glass slides. Stingley used the same approach as Chambers did earlier that year for HCMV and used inhibitors of protein (cycloheximide) and viral DNA synthesis (phosphonoacetic acid) to classify viral transcripts into the three kinetic classes of IE, E, and L. Like Chambers, the authors of the HSV-1 array reported that the microarray classification of viral ORFs was in good agreement with previous studies of individual transcripts using more traditional methods such as Northern blots and RNase

protection.

In 2001, two independent groups reported on the construction of gene expression microarrays for Kaposi's sarcoma-associated herpesvirus (KSHV or human herpesvirus-8) and their use in determining KSHV's transcription program in both latent and lytic infection (Jenner *et al.*, 2001 and Paulose-Murphy *et al.*, 2001). Both groups used PCR products as probes for all known KSHV ORFs but the difference is that Jenner *et al.* constructed nylon membrane arrays and Paulose-Murphy *et al.* glass slide arrays. Both groups induced latently infected cells and collected RNA at multiple time points thereafter. Both papers stated that the transcription patterns determined with their respective KSHV arrays correlate well with results of previously published studies of KSHV transcription and also with the functions of specific genes.

In 2002, Ahn *et al.* published a paper describing the construction of cDNA nylon array for murine gammaherpesvirus 68 (MHV-68) and the use of the array, along with inhibitors of protein and viral DNA synthesis, to characterize viral transcription and classify genes according to their transcription kinetics. With their array, the authors were able to classify 73 of MHV-68 genes into the three temporal classes of IE, E and L. The authors concluded that their cDNA array was reproducible with a large range of detection and that the array results were in agreement with transcription kinetics determined for the few genes that have been studied by other methods such as Northern blots and RNase protection assays. In 2003, another array was made for MHV-68 (Ebrahimi *et al.*) using 75-mer oligonucleotide probes in a glass array. The authors used their array in conjunction with inhibitors of protein and viral DNA synthesis to classify 77 of the viral genes into three kinetic classes. The authors noted that their data correlated well with

previous studies with respect to specificity and sensitivity but that there were discrepancies. For instance, the earlier MHV-68 cDNA array classified more of the 80 or so viral genes as having late expression (42) than did the Ebrahimi oligo array (29) and only 22 were classified as late with both arrays. The authors of the oligo array suggested that the discrepancies may be due in part to the fact that double stranded PCR product probes were used in the earlier study which do not allow for strand specific hybridization for analyzing a genome that often has ORFs overlapping on opposite strands.

Since gene expression microarrays became widely accessible, custom arrays have been made for many herpesviruses and been used successfully in studying transcription kinetics of these viruses.

### 1.6.2 Single Nucleotide Polymorphism Genotyping Arrays

Studies have been published describing the use of the microarray format for determining the identity of SNPs in biological samples ranging from detecting breast cancer *BRCA1* and cystic fibrosis mutations to ciprofloxacin resistance in *Neisseria gonorrhoeae* (Hacia *et al.*, 1996, Cronin *et al.*, 1996, and Ng *et al.*, 2002). In these three instances, the identity of SNPs was determined by hybridization to arrays spotted with short (15-25 nucleotides) oligonucleotide probes that contain the polymorphic base in the central position. Separate probes were included for both wild type and mutant bases resulting in a perfect match and mismatch for any homozygous DNA sample that is hybridized. The presence of a mismatch in the central position in the short oligo probe destabilizes hybridization enough that under sufficiently stringent conditions, mismatched DNA will not remain bound. Thus the identity of an unknown DNA sample is ascertained by determining its hybridization pattern to the probes, with the probe

showing the highest fluorescent intensity corresponding to the identity of the SNP. This type of SNP array format is known as allele specific hybridization or sequencing by hybridization (Fig 4). Hundreds of SNPs have been documented for VZV, some of which are listed above and in Table 2. A high throughput technique such as a DNA microarray would allow for the screening of all SNPs of interest in parallel. Many strains of VZV could be screened and by comparing their results, insights into the processes that lead to virulence or attenuation could be gained.

## 1.7 Objectives

1. To develop and test a DNA microarray which can then be used to study the transcription profile of VZV and classify viral genes according to their kinetics.

2. To develop a genotyping array for VZV that would allow the screening of strains for documented mutations.

## 1.8 Hypothesis

DNA microarrays allow for the analysis of hundreds or thousands of target nucleic acids in a single hybridization. They have proven to be useful in the determination of transcription kinetics for many herpesviruses. We hypothesized that effective microarrays could be developed for studying the transcriptome and genome of VZV.

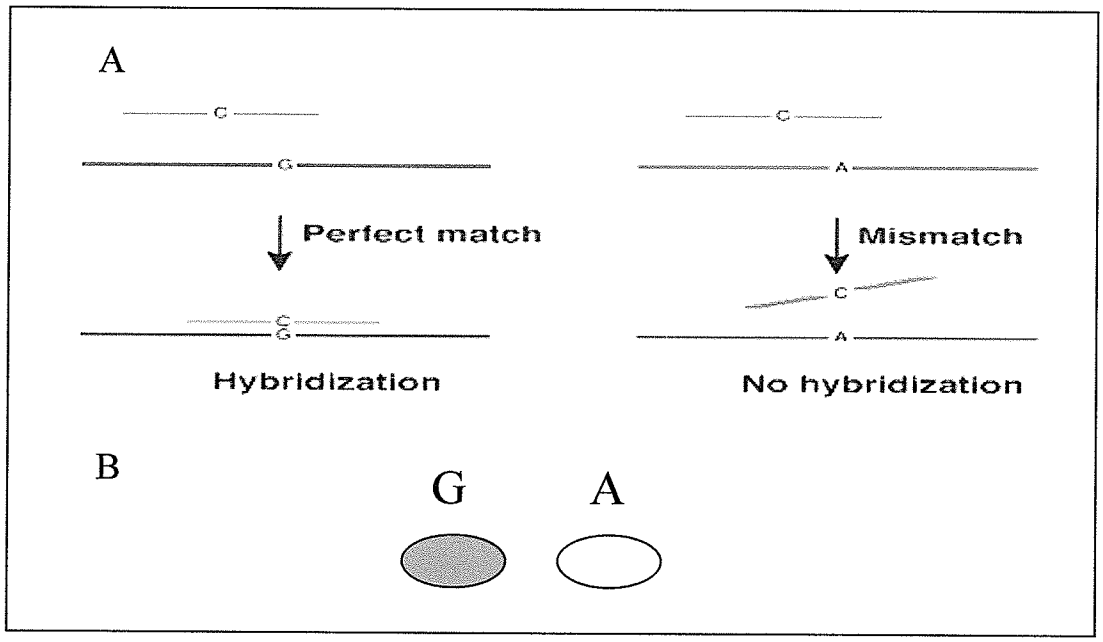



**Figure 4.** Schematic of SNP genotyping array using a sequencing by hybridization format. A depicts a perfect match (on the left) and a mismatch probe (on the right) for the target sample to be hybridized. On the left, the probe with a G in the central polymorphic position is a perfect complement to the target sample and thus the target hybridizes. On the right, the presence of a mismatched A in the central position destabilizes duplex formation enough that under appropriate conditions the mismatched target sample does not remain bound.
(www.lifesciences.perkinelmer.com/areas/snps/protocol.asp#hybridization.)
B demonstrates the expected fluorescent image after hybridization of the target sample in A to its perfect match (G) and mismatch (A) probes. Hybridization to the G probe is indicated in the filled spot and lack of hybridization to the A probe is indicated by the lack of fluorescence or colour in its corresponding spot.

**Table 2.** Single nucleotide polymorphisms compiled from literature.

| SNP | ORF | aa Change | Notes | Reference |
|---|---|---|---|---|
| a69349g | 38 | No change | Loss of *Pst*I site in Oka | Takayama *et al.*, 1996 |
| t94167c | 54 | No change | Creation of *BstX*I site | Argaw *et al.*, 2000 |
| a94632g | 54 | No change | | Argaw *et al.*, 2000 |
| a94641t | 54 | No change | | Argaw *et al.*, 2000 |
| t95241c | 54 | No change | Creation of *Bgl*I site in Oka | LaRussa *et al.*, 1992 |
| c95300t | 54 | Gly -> Ser | | Barrett-Muir *et al.*, 2001 |
| g95546a | 54 | No change | | Argaw *et al.*, 2000 |
| t95601g | 54 | Glu -> Asp | | Stanford web site |
| g105234a | 62 | No change | | C Grose, personal communication |
| g105264c | 62 | | | Argaw *et al.*, 2000 |
| a105310g | 62 | Leu -> Ser & Leu | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| a105312g | 62 | No change | | Argaw *et al.*, 2000 |
| t105356c | 62 | Ile -> Val | Loss of *Sfa*NI site | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| a105371c | 62 | Ser -> Ala | | Faga *et al.*, 2001 |
| t105406c | 62 | Glu -> Gly | | Faga *et al.*, 2001 |
| t105413c | 62 | Ser -> Gly | | Faga *et al.*, 2001 |
| a105451g | 62 | Leu -> Pro | | Gomi *et al.*, 2000 |
| t105490c | 62 | Gln -> Arg | | Faga *et al.*, 2001 |
| t105510c | 62 | No change | | Faga *et al.*, 2001 |
| a105512c | 62 | Ser -> Ala | | Gomi *et al.*, 2000 |
| t105532c | 62 | His -> Arg | | Faga *et al.*, 2001 |
| a105544g | 62 | Val -> Ala | Creation of *Acc*II site | Gomi *et al.*, 2000 |
| t105699c | 62 | No change | | Faga *et al.*, 2001 |
| t105705c | 62 | Ala -> Ala | Creation of *Sac*II site | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| t105855c | 62 | No change | | Faga *et al.*, 2001 |
| t105894c | 62 | No change | | Faga *et al.*, 2001 |
| a105915g | 62 | No change | | C Grose, personal communication |
| t105919c | 62 | Gln -> Arg | | Faga *et al.*, 2001 |
| g105923t | 62 | No change | | Faga *et al.*, 2001 |
| t105964c | 62 | Gln -> Arg | | Faga *et al.*, 2001 |
| t106029c | 62 | No change | | Faga *et al.*, 2001 |
| a106247g | 62 | No change | | Faga *et al.*, 2001 |
| t106262c | 62 | Arg -> Gly | Loss of *Bst*NI site and creation of *Sma*I site | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| t106497c | 62 | No change | | Faga *et al.*, 2001 |

| SNP | ORF | aa Change | Notes | Reference |
|---|---|---|---|---|
| c106569t | 62 | No change | | Wagenaar *et al.*, 2003 |
| a106710g | 62 | No change | | Gomi *et al.*, 2000 |
| a106905g | 62 | No change | | Faga *et al.*, 2001 |
| a107026g | 62 | Val -> Ala | | Faga *et al.*, 2001 |
| t107070c | 62 | No change | | Faga *et al.*, 2001 |
| t107136c | 62 | Ala->Ala | Creation of *BssH*II site in Oka | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| c107165t | 62 | Ala -> Thr | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| t107252c | 62 | Ser -> Gly | Creation of *Nae*I site in Oka | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| t107307c | 62 | No change | | Argaw *et al.*, 2000 |
| g107329a | 62 | Ala -> Val | | Faga *et al.*, 2001 |
| c107586t | 62 | No change | | Faga *et al.*, 2001 |
| a107599g | 62 | Val -> Val & Ala | | Gomi *et al.*, 2000 |
| c107607a | 62 | Thr -> Thr | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| t107715c | 62 | Ala -> Ala | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| a107797g | 62 | Leu -> Leu & Pro | | Gomi *et al.*, 2000 |
| t108111c | 62 | Pro -> Pro | Loss of *BsrI* site | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| c108189t | 62 | No change | | C Grose, personal communication |
| t108551c | 62 | Asn -> Asp | | Faga *et al.*, 2001 |
| t108564c | 62 | No change | | Faga *et al.*, 2001 |
| c108591t | 62 | No change | | C Grose, personal communication |
| a108618g | 62 | No change | | Faga *et al.*, 2001 |
| t108741c | 62 | No change | | Faga *et al.*, 2001 |
| a108747g | 62 | Leu -> Leu | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| a108838g | 62 | Met -> Met & Thr | | Gomi *et al.*, 2000 |
| g108951a | 62 | No change | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| a109010c | 62 | Ser -> Ala | | Faga *et al.*, 2001 |
| c109044g | 62 | Ala -> Ala | | Argaw *et al.*, 2000 and Gomi *et al.*, 2000 |
| a109137g | | Noncoding | | Gomi *et al.*, 2000 |
| g116255a | 68 | Asp -> Asn | Loss of 3B3 B cell epitope | Santos *et al.*, 1998 |

## 2.0 MATERIALS AND METHODS

### 2.1 Cell Culture

Two cell lines were used to culture VZV: MRC-5 human lung fibroblasts (American Type Culture Collection) and MeWo human melanoma cells (graciously provided by Dr. Charles Grose). Both cell lines were cultured in filter sterilized Eagle's Minimum Essential Medium (MEM) (pH 7.2) (Invitrogen) supplemented with 1 mM sodium pyruvate (Invitrogen), 0.1 mM non essential amino acids (Invitrogen), 2 mM glutamine (Invitrogen), and 3.5 g/l glucose (Invitrogen). The cells were supplemented with 10% heat inactivated (56°C, 30 minutes) fetal bovine serum (FBS) (CanSera) for growth; the FBS was decreased to 2% for maintenance and infection. The cells were also supplemented with 100 units/ml penicillin G sodium and 100 µg/ml streptomycin sulfate (Invitrogen) for infection with VZV. The cells were grown in a 37°C humidified incubator with 5% $CO_2$.

For growth of MRC-5 cells in a microcarrier spinner culture, Cytodex™ 1 (Amersham Biosciences) microcarriers were used. Prior to use, the microcarriers were swollen and hydrated overnight in PBS (calcium and magnesium free, 100 ml/g Cytodex™ 1) at room temperature. The following day, the supernatant was decanted and the beads rinsed in fresh PBS (50 ml/g Cytodex™ 1). The supernatant was discarded and the volume brought up with fresh PBS to a final Cytodex™ 1 concentration of 20 g/l. The bead suspension was sterilized by autoclaving and stored at 4°C.

To use the beads, they were dispersed by shaking and the appropriate volume (for the appropriate amount of beads) was removed by pipette and transferred to a sterile 50

ml tube. The beads were allowed to settle, the supernatant decanted and they were washed twice in warm growth medium (100 ml/g Cytodex$^{TM}$ 1) by pipetting up and down. The beads were allowed to settle and the supernatant removed. Warm growth medium was added such that the final volume after the cells were seeded would yield a Cytodex$^{TM}$ 1 concentration of 2 g/l and the beads were transferred to a sterile glass spinner flask. The spinner culture was seeded with MRC-5 cells such that there were 20 cells for every bead (~1:1.5 split ratio by surface area). (e.g. For a 50 ml 2g/l Cytodex$^{TM}$ 1 culture there is 2 x 0.05 = 0.1 g of beads. There are approximately 4.3 x $10^6$ beads/g dry weight of Cytodex$^{TM}$ 1 therefore there are 4.3 x $10^6$ x 0.1 g = 4.3 x $10^5$ beads. Thus for an ideal cell to bead seeding ratio of 20 to 1, 4.3 x $10^5$ x 20 = 8.6 x $10^6$ cells must be used.) The beads were returned to the incubator but left without spinning for approximately an hour to allow the cells to adhere. After this attachment period, the magnetic spinner was set to the minimum rpm that would keep the beads in suspension.

## 2.2 VZV Strains

Several VZV strains were used: Ellen, a lab reference strain (ATCC vr1367); Oka vaccine strain (Merck Frost); VZV-BC (Tipples *et al.*, 2002), a zoster isolate from British Columbia, Canada; and eight clinical isolates from zoster patients from George-L. Dumont Regional Hospital in Moncton, New Brunswick, Canada (referred to as VZV numbers 6, 8, 9, 16, 20, 21, 31, and 55).

## 2.3 Cell Free VZV Preparation using Microcarrier Culture

As much growth medium that could be removed from a 50 ml 2 g/l Cytodex$^{TM}$ 1 MRC-5 culture (seeded the day before) was removed (~43 ml). Two T-25s of VZV-Ellen infected MRC-5s displaying very high CPE (close to 100%) were scraped and the

cells along with the overlaying medium were added to the microcarriers for an infection rate of approximately 1 infected cell to 6 non-infected cells. The culture was put in the incubator with approximately half the volume of medium and left without stirring for one hour. After the attachment period, the volume was brought up to 50 ml with medium lacking FBS and the flask was put on the stir plate in the incubator.

The following day (21 hours post infection), half of the infected microcarrier culture (25 ml) was removed while the remaining was diluted to 50 ml with medium lacking FBS and returned to the incubator. A protocol from Dr. Jennifer Moffat (Upstate Medical University) for preparing cell free VZV by sonication of infected cells was followed on the half of the culture removed except that it was adapted for use with a microcarrier culture rather than tissue culture flasks. The 25 ml culture aliquot was divided evenly into two sterile 50 ml tubes. The beads were allowed to settle and the medium aspirated. The beads with washed three times with cold PBS. Five ml of 0.1% EDTA in PBS was added to both tubes and the cells were left for five minutes at room temperature. After the incubation, one tube was shaken by hand while sterile glass beads were added to the second tube of cells before it was shaken. The supernatant (containing the cells and the microcarriers) was aspirated from the glass beads and placed in a fresh tube. The glass beads were washed with another 5 ml of PBS-EDTA (0.1%) and this was added to the first 5 ml. Five ml of PBS-EDTA (0.1%) was added to the first aliquot of cells that had not been shaken with glass beads. These two 10 ml aliquots of cells were spun at approximately 400x g for 10 minutes at 4°C. The supernatant was removed and the cell (and microcarrier) pellets were resuspended in 2 ml PBS-sucrose-glutamate-serum storage buffer (PSGC: 1x PBS, 50 g/l sucrose, 1 g/l sodium glutamate, 10%

(vol/vol) heat-inactivated FBS) (Harper *et al.*, 1998). These two aliquots were sonicated in a water bath sonicator for 15 seconds and put on ice. The sonicated suspensions were both aliquoted into 4 pre-chilled 1.5 ml cryovials each and stored at –80°C.

After another two days (three days post infection), the remaining half of the infected microcarrier culture (now at 50 ml but still only half of the original infected culture) was split evenly into two sterile 50 ml tubes for polyethylene glycol (PEG) precipitation (Grose *et al.*, 1979 and Grose *et al.*, 1983). Sterile glass beads were added to one aliquot and the tube was shaken by hand. The supernatant (containing the microcarriers) was aspirated from the glass beads and placed into a fresh tube. The second aliquot of microcarriers was shaken without the addition of glass beads. The cells (and the microcarriers) were pelleted at approximately 400x g for 15 minutes. The supernatants were aspirated and put into fresh centrifuge tubes (~21.5 ml each). Sterile sodium chloride and PEG 8000 were added to a final concentration of 0.5 M (using a 5 M stock) and 8% (w/v) (from a 50% stock) respectively. The virus-containing supernatants were precipitated at 4°C overnight and then sedimented at 9380x g for 1.5 hours at 4°C. The pellets were each resuspended in 2 ml PSGC storage buffer, then aliquoted into 4 cryovials each and put at -80°C.

## 2.4 VZV Plaque Titrations

The methods used here are adapted from Grose and Brunell (1978), Grose *et al.* (1979) and Husson-van Vliet *et al.* (1987). One 24-well plate per sample to be titrated was seeded the day before at a split ratio of 1 to 2 (i.e. one T-185 flask to seed two 24-well plates) so that they would be ~80% confluent at the time of infection. An aliquot of the cell-free VZV preparation was thawed and 7 dilutions in infection medium (2% FBS)

were made: $2^{-1}$ (PSGC is toxic to the cells therefore it must be diluted 1:2) and a five-fold dilution series ($5^{-1}$ to $5^{-6}$) in a total of 1 ml. Three wells of the 24-well plate were inoculated with 100 µl of each dilution. The remaining three wells were mock infected with 100 µl of infection medium. The plate was put in the incubator to allow the virus to adsorb for an hour. The inoculum was aspirated and the cells were overlaid with 0.5% carboxymethylcellulose (CMC) overlay (1x MEM, 2% heat inactivated FBS, 100 units/ml penicillin G sodium, 100 µg/ml streptomycin sulfate, 1 mM sodium pyruvate, 0.1 mM non essential amino acids, 2 mM glutamine, 3.5 g/l glucose and 0.2% w/v sodium bicarbonate). The plates were returned to the incubator for 6 days to allow plaques to develop.

After a 6 day incubation the 0.5% CMC overlay was removed and the cells were fixed with 10% neutral buffered formalin (~1 ml per well) for 30 minutes at room temperature. The formalin was removed and the fixed cells were stained with a 1% crystal violet solution (a few drops per well) at room temperature for 30 minutes. The wells were rinsed with water and allowed to dry. The plaques were then counted in each well and the average determined from the replicates for each dilution. The average plaque count was determined by multiplying the average count for each dilution by its dilution factor (so the average count for the $5^{-2}$ dilution was multiplied by $5^{2}$) and then taking the average of these corrected counts. To determine the titre, the average plaque count was divided by the inoculum volume (in this case 0.1 ml) to generate a PFU/ml.

## 2.5 VZV Nucleocapsid DNA Isolation

Nucleocapsid DNA was isolated from T-185s of VZV-Ellen infected MRC-5s in a protocol adapted from Martin *et al.* (1982). Cells showing high cytopathic effect (CPE)

were scraped from the surface, pelleted, and washed once in cold PBS. The washed and pelleted cells were suspended in 6 to 12 ml of freshly made lysis buffer [0.5% sodium deoxycholate, 0.5% NP-40, 30 mM Tris (pH 7.5), 5.0 mM magnesium acetate, 125 mM KCl, 0.5 mM EDTA, 3.6 mM $CaCl_2$, 6.0 mM β-mercaptoethanol]. The cells were disrupted by sonicating on ice with a probe sonicator at 50% duty and ten 10 second pulses. Nucleic acid was digested at 30°C by the addition of 60 μg RNaseA and three rounds of DNaseI (1000 units) at 30 minute intervals. After the third and final DNaseI treatment, the nucleocapsids were extracted by the addition of 6 ml of 1,1,2-trichloro-trifluoroethane (TCTFE) and the solution was vortexed for 2 minutes. The layers were separated by centrifugation at 800x g for 10 minutes. A 5 to 40% step gradient glycerol cushion was set up in a 13.2 ml ultracentrifugation tube by layering 3.5 ml 5% glycerol in lysis buffer on top of 3.5 ml 40% glycerol in lysis buffer. The aqueous layer of the TCTFE extraction was gently layered on top of the glycerol cushion to within 1 to 2 mm from the top of the tube. The nucleocapsids were pelleted by ultracentrifugation at 35,000 rpm in a SW41 rotor for 1 hour at 4°C. The pellet was resuspended in 2 ml of 2x STEP buffer [2% SDS, 0.1 M Tris (pH 7.5), 20 mM EDTA]. The nucleocapsids were digested by the addition of 1 ml of 2x STEP buffer containing 3 mg/ml proteinase K (Invitrogen) and incubated at 56°C for 2 hours to overnight. The DNA was extracted from the digested nucleocapsids by phenol-chloroform extraction. An equal volume of phenol:chloroform (1:1) was added and the solution was briefly vortexed. The layers were separated by centrifugation at 1300 rpm (290x g) for 2 minutes. The DNA was precipitated from the aqueous layer by the addition of one-tenth volume of 3 M sodium acetate (pH 5.6) and 3 volumes of 100% ethanol. The DNA was precipitated at -20°C for

1 hour to overnight and then pelleted at 6500x g for 5 minutes. The pelleted DNA was washed with 70% ethanol, pelleted again and allowed to air dry. The nucleocapsid DNA was purified with a Genomic Tip 500 (Qiagen) anionic exchange column according to the manufacturer's instructions for the purposes of transfecting. The DNA was eluted with 10 ml of elution buffer, pre-warmed to 50°C, and precipitated by the addition of 10 ml room temperature isopropanol. The DNA was pelleted by centrifugation at 10,000x g for 45 minutes at 4°C. The pellet was washed with 4 ml cold 70% ethanol and pelleted at 10,000x g for 20 minutes. The DNA was air dried and resuspended in 10 mM Tris (pH 8.5). The quality of the nucleocapsid DNA was determined by PCR using primers for viral open reading frame (ORF) 38 (5'-TTGAACAATCACGAACCGTT-3' and 5'-CGGGTGAACCGTATTCTGAG-3') and cellular β-globin (5'-CAACTTCATCCACGTTCACC-3' and 5'-GAAGAGCCAAGGACAGGTAC-3').

## 2.6 PCR Template DNA Extraction

Viral and cellular DNA was extracted from VZV-infected MRC-5 cells showing approximately 70% CPE using a QIAamp™ DNA Blood Mini Kit (Qiagen) following the manufacturer's instructions for cultured cells. The DNA (mix of cellular and viral) was used as template for PCR.

## 2.7 RNA Extraction

RNA was extracted from frozen VZV-Oka-infected and non-infected MRC-5 cells for the trial oligonucleotide and PCR product arrays and from T-75s of Oka-infected cells and non-infected cells for the complete gene array. In either case, the cells were scraped from the tissue culture flasks with a sterile cell scraper and pelleted at 500x g. If

cells were to be frozen, they were resuspended in approximately 3 ml of growth medium supplemented with 10% DMSO, 100 units/ml penicillin G and 100 µg/ml streptomycin and frozen at -80°C in 1 ml aliquots. If RNA was extracted immediately, the supernatant was poured off and the pelleted cells were used for the extraction. Total RNA was extracted from the cells using a Qiagen RNeasy™ kit according to the manufacturer's instructions. The cells were homogenized by passing them seven times through a 27G½ needle. The RNA was eluted from the Qiagen column twice with 30 µl of elution buffer. The absorbance at 260 nm was measured to determine the RNA concentration (ng/µl) (40 µg of single stranded RNA has an absorbance at 260 nm of 1). The RNA was either reverse transcribed immediately or stored at -80°C for later use.

## 2.8 Microarray Probe Design

### 2.8.1 VZV Gene Array

#### 2.8.1.1 PCR Products

Polymerase chain reaction (PCR) primers were designed for three VZV ORFs (10, 29, and 62) with ArrayDesigner 1.15 software (Premier Biosoft) using the default criteria and the published VZV-Dumas sequence (GenBank accession number X04370) (Davison and Scott, 1986). Two pairs of primers were designed for each ORF near the 3' ends to generate two products of different sizes, one approximately 350 base pairs (bp) and the other approximately 800 bp (Table 11). Two additional pairs of primers were also made near the 5' end of ORF 62 generating products of approximately 350 and 800 bp. The primers were made by the DNA Core Facility at the National Microbiology Laboratory (NML) in Winnipeg, Canada.

#### 2.8.1.2 Oligonucleotides

Regions in each ORF that were conserved between vaccine and parental Oka were determined from the Stanford ORF Comparisons website (http://cmgm.stanford.edu/~jjcheng/VZV/). A list of candidate sites for oligonucleotide probe positions was made based on these conserved regions for every ORF. ArrayDesigner 1.15 (Premier Biosoft) was used to design probes within 50 nucleotides on either side of these sites using the published VZV-Dumas sequence (GenBank accession number X04370) (Davison and Scott, 1986). The design settings that resulted in probes for every ORF were: length of 70 nucleotides, target Tm of 78±5°C, hairpin maximum $\Delta G$ of –5.5, self dimer maximum $\Delta G$ of –7.0, and a maximum run/repeat length of 6. Both sense and antisense probes were designed by the software and the probes that had the same sequence as the coding sequence for every ORF were selected. A BLAST (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) search was conducted on the sense (i.e. same sequence as the coding sequence) probe that ranked the highest for every ORF to ensure that no significant cross-hybridization potential existed for either VZV or cellular (human) sequences. The 73 70-mer oligonucleotide probes were synthesized by Sigma Genosys with a 5' C6 amine modification and were supplied lyophilized. The probes and their sequences are listed in Table 14.

#### 2.8.1.3 Controls

Control probes included both internal and foreign controls. Internal controls (70-mer oligonucleotides) were designed using ArrayDesigner 1.15 to detect three human housekeeping RNA species: glyceraldehyde phosphate dehydrogenase (GAPD), beta-actin and 18S ribosomal RNA (accession numbers U34995, NM_001101, and AL035413

respectively). These probes were also synthesized by Sigma Genosys with a 5' C6 amine modification. Foreign controls were purchased from Stratagene and included oligonucleotide (provided lyophilized) or PCR product probes (provided at 0.1 μg/μl in 3x SSC) for *Arabidopsis thaliana* genes which are involved in plant-specific processes. The corresponding mRNA transcripts were also purchased from Stratagene. The five genes for which probes and transcripts (1-5) were purchased are chlorophyll a/b-binding protein (accession X56062), RUBISCO activase (accession X14212), ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (accession U91966), lipid transfer protein 4 (accession AF159801), and lipid transfer protein 6 (accession AF159803).

### 2.8.2 VZV SNP Array Design

All 61 of the known single nucleotide polymorphisms (SNPs) in VZV ORFs 54 and 62 were compiled from the literature, sequence information available through NML, data from Dr. Charles Grose, and the Stanford ORF Comparisons website (http://cmgm.stanford.edu/~jjcheng/VZV/) (Table 2). Short oligonucleotides were designed to probe each of the 61 SNPs plus the *Pst*I site in ORF 38 (Takayama *et al.*, 1996) and the 3B3 site in ORF 68 (glycoprotein E) (Santos *et al.*, 1998). The polymorphic site was designed to be near the centre of the oligos. ArrayDesigner 2.0 (Premier Biosoft) software was used to design the probes using the published VZV-Dumas sequence (Davison and Scott, 1986, GenBank accession number X04370). The settings (i.e. Tm, length, hairpin ΔG, self dimer ΔG, and run/repeat length), which had to be relaxed considerably for probes to be designed for all 63 sites, were: Tm of 55±5°C, length of 15 to 30 nucleotides, hairpin maximum ΔG of –9.0, self dimer maximum ΔG of –16.0, and a maximum run/repeat length of 9. The sequences of the probes were

identical to the coding strand of the ORFs and complementary to the noncoding antisense strand. Four probes were designed for each SNP, differing only by one nucleotide (A, C, G or T) at the central position. These four probes for each SNP comprise a probe set. A few of the probe sets (19) overlapped another nearby SNP. In these cases, another probe set was made which contained the documented mutation for the neighbouring SNP. The probes (a total of 352) were synthesized by Qiagen Operon with a C6 amine group on the 5' end to serve as both a spacer and for attachment. The probes were desalted and received lyophilized. As well, replicate probes for the *Pst*I, *Bgl*I (ORF 54), *Sma*I (ORF 62) and 3B3 sites were manufactured by Sigma Genosys in the same manner. The complete list of the probes and their sequences are found in Table 15.

## 2.9 Generation of PCR Product Probes

### 2.9.1 VZV Probe PCR

PCR was performed to amplify segments of viral ORFs 10, 29 and 62 from DNA extracted from VZV-Oka vaccine infected MRC-5 cells. The reactions were set up in 100 µl volumes with a final concentration of 1x PCR buffer [20 mM Tris·HCl (pH 8.4) 50 mM KCl] (Invitrogen), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 10 µM primer #1, 10 µM primer #2, 5 units of Platinum™ Taq DNA polymerase (Invitrogen) and approximately 200 ng template DNA. Reactions lacking template DNA were also set up with each primer pair. The primer pairs and resulting product size are listed in Table 11. The reactions were amplified in a PTC-200 Thermal Cycler DNA Engine (MJ Research). Reactions using primer pairs 1, 2, 3, 4, 5, and 6 were run with the following conditions: initial denaturation at 96°C for 3 minutes, 40 cycles of 95°C denaturation for 30 seconds,

54.1°C annealing for 30 seconds, 72°C extension for 1 minute and a final extension at 72°C for 5 minutes. Reactions using primer pairs 7, 8 and 9 were run with an initial denaturation at 96°C for 3 minutes, 40 cycles of 95°C for 30 seconds, 60°C annealing for 30 seconds, 72°C for 1 minute, and a final extension at 72°C for 5 minutes. A 10 µl aliquot of each reaction electrophoresed onto a 2% agarose gel made with 1x TAE buffer (40 mM Tris·acetate, 1 mM EDTA) to determine the quality of the PCR reaction. If only the correct band was present, the remainder of the PCR reaction was purified (primer pairs 1 – 6). If more than one band was observed, the appropriately sized band was gel purified (primer pairs 7 – 9).

### 2.9.2 PCR Purification

The PCR products from the reactions using primer pairs 1, 2, 3, 4, 5, and 6 were purified to remove primers, unincorporated deoxynucleotides, enzyme, and buffers using a QIAquick™ PCR Purification Kit (Qiagen). The DNA was eluted from the QIAquick™ column with 50 µl of Buffer EB (10 mM Tris·HCl, pH 8.5). The concentration of the purified PCR product was determined by measuring the absorbance at 260 nm of 3 µl of the DNA in a final volume of 300 µl of water. To calculate the concentration of the purified DNA (ng/µl), the absorbance of the diluted sample was multiplied by the dilution factor (100) and then by 50 (50 µg of double stranded DNA has an absorbance at 260 nm of 1).

### 2.9.3 Gel Purification

The PCR reactions using primer pairs 7, 8 and 9 contained more than one product, thus, the correct product (704, 351, and 303 bp respectively) was gel purified. The

remainder of the 100 µl PCR reaction was mixed with 25 µl of loading buffer and loaded into a double well (2 wells produced by taping together 2 teeth of a comb) of a 2% low melting agarose gel in TAE buffer. The gel was run at 70 volts for 1 hour and 45 minutes after which time the DNA bands were visualized using a UV transilluminator (VWR). The band of the appropriate size was cut out with a scalpel and the DNA was extracted from the gel slice using a QIAquick™ Gel Extraction Kit (Qiagen) following the manufacturer's instructions for use with a vacuum manifold. The concentration of the purified DNA was determined by reading the absorbance of a 1 in 100 dilution at 260 nm.

## 2.10 Microarray Probe Preparation

### 2.10.1 VZV Gene Array Probes

#### 2.10.1.1 PCR Product Probes

Purified PCR products for the three viral ORFs were lyophilized in a Speed Vac™ (Savant) at medium speed for approximately 20 to 30 minutes. They were resuspended to a final concentration of 0.3 µg/µl in 1.5 M Betaine (Sigma) in 3x saline sodium citrate (SSC). A 5 µl aliquot was transferred to a 384 well plate to be subsequently used for spotting slides. Two *A. thaliana* PCR products from Stratagene were included in the plate at a concentration of 0.1 µg/µl in 3x SSC as were two wells which contained only the printing buffer (1.5 M Betaine in 3x SSC) to serve as negative controls.

#### 2.10.1.2 Oligonucleotide Probes

The lyophilized viral and control (internal and foreign) oligonucleotide probes were resuspended to a final concentration of 0.1 μg/μl in NoAb BioDiscoveries Print Buffer (150 mM sodium phosphate, pH 8.5). A 5 μl aliquot of each was transferred to a 384 well plate. For the preliminary trial array (3 VZV probes), two *A. thaliana* PCR products (Stratagene) were included as positive controls (0.1 μg/μl in 3x SSC) and 2 wells with only print buffer as negative controls. The final array contained probes for the full viral gene complement (73 probes), 3 human housekeeping genes, 5 oligonucleotide probes for *A. thaliana* (Stratagene) and 19 wells of print buffer.

### 2.10.2 VZV SNP Array Probes

The 368 SNP probes were resuspended in NoAb print buffer (150 mM sodium phosphate, pH 8.5) to a final concentration of 100 μM. A Biomek™ 2000 liquid handling robot (Beckman) transferred 25 μl aliquots of the probes to four 96-well plates. Four wells in each of the plates were filled with 25 μl of print buffer or 1 μg/μl of sheared salmon sperm DNA (Invitrogen) in print buffer. The robot then added 75 μl of print buffer and mixed the samples to give a final printing concentration of 25 μM. A 10 μl aliquot from each well of the four 96-well plates was transferred to 384-well plates to serve later as printing plates (each 384-well plate contained the entire probe complement). The plates were sealed with Beckman foil seals (Biomek™ seal and sample aluminium foil lids) and stored at -20°C until used.

## 2.11 Array Printing

For printing, one of the 384-well printing plates containing the PCR products,

gene array oligos (trial or full sized) or SNP probes, was thawed, vortexed and spun down. A Virtek ChipWriterPro™ robot and Telechem SMP3 pins were used to print the arrays in a dust free humidified (50%) chamber (Virtek). One pin was used to print the PCR product gene array onto CMT™-GAPS (gamma amino propyl silane) coated slides (Corning) or the oligo probes onto hydrogel aldehyde slides (NoAb Biodiscoveries). Four pins (2x2) were used to print the SNP array onto hydrogel aldehyde coated slides. The pins drew up 0.25 µl of solution, blotted five times on a blotting slide and deposited 0.7 nl per spot with a 100 µm diameter. The distance between spots in both x and y axes was set at 300 µm for the SNP and complete gene arrays and 200 µm for the trial gene arrays (PCR products and oligos). Five replicate spots were printed side by side for the trial gene array yielding grid dimensions of 13x5 (PCR products) and 7x5 (oligos). The samples for the full sized gene array were printed in a 10x10 grid, which was repeated 4 times on the slide for a total of 4 replicate grids (4.5 mm between them) each of which contained the full probe complement. For the SNP array, each of the four pins printed a 12x8 grid containing 96 spots representing one quarter of the probes. The full probe complement was printed in 4 neighbouring 12x8 grids (2x2). These 4 grids were repeated in quadruplicate across the slide (2x2) for a total of 4 replicate spots for each probe. After each sample spotting, the pins were washed in nuclease free water (Ambion) for 2 seconds and then vacuum dried (1 second) before picking up and spotting the next sample(s). Upon completion of printing, the plate was removed and sealed up for storage and the humidifier turned off.

### 2.12 Post Printing: Attachment of Probes to Slides

PCR product probes were immobilized onto CMT™-GAPS slides by incubating

at 80°C for 3 hours. The hydrogel aldehyde slides were allowed to dry overnight in the chamber to facilitate covalent attachment of the amine-modified oligonucleotide probes to the aldehyde groups on the slide surface through a Schiff's base reaction. After probes were immobilized to both slide types, the slides were visually inspected, the outlines of the array were etched into the slide using a diamond scriber (Sigma), and they were stored dessicated at room temperature.

## 2.13 Generation of Labelled cDNA Targets

This section along with section 2.15 Array Hybridization is summarized in the flow chart presented in Fig 5.

### 2.13.1 Reverse Transcription

Ten μg of RNA (or less if the yield was low) was reverse transcribed into cDNA. The volume was reduced either by concentrating to less than or equal to 16 μl with a speed vac on medium for about 15 minutes or by ethanol precipitation if the entire 60 μl RNA extraction was required for reverse transcription (RT). Ethanol precipitation was performed according to Miles Wilkinson (1991). One tenth the volume of 3 M sodium acetate (pH 5.2, pretreated with 0.15% DEPC) was added followed by 2 volumes of 100% ethanol. The solution was briefly vortexed and the RNA precipitated at -20°C for approximately 30 minutes. The RNA was then pelleted at maximum speed in a microfuge (20800x g) for 10 minutes at room temperature. The supernatant was carefully removed, the RNA pellet washed with 1 ml of 80% ethanol and pelleted again at maximum speed for 1 minute at room temperature. The supernatant was decanted and the RNA pellet air dried for 15 minutes. Commercially prepared *A. thaliana* transcripts

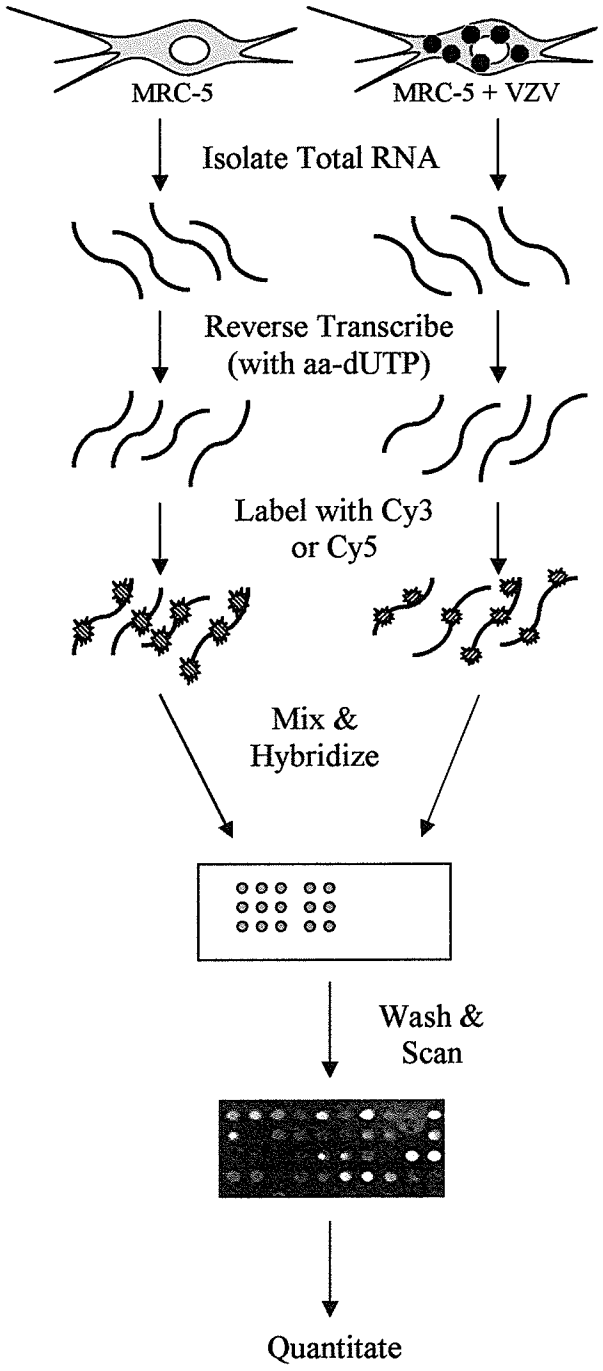



**Figure 5.** Flow chart of methods used for the VZV gene expression array. Details are given in the text and the protocols are adapted from Standard Operating Procedures from The Institute for Genomic Research (TIGR).

(Stratagene), corresponding to probes spotted on the array, were added to the concentrated or pelleted RNA. As little as 0.3 ng to as much as 1 ng of the transcripts were spiked into the reverse transcription reaction. Reverse transcription was performed according to a protocol from The Institute for Genomic Research (TIGR) (Hasseman, 2002a). Six μg of random hexamer primers (Invitrogen) were added to the 10 μg (or less) of RNA and the volume brought up to 18.5 μl with DNase and RNase free water (Invitrogen). The solution was mixed thoroughly and incubated at 70°C for 10 minutes and then centrifuged briefly in a microfuge. The RNA was reverse transcribed in a reaction containing 1x First Strand Buffer [50mM Tris·HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$] (Invitrogen), 0.01 M DTT (Invitrogen), 0.5 mM dATP (Invitrogen), 0.5 mM dCTP (Invitrogen), 0.5 mM dGTP (Invitrogen), 0.3 mM dTTP (Invitrogen), 0.2 mM aminoallyl-dUTP (aa-dUTP) (Sigma), and 400 units SuperScript II$^{TM}$ RNase H$^-$ Reverse Transcriptase (Invitrogen). The reaction was mixed and incubated at 42°C overnight.

### 2.13.2 Indirect Labelling of Target cDNA

#### 2.13.2.1 Hydrolysis of RNA Template

The RNA template was hydrolysed according to the TIGR protocol (Hasseman, 2002a). Sodium hydroxide and EDTA were added to the RT reaction to a final concentration of 0.2 M and 0.1 M respectively. The hydrolysis reaction was mixed and incubated at 65°C for 15 minutes after which time the cDNA was purified.

#### 2.13.2.2 cDNA Purification

Unincorporated aa-dUTP and free amines were removed from the reaction using

Microcon YM-30 columns (Millipore). The volume was brought up to 500 µl with 4 mM phosphate buffer ($KPO_4$, pH 8.5) and was applied to the ultrafiltration column. The column and collection tube were spun at 13,600x g in a microfuge for 12 minutes. The eluate was discarded and the column inverted into a new collection tube. The cDNA was eluted in the retained buffer by centrifuging at 1000x g for 3 minutes.

#### 2.13.2.3 Coupling aa-cDNA to Cy Dye Ester

The purified cDNA was dried in a speed vac set on medium and subsequently resuspended in 4.5 µl of 0.1 M sodium carbonate buffer ($Na_2CO_3$, pH 9.0, made fresh every month) (Hasseman, 2002a). One lyophilized aliquot of mono-reactive Cy3 (100 µg) or Cy5 (62.5 µg) ester was resuspended in 4.5 µl of DMSO and added to the cDNA. The tubes were wrapped in foil and the reaction was incubated for 1 hour at room temperature in the dark. The labelling reaction was stopped with the addition of 35 µl of 100 mM sodium acetate (pH 5.2).

#### 2.13.2.4 Reaction Purification

Uncoupled Cy dye was removed from the labelled cDNA using Microcon YM-30 ultrafiltration columns as previously described. A sample was removed for spectrophotometry. The amount of cDNA and label (Cy3 or Cy5) was determined by measuring the absorbance at 260 nm and 550 nm for Cy3, or 650 nm for Cy5 (Hasseman, 2002a). The concentration of the cDNA was determined by multiplying the absorbance at 260 nm by the dilution and then by 33 (33 ng of single stranded DNA has an absorbance of 1 at 260 nm). The amount of Cy3 (picomoles) incorporated was determined by multiplying the absorbance at 550 nm by both the dilution and the volume

and then dividing by 0.15. The amount of Cy5 (picomoles) incorporated was determined by multiplying the absorbance at 650 nm by both the dilution and the volume and then dividing by 0.25.

## 2.14 Generation of Labelled RNA Targets

Levente Bodrossy's protocol for generating labelled RNA targets (available online at http://www.diagnostic-arrays.com/) was followed with some minor changes (indirect corporation of Cy label rather than direct for example) (Bodrossy, in press). This section along with section 2.15 Array Hybridization is summarized in the flow chart presented in Fig 6.

### 2.14.1 PCR Amplification of Targets

PCR primers were designed (PrimerSelect software in the DNA Star suite) to amplify the four regions of interest (*Pst*I in ORF 38, ORF 54, ORF 62, and 3B3 in ORF 68) and contained a T7 promoter (5'-TAATACGACTCACTATAG – unique primer sequences-3') at the 5' end of the reverse primers (Table 12). The primers were synthesized by the DNA Core Facility (NML). One hundred µl PCR reactions were set up for all 4 primer pairs as previously described using Platinum$^{TM}$ Taq (Invitrogen). Two µl of DNA (approximately 200 ng) extracted from virally infected cells was used as the template. The cycling parameters were as follows. Reactions containing the *Pst*I primers were initially denatured and the enzyme activated at 96°C for 3 minutes, and amplified by 30 cycles of 95°C for 30 seconds, 50°C annealing for 30 seconds and a 1 minute extension at 72°C followed by a final extension of 8 minutes at 72°C. Reactions containing the ORF 54 primers, were initially denatured at 96°C for 3 minutes and

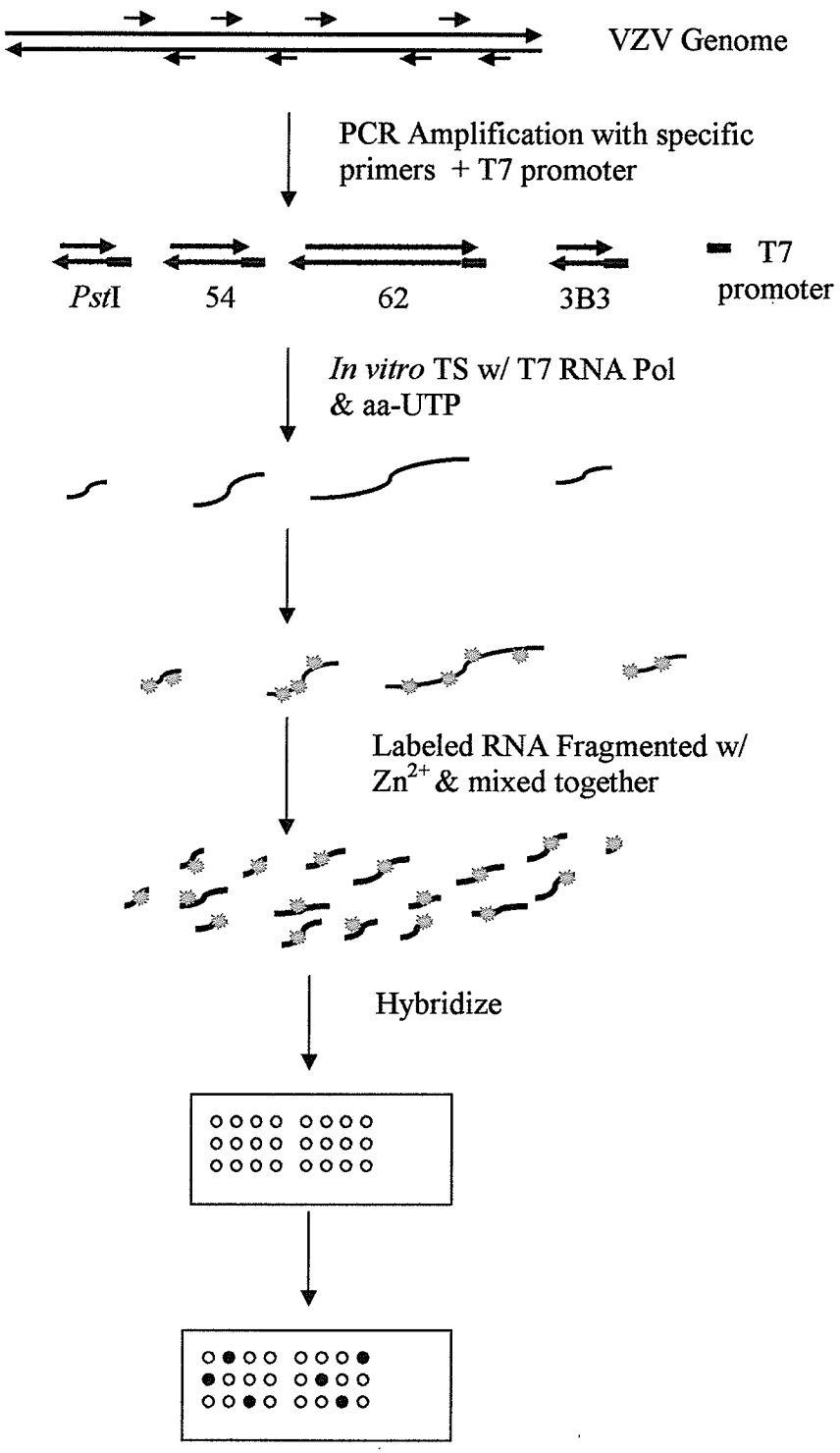



**Figure 6.** Flow chart of methods used for the VZV genotyping array. Details are given in the text and the protocols are adapted from www.diagnostic-arrays.com and Standard Operating Procedures from The Institute for Genomic Research (TIGR).

amplified by 30 cycles of 95°C for 30 seconds, 55°C annealing for 30 seconds, and extension at 72°C for 3 minutes 30 seconds followed by a final extension of 15 minutes at 72°C. Reactions containing the ORF 62 primers were denatured at 96°C for 5 minutes and amplified by 30 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 5 minutes followed by a final extension of 72°C for 20 minutes. PCR reactions with 3B3 primers were set up with the following cycling parameters: 96°C for 3 minutes, followed by 30 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and a 1 minute extension at 72°C and a final extension for 8 minutes at 72°C. Ten µl of the PCR reactions were electrophoresed through an agarose gel to determine product size. The products were purified with a QIAquick™ PCR Purification Kit (Qiagen) according to the manufacturer's instructions.

### 2.14.2 *In Vitro* Transcription of PCR Products

Four hundred ng of purified PCR product was *in vitro* transcribed with T7 RNA polymerase in a 20 µl reaction containing 1x T7 RNA polymerase buffer [0.04 M Tris·HCl (pH 8.0), 8 mM $MgCl_2$, 2 mM spermidine-$(HCl)_3$, and 25 mM NaCl] (Invitrogen), 10 mM DTT (Invitrogen), 0.5 mM ATP (Invitrogen), 0.5 mM CTP (Invitrogen), 0.5 mM GTP (Invitrogen), 0.25 mM UTP (Invitrogen), 0.25 mM amino-allyl-UTP (Ambion), 20 units RNasin (Promega), and 40 units T7 RNA polymerase (Invitrogen). The reactions were incubated in RNase free tubes for 4 hours at 37°C.

### 2.14.3 RNA Purification

Unincorporated nucleotides, enzymes, salts, and DNA template were removed with a Qiagen RNeasy™ Mini Kit according to the manufacturer's protocol with the

exception of the second wash which was performed with 80% ethanol in nuclease free water (Invitrogen) instead of Buffer RPE (proprietary composition) to remove any residual free amines. The RNA yield was determined by measuring its absorbance at 260 nm (40 μg of single stranded RNA has an absorbance at 260 nm of 1). It was then dried to completion in a speedvac on low for 30 minutes.

### 2.14.4 Labelling of RNA Target

Labelling of the RNA was done exactly as described previously for cDNA (2.11.2.3) except special care was taken to ensure that reagents were RNase free. After the labelling reaction, 35 μl of 100 mM sodium acetate (pH 5.2, pretreated with 0.15% DEPC) was added and the volume was brought up to 500 μl with DNase, RNase free water (Invitrogen). The labelled RNA was purified with Microcon YM-30 columns (Millipore) according to the manufacturer's instructions and as described previously for cDNA (2.11.2.2). The volume of the eluted RNA was brought up to 50 μl with DNase, RNase free water (Invitrogen).

### 2.14.5 $Zn^{2+}$ Fragmentation of RNA

The labelled RNA was fragmented to an average length of 50 nucleotides by zinc treatment. The purified, labelled RNA was mixed with Tris·HCl (pH 7.4, Ambion) and $ZnSO_4$ (pretreated with 0.15% DEPC) to final concentrations of 25 mM and 10 mM, respectively. The reaction was incubated at 60°C for 30 minutes, stopped with the addition of 1.43 μl of 500 mM EDTA (pH 8.0, Ambion) and put on ice for 1 minute. The four RNA species for a particular viral strain (PstI, ORF54, ORF62, and 3B3), labelled with either Cy3 or Cy5, were mixed together with the four differentially labelled RNAs

for another strain. RNasin (80 units) (Promega) was added to the fragmented RNA. The mixed RNA was split equally into 4 tubes to be dried down and hybridized to 4 replicate arrays as described below.

## 2.15 Array Hybridization

The TIGR protocol for array hybridization (Hasseman, 2002b) was followed with a few modifications.

### 2.15.1 Prehybridization

The printed slides were pre-hybridized in freshly made prehybridization buffer (5x SSC, 0.1% SDS, 1% BSA, pre-warmed to 42°C) for 45 minutes at 42°C. The slides were washed in three Coplin jars as follows: 1) completely submerged in MilliQ water four or five times and then the array portion only submerged in MilliQ water five times, 2) array only submerged in MilliQ water, and 3) submerged completely in isoproponal. The slides were then dried by centrifuging in a slide rack with a microtitre plate rotor in a table top centrifuge at 500 rpm (28x g) for 5 minutes. The appearance of the slides was noted and the wash steps repeated if necessary. The slides were hybridized immediately following prehybridization.

### 2.15.2 Hybridization

The differentially labelled targets (cDNA or RNA) to be hybridized to one array were mixed together and dried to completion in a speed vac on medium. The dried labelled cDNA or RNA target was resuspended in 7.5 or 20 µl, respectively, of hybridization buffer [DIG Easy Hyb Buffer (Roche) plus 0.5 µg/µl calf thymus DNA (Sigma) pre-warmed to 42°C]. The cDNA target was denatured at 95°C for 3 minutes,

snap cooled on ice for 30 seconds, and then centrifuged at maximum speed (20,800x g) for 1 minute. The RNA target was denatured at 65°C for 15 minutes.

#### 2.15.2.1 VZV Gene Array

A prehybridized slide was placed in the bottom half of a single slide hybridization chamber (Corning) and a hydrophobic coverslip (Grace Biolabs) was cut to size using the etched marks on the slide as a guide. The labelled cDNA was applied to the array near one end with a pipette and the coverslip was gradually lowered onto the slide with forceps starting at the edge where the target had been applied. Any large bubbles trapped under the coverslip were gently worked out by tapping the coverslip. Ten µl of DIG Easy Hyb buffer (without calf thymus DNA) was added to the wells of the chamber on either side of the slide. The top half of the chamber was placed and it was sealed. The assembled chamber was wrapped in foil and placed in a 42°C water bath to hybridize overnight.

#### 2.15.2.2 SNP Array

A dusted 22x25 mSeries LifterSlip™ (Erie Scientific) was placed on top of the prehybridized array using the etch marks as a guide. The labelled target was applied slowly with a pipette at one corner of the slip and then the diagonal corner. The solution was drawn under the slip and spread out across the array via capillary action. The slide was placed in a Genetix hybridization chamber and 1 ml of DIG Easy Hyb buffer (without calf thymus DNA) was added to the well of the chamber. The chamber was sealed and placed in a 42°C hybridization oven (DiaMed) overnight. Four replicate arrays were performed in parallel for each viral strain.

### 2.15.3 Post Hybridization Washes

All wash steps were performed in light tight slide wash boxes (DiaMed). The hybridization chamber was taken apart and the hybridized slides immediately placed in warm (42°C), low stringency wash buffer (1x SSC, 0.2% SDS) to remove the coverslips. The slides were washed for 4 minutes on a shaker in the warm, low stringency wash buffer. They were then washed with agitation an additional 4 minutes in high stringency wash buffer (0.1x SSC, 0.2% SDS). All traces of SDS were removed by two successive washes in 0.1x SSC for 2.5 minutes each on a shaker. The slides were dipped in MilliQ water several times and spun dry in a table top centrifuge at 500 rpm for 5 minutes. They were stored in a light tight slide box (Fisher) prior to being scanned.

## 2.16 Slide Scanning and Image Acquisition

Hybridized arrays were scanned with a Virtek ChipReader$^{TM}$ 2.0 with a gain of 1, 1 scan, scan speed of 25, and pixel (micron) resolution of 10. The Cy3 and Cy5 laser power and detector sensitivity were adjusted as necessary to decrease saturation and optimize signal and the images were saved as tif files. The black and white raw images were imported into Array-Pro$^{TM}$ 4.0 or 4.5 (SNP arrays) software (Media Cybernetics) and the Cy3 and Cy5 images were assigned a green and red colour respectively. Grids, consisting of circles (12 by 12 pixels) that surrounded each spot, were placed on the image to define where the intensity data was extracted from. The images were scanned by eye and any spots that were in areas of high background (e.g. near a piece of dust) were flagged and removed from further analysis. The signal from each spot was corrected for background by subtracting the mean signal of the corners of an imaginary square surrounding the spot (local corners) from the density of the spot signal.

## 2.17 Data Analysis

Background corrected signal intensity data for the all arrays was imported into Microsoft Excel where the data from the four replicate spots per array and replicate arrays was compiled and analyzed.

For the gene expression array, the average net intensity was determined from the replicate spots. The averages for all unique spots (excluding the *A. thaliana* spots because they were not always hybridized in the same amounts) were totalled for the channel that had been hybridized with the VZV cDNA. The net intensity of each probe (average of replicate spots) was normalized to this total signal for the slide by dividing each by the total and expressing this relative signal intensity as a percentage. The data from the replicate arrays was then combined by taking the average of these percentages for each probe.

For the genotyping array, for each SNP site, each probe was expressed as the proportion of the total for the probe set (e.g. net signal for A probe divided by sum of net signal for A, C, G and T probes for that site, etc) for every replicate spot ($n \leq 16$). One-tailed t-tests (unequal variances) were performed on the two probes with the highest proportion of the signal for the probe set for every SNP and the p-value was reported. The difference between the two probes with the highest signal intensity, for every SNP, was considered statistically significant if the p-value was less than 0.05, and the identity of the base at that particular SNP was called as the one with the highest intensity.

## 2.18 IE62 Cloning

### 2.18.1 PCR Amplification

Primers were designed to amplify the immediate early (IE) 62 ORF (3933 bp) and

incorporated *Eco*RI and *Xba*I restriction sites at their 5' ends for directional cloning (DNA Core Facility at the NML) (Fig 15). The 50 µl PCR reaction was set up as follows: 1x *Pfx* Amplification Buffer (Invitrogen, proprietary composition), 1 mM $MgSO_4$, 0.3 mM dNTPS, 0.3 µM forward primer (5'-AGATAGAATTCGTACGTCTAAATTCACCCCAGTGC-3'), 0.3 µM reverse primer (5'- GATATTCTAGACGCCAGTGGCGCTCACG-3'), 2.5 units Platinum™ *Pfx* DNA polymerase and 180 ng of DNA extracted from VZV-Ellen (5th passage) infected MRC-5s. The reaction was carried out in a PTC-200 Thermal Cycler DNA Engine (MJ Research) with the following conditions: initial denaturation at 94°C for 5 minutes; 30 cycles of 94°C for 30 seconds, 55°C annealing for 30 seconds, and 68°C extension for 10 minutes; followed by a final extension at 68°C for 20 minutes. Five µl of the PCR reaction was electrophoresed through a 0.8% agarose gel to check for quality.

### 2.18.2 Purification of Amplicon

The remainder of the reaction (45 µl) was gel purified from a 0.8% agarose gel. The band was visualized with a longwave UV light (to prevent damage to the DNA) and excised. The 4 kb amplicon was extracted from the gel slice using QIAquick™ Gel Extraction Kit (Qiagen) as previously described. The concentration of the purified DNA was determined by measuring its absorbance at 260 nm.

### 2.18.3 Restriction Endonuclease Digestion of Insert and Vector DNA

All of the purified amplicon DNA (~2 µg) and 4 µg of vector pTracer™ –CMV2 (Invitrogen) DNA were digested with restriction enzymes *Eco*RI and *Xba*I; insert in a final volume of 40 µl and vector in 100 µl. The reaction contained the following

components: 1x REact™ 2 buffer [50 mM Tris·HCl (pH 8.0), 10 mM $MgCl_2$, and 50 mM NaCl] (Invitrogen), and 10 units each of *Eco*RI (Invitrogen) and *Xba*I (Invitrogen). The reactions were incubated at 37°C for 2 to 4 hours. The digested DNA was precipitated for 1 hour at -20°C with one tenth the volume of 3 M sodium acetate (pH 5.6) and 2.5 volumes of cold 100% ethanol. The DNA was pelleted by centrifugation at 13,000 rpm (17,900x g) for 5 minutes, washed with 70% ethanol and air dried. The insert and vector DNA pellets were resuspended in 22.5 and 50 µl 10 mM Tris·HCl (pH 8.5), respectively and gel purified using longwave UV and a QIAquick™ Gel Extraction Kit (Qiagen). The insert DNA was eluted in 30 µl of warm Buffer EB (10 mM Tris·HCl, pH 8.5) and the vector in 50 µl.

### 2.18.4 Dephosphorylation of Digested Vector DNA

The 5' phosphates of the vector DNA were removed by treatment with calf intestinal phosphatase (CIP) (Roche). Five µl of 10x CIP buffer (50 mM Tris·HCl, 0.1 mM EDTA, pH 8.5) (Roche) and 0.2 µl of CIP (20 units/µl, Roche) were added to the vector. The reaction was incubated at 37°C for 1 hour. The vector DNA was extracted with an equal volume of phenol:chloroform (1:1). The mixture was briefly vortexed and the layers separated by centrifugation at 13,000 rpm (17,900x g) for 2 minutes. The aqueous layer was transferred to a new 1.5 ml microfuge tube and the DNA was ethanol precipitated for 1 hour at –20°C as previously described. The vector DNA pellet was resuspended in 40 µl 10 mM Tris·HCl (pH 8.5) and quantified by absorbance at 260 nm.

### 2.18.5 Ligation of Insert and Vector DNA

The digested IE62 insert was ligated with the digested and dephosphorylated

vector pTracer-CMV2 in a 3 to 1 ratio by weight (μg). Briefly, 450 μg of insert was mixed with 150 μg vector and 1x Ligase buffer [50 mM Tris·HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM ATP, 1 mM DTT, 5% (w/v) polyethylene glycol 8000] (Invitrogen), 0.5 mM ATP (Invitrogen), and 1 unit T4 DNA Ligase (Invitrogen). The reaction was incubated at 15°C overnight. The recombinant plasmid was precipitated at -20°C for 6 hours with the addition of one-tenth volume 3 M sodium acetate (pH 5.6) and 3 volumes of cold 100% ethanol. The DNA was pelleted by centrifugation at 17,900x g for 5 minutes and washed with 70% ethanol. The pellet was briefly air dried and then resuspended in 20 μl of water.

### 2.18.6 Electroporation of Electrocompetent *E. coli* DH5α Cells

ElectroMAX™ DH5α-E™ *E. coli* cells were transformed by electroporation with the recombinant plasmid. One vial of the cells was thawed on ice. Two and 4 μl aliquots of the recombinant plasmid were each mixed with 40 μl aliquots of the cells in sterile prechilled 1.5 ml microfuge tubes placed on ice. The mixture was transferred to ice-cold electroporation cuvettes (BioRad Gene Pulser™ II Cuvettes, 0.1 cm electrode gap). The cells were electroporated with a Bio-Rad Gene Pulser™ II Electroporator set to 1.25 kV with a 25 μF capacitor and a pulse controller set to 400 Ω. Immediately following the pulse 1 ml of warm SOC medium (Sigma) was added to the cells. The cells were transferred to a 5 ml culture tube and incubated at 37°C for 15 minutes. Ten and 100 μl aliquots were spread (in a total volume of 100 μl LB) onto LB plates (LB medium, 15 g/l Bacto-agar) supplemented with carbenicillin (50 μg/ml). The plates were incubated overnight at 37°C and single colonies were picked the following day for screening.

### 2.18.7 Colony Screening and Verification of Insert

Twenty four single colonies were picked and screened for the insert by PCR using lysed cells. Briefly, a single colony was suspended in 70 µl of water, boiled for 1 minute to lyse the cells and centrifuged at maximum speed (1810x g) in a table top centrifuge with a microtitre plate rotor for 10 minutes to pellet the cell debris. Two µl of this lysed cell suspension was used as a template in a 50 µl PCR reaction using primers that amplify approximately 500 bp of the insert (5'-ATTACTGTCGACCCGAGACC-3' and 5'-AGGTTGGCAAACGCAGTC-3'). Twenty two colonies were positive for the insert by PCR. Five were selected and each inoculated into two 2 ml LB plus carbenicillin (25 µg/ml) cultures and incubated at 37°C with shaking (200 rpm) overnight. Plasmid DNA was purified using a QIAprep™ Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. The presence of an insert was confirmed by digestion of plasmid DNA with *Eco*RI and *Xba*I restriction enzymes. The purified plasmids were sequenced (DNA Core Facility of the NML) using 16 primers staggered every 500 bases on both strands of the insert and 2 vector primers on either side of the multiple cloning site (Table 13).

### 2.18.8 Large-Scale Plasmid DNA Isolation for Transfection

Clones of the vector (pTracer) and the vector with the insert (pTracerIE62) were spread onto LB plates supplemented with carbenicillin (50 µg/ml) and incubated overnight at 37°C. The following day a single colony from each was picked, inoculated into 2 ml of LB supplemented with 50 µg carbenicillin and incubated for 9 hours at 37°C. A 250 µl aliquot was used to inoculate 100 ml of LB plus carbenicillin (25 µg/ml) in a

500 ml flask. These cultures were incubated at 37°C with shaking (200 rpm) overnight. The following day, the plasmids were purified from the 100 ml cultures using a QIAfilter Plasmid Maxi Kit (Qiagen) following the manufacturer's protocol.

### 2.19 Transfection

MeWo ($8 \times 10^4$ per well) and MRC-5 ($6 \times 10^4$ per well) cells were seeded in 4 wells each per slide in three 8-well chamber slides (VWR) the day before for ~75 (MeWo)-90% (MRC-5) confluency at the time of transfection. The three slides were set up the same way: the top 4 wells were seeded with MRC-5s and the bottom 4 with MeWos, but differed in the amount of DNA they received, either 1x (19 ng), 2x (38 ng) or 3x (57 ng). The first well of each received only transfection medium (Opti-MEM™ I Medium from Invitrogen), the second well only sheared salmon sperm DNA (always 381 ng), the third well received only pTracer, and the last well of each received equal amounts of pTracerIE62 and VZV-Ellen nucleocapsid DNA. The cells were transfected with Lipofectamine™ 2000 reagent from Invitrogen according to the manufacturer's instructions for transient transfection of adherent cells in a surface area of 0.8 $cm^2$. Briefly, the DNA was diluted to a final volume of 20 µl with Opti-MEM™ I. The DNA used was as follows: 381 ng sheared salmon sperm DNA; 19 ng (1x), 38 ng (2x) or 57 ng (3x) pTracer; 19 ng (1x), 38 ng (2x) or 57 ng (3x) pTracerIE62 + 19 ng (1x), 38 ng (2x) or 57 ng (3x) VZV-Ellen DNA. The amount of DNA in each well (except for the mock-transfected wells) was kept constant at 381 ng with sheared salmon sperm DNA. For each well, 1.14 µl of Lipofectamine™ 2000 reagent was diluted with 18.86 µl Opti-MEM™ I to give a DNA to transfection reagent ratio of 1:3 and it was incubated at room temperature for 5 minutes. Equal volumes of the diluted transfection reagent were mixed

with the diluted DNA and incubated at room temperature for 20 minutes to allow the DNA-Lipofectamine™ 2000 complexes to form. After the incubation, 115.5 µl of Opti-MEM™ I was added to the DNA complexes. The growth medium was removed from each well of the chamber slides and 150 µl of the diluted DNA-Lipofectamine™ 2000 complexes was added to each well. The slides were mixed gently by rocking and returned to the incubator for 4-5 hours after which time 150 µl growth medium with 20% FBS was added to each well for a final concentration of 10% FBS. The cells were assayed for the presence of green fluorescent protein 2 days post transfection by confocal microscopy.

### 2.20 Confocal Microscopy

Two days post transfection the medium was aspirated from the wells of the chamber slides, the chamber portion and gasket of the slides removed and the cells air dried for approximately 30-60 minutes. The cells were fixed in ice-cold methanol for 6 minutes and rehydrated in two five minute changes of PBS. A couple of drops of glycerol were placed on the slides and coverslips mounted. The success of the transfection was determined by detecting the presence of green fluorescent protein with an Olympus IX70 confocal microscope using an argon laser with emission filter set at 488 nm. The transfection efficiency was estimated from the proportion of green cells detected.

# 3.0 RESULTS

## 3.1 VZV Gene Array

### 3.1.1 Array Design: Oligos or PCR Products

Arrays were made with either 70-mer oligonucleotide or PCR product (800 and 350 bp) probes for viral ORFs 10, 29 and 62. RNA extracted from VZV-Oka infected MRC-5s was reverse transcribed and labelled with Cy5. The labelled cDNA was divided equally onto the two array types and hybridized overnight. The hybridized slides were washed and scanned and the images compared (Fig 7). Hybridization to the *Arabidopsis thaliana* positive control spots (PCR products on both slides) resulted in long tails on the PCR product array (Fig 7b) but on the oligo array (Fig 7a), these spots were much more distinct without any tails. The negative control spots directly above the positive controls had a detectable fluorescence on the PCR product array but this was not apparent on the slide spotted with oligo probes. In addition, the PCR product probes for the viral ORFs did not appear to hybridize well as indicated by the low signal from these spots. In contrast, there was a readily detectable signal from the oligo viral probes indicating hybridization. In general, the oligo probes appeared to have cleaner brighter signals than the PCR products. Thus the array for the entire viral gene complement was made using oligos rather than PCR products as probes.

### 3.1.2 Evaluation of VZV Gene Array

Oligo probes were designed to detect transcripts from viral ORFs, including spliced transcripts from ORFs S/L and 42-45. They were designed to be 70 nucleotides in length, have a similar melting temperature of about 78°C, and were checked for

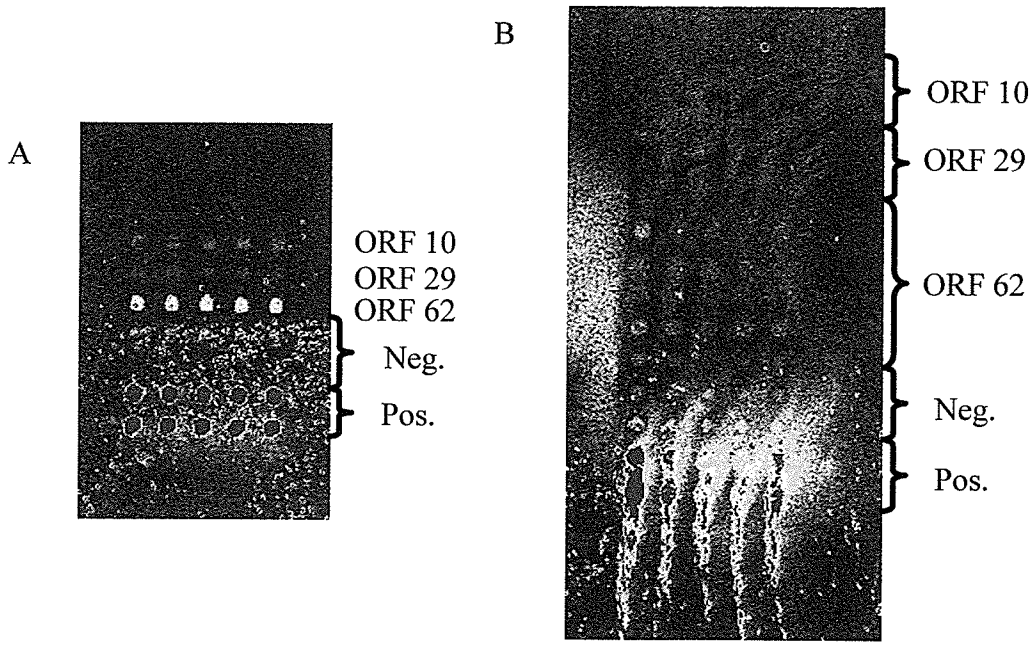
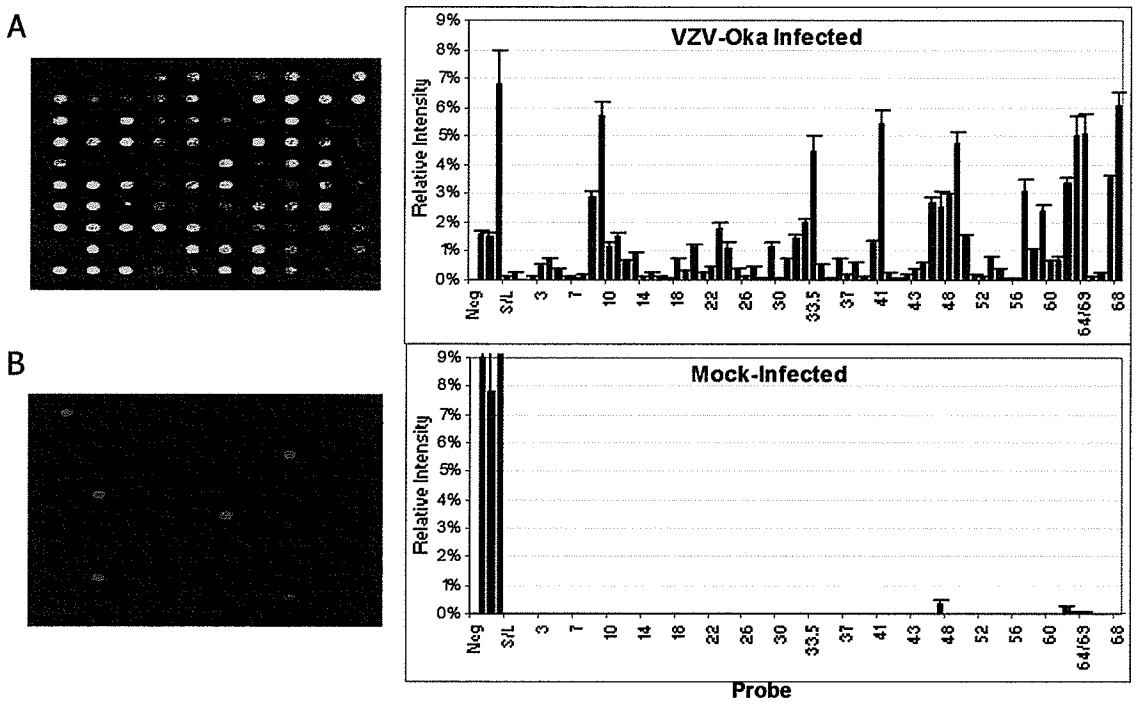



**Figure 7.** Images of slides spotted with 70-mer oligonucleotides (A) or PCR products (B) for VZV ORFs 10, 29, and 62 hybridized with labelled cDNA generated from VZV-Oka infected MRC-5s. In A, each of the three ORFs was probed with one oligo spotted five times. In B, each of the three ORFS was probed with two PCR products, one ~800 bp and the other ~350 bp in length spotted five times each to generate one row. For ORF 62, three additional PCR products were spotted of ~800 and ~300 bp for a total of five probes each spotted five times. In both A and B, printing buffer alone was spotted in the two rows labelled "Neg." with five spots each. The two rows in both A and B labelled "Pos." are two *Arabidopsis thaliana* PCR products spotted five times each on both types of slides. 0.5 ng and 1.0 ng of the corresponding RNAs were spiked in with the experimental RNA and thus hybridized to the slide.

potential cross-hybridization *in silico* (BLAST) against sequences available in the National Centre for Biotechnology Information (NCBI) database. The complete array consisted of 73 probes for viral transcripts, 3 probes for human housekeeping genes (18S rRNA, glyceraldehyde phosphate dehydrogenase and β-actin), and 5 oligo probes for *A. thaliana* transcripts (Table 14). The total number of samples to be spotted was brought up to 100 by the addition of 19 negative control spots consisting of print buffer alone. The VZV gene array was constructed by spotting the 100 samples in a 10 by 10 grid and this grid was repeated in quadruplicate. The array was assessed for the ability to detect viral transcription by hybridization. Briefly, RNA was extracted from T-75 tissue culture flasks of VZV-Oka infected MRC-5 cells showing approximately 80% CPE (two days post infection) and mock-infected cells. Ten μg of RNA was reverse transcribed, labelled with Cy3 or Cy5 and hybridized to the array overnight at 42°C. The slides were washed and scanned and the image of one of five replicate arrays (including two with a dye swap) is shown in Fig 8. Spots containing viral probes hybridized to labelled cDNA from infected cells (Fig 8a) but did not hybridize significantly with cDNA from mock-infected cells (Fig 8b). Human housekeeping probes hybridized to cDNA from both infected and mock-infected cells but as expected their signal intensity level was much lower in the infected cells. Spots containing *A. thaliana* probes hybridized only when the corresponding RNA transcript had been spiked into the reverse transcription reactions. Negative control spots of printing buffer alone did not hybridize when assayed with either sample. The array detected as little as 0.3 ng of transcript (spikes 1 and 5 in Cy5 image and spikes 2 and 5 in Cy3 image).

The net signal intensity was determined for all probes for both VZV and mock




D

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *A.th.* #1 0.3 ng | N | N | SL | SL-I | 1 | 2 | 3 | N | 4 |
| 5 | 6 | 7 | 8 | *A.th.* #2 0.6 ng | N | 9A | 9 | 10 | 11 |
| 12 | N | 13 | 14 | 15 | 16 | 17 | hu 18S rRNA | N | N |
| 18 | 19 | 20 | 21 | 22 | N | 23 | 24 | 25 | 26 |
| 27 | *A.th.* #3 0.6 ng | N | N | 28 | 29 | 30 | 31 | 32 | N |
| 33 | 33.5 | 34 | 35 | 36 | hu GAPDH | N | 37 | 38 | 39 |
| 40 | 41 | N | 42 | 45 | 42+45 | 43 | 44 | *A.th.* #4 0.3 ng | N |
| 46 | 47 | 48 | 49 | 50 | N | 51 | 52 | 53 | 54 |
| 55 | hu ß-actin | N | 56 | 57 | 58 | 59 | 60 | N | 61 |
| 62/71 | 63/70 | 64/69 | 65 | 66 | 67 | 68 | *A.th.* #5 0.3 ng | N | N |

**Figure 8.** Hybridization to the VZV gene array. The array was hybridized with Cy3-labelled cDNA generated from VZV-Oka infected MRC-5s 2 days after infection (A) and Cy5-labelled cDNA generated from mock-infected MRC-5s (B). The bar graphs in A and B are the mean signal intensity relative to the total intensity when hybridized with infected cDNA for each viral probe (S/L and numbers) as well as the three housekeeping genes and negative controls (first four bars) from five replicate arrays. The housekeeping probes follow the negative control in the graph and are ordered β-actin, GAPDH, and 18s rRNA. The viral probes are in numerical order following S/L (the same order as in D). The error bars are the standard error of the mean. The two images from A and B are merged in C. Red spots are those that only hybridized with Cy5-labelled cDNA (ie mock-infected MRC-5s), green spots only hybridized with Cy3-labelled cDNA (ie VZV-infected MRC-5s), and yellow spots hybridized with both. D shows the layout of spots in the grid. N refers to negative control spots (Neg in the graphs) that do not contain probe DNA, *A.th* refers to the *A. thaliana* positive controls, and the numbers refer to viral ORFs. The images in A, B, and C are all the same 10x10 grid containing all 100 samples spotted (shown in D) on one slide. This grid is repeated 4 times on the slide.

infected samples and expressed as a percentage of the total signal (excluding the *A. thaliana* controls) of the channel hybridized with the VZV-infected cDNA. The mean was determined for all five replicate arrays and plotted in the bar graphs shown in Fig 8. The results are summarized in Table 3 which lists each probe in descending order of average relative intensity. The five probes with the highest relative intensity (greater than 5%), indicating the five most abundant transcripts detected, are the probes for ORFs 68, 9, 41, 64/69 and 63/70. Only three of these five have had the functions of the encoded protein determined: ORFs 68, 9 and 63/70. All three encode structural proteins. Notably, the probes with the two highest intensity correspond to proteins which are known to be abundant, namely ORF 68, the most abundant glycoprotein and ORF 9 which is an abundant tegument protein. ORF 63/71 is known to be an immediate early gene and it is likely that ORF 68, a glycoprotein, is a late gene demonstrating the asynchronous nature of the infection.

Probes for a recently discovered spliced ORF, ORF S/L (Kemble *et al.*, 2000) that spans the termini of the genome were included. One probe spanned the exon-intron junction (S/L) while the other spanned the splice junction (S/L-I). Both spliced and non-spliced transcripts were detected with the average relative intensity for the spliced transcript 1.6 fold higher than that for the non-spliced transcript. When a t-test (two-tailed, unequal variance) was performed on the relative net intensities for the two probes in the five replicate arrays, the difference was found to be statistically significant ($p = 0.012$).

ORFs 28 and 29 encode the DNA polymerase (large subunit) and single stranded DNA binding protein respectively. These two ORFs are on opposite strands but have

**Table 3.** Relative expression of VZV-Oka transcripts as determined by hybridization to five replicate gene arrays. The signal intensity for each probe was expressed as a percentage of the total array signal and averaged for the five replicate arrays. The probes were then ranked in order of decreasing relative intensity.

| Probe | Relative Intensity (%) | Gene function |
|---|---|---|
| 68 | 6.02 | Glycoprotein (gE) |
| 9 | 5.72 | Abundant tegument phosphoprotein |
| 41 | 5.46 | |
| 64/69 | 5.11 | |
| 63/70 | 5.03 | Tegument protein |
| 49 | 4.77 | |
| 33.5 | 4.47 | Assembly protein |
| 67 | 3.59 | Glycoprotein (gI) |
| 62/71 | 3.38 | Transactivator, tegument protein |
| 57 | 3.08 | Cytoplasmic protein |
| 48 | 2.96 | |
| 9A | 2.91 | Syncytia formation, virion protein |
| 46 | 2.71 | |
| 47 | 2.52 | Protein kinase |
| 59 | 2.42 | Uracil-DNA glycosylase |
| 33 | 1.97 | Protease |
| 23 | 1.80 | |
| 50 | 1.51 | |
| 11 | 1.49 | |
| 32 | 1.42 | Phosphoprotein, probable substrate for ORF 47 kinase |
| 40 | 1.31 | Major nucleocapsid protein |
| 20 | 1.19 | |
| 10 | 1.17 | Transactivator, tegument protein |
| 29 | 1.14 | Single stranded DNA binding protein |
| 24 | 1.12 | |
| 58 | 1.06 | |
| 13 | 0.89 | Thymidylate synthetase |
| 53 | 0.73 | |
| 18 | 0.73 | Ribonucleotide reductase, small subunit |
| 61 | 0.70 | Transactivator, transrepressor |
| 31 | 0.67 | Glycoprotein (gB), fusogen |
| 4 | 0.67 | Transactivator, tegument protein |
| 36 | 0.66 | Thymidine kinase |
| 12 | 0.63 | |
| 38 | 0.62 | |
| 60 | 0.60 | Glycoprotein (gL), chaperone for gH |
| 44 | 0.55 | |
| 34 | 0.53 | |
| 3 | 0.46 | |
| 22 | 0.45 | |
| 27 | 0.42 | |
| 43 | 0.38 | |
| 25 | 0.37 | |
| 54 | 0.37 | |

| Probe | Relative Intensity (%) | Gene function |
|---|---|---|
| 5 | 0.36 | Glycoprotein (gK) |
| 19 | 0.35 | Ribonucleotide reductase, large subunit |
| 66 | 0.24 | Protein kinase |
| S/L-I | 0.24 | Cytoplasmic protein; probe spans exon junction in Oka |
| 42 | 0.24 | |
| 21 | 0.23 | |
| 15 | 0.23 | |
| 51 | 0.22 | Origin binding protein |
| 37 | 0.18 | Glycoprotein (gH) |
| 8 | 0.17 | Deoxyuridine triphosphatase |
| 42+45 | 0.16 | Probe spans 42 + 45 junction |
| 65 | 0.15 | Virion protein |
| S/L | 0.15 | Cytoplasmic protein; probe spans intron in Oka |
| 2 | 0.13 | |
| 26 | 0.13 | |
| 52 | 0.13 | |
| 6 | 0.13 | |
| 14 | 0.11 | Glycoprotein (gC) |
| 39 | 0.11 | |
| 16 | 0.10 | |
| 7 | 0.09 | |
| 30 | 0.09 | |
| 28 | 0.08 | DNA polymerase (large subunit) |
| 45 | 0.08 | |
| 35 | 0.08 | |
| 56 | 0.06 | |
| 17 | 0.06 | |
| 55 | 0.04 | |
| 1 | 0.03 | Membrane Protein |

overlapping promoters. The activation of one promoter cannot occur without the activation of the other (Meier and Straus, 1993) and thus the levels of the two transcripts are expected to be similar if not the same. This was not seen here after hybridization to cDNA from VZV-Oka infected cells. The level of hybridization to the ORF 28 probe was much lower than that to the ORF 29 probe (Fig 8a bar graph). In fact, inspection of the image of the hybridized array (Fig 8a) shows little if any signal from the ORF 28 probe (row 5, column 5) while the adjacent (row 5, column 6) probe for ORF 29 clearly shows hybridization. Ranking of the viral probes in order of relative signal intensity places ORF 28 43 positions below ORF 29 (rank of 24 versus 67) (Table 3). Nothing has been published to indicate that VZV-Oka has low DNA polymerase expression therefore the reduced signal seen here, in comparison to ORF 29, might be due to a poor probe resulting in reduced hybridization. It is important to remember that these are only preliminary results and must be validated by another method such as RT-PCR.

## 3.2 Synchronous Infection

### 3.2.1 Cell Free VZV

The custom VZV gene array was made to study VZV temporal gene expression in order to classify ORFs as having immediate early (IE), early or late expression. In order to do so a synchronous infection was required so that every infected cell would be at the same point in the viral life cycle. VZV remains highly cell associated in tissue culture and is generally propagated *in vitro* by passaging infected cells onto non-infected cells. A synchronous infection requires that the cells be infected with cell free virus rather than infected cells. Methods for obtaining cell free VZV, sonication of infected cells for

example, generally give low titres. We attempted to achieve enough cell free virus by scaling up the infection. Microcarriers are beads that anchorage-dependent cells can adhere to and grow on. The advantage of microcarriers is that adherent cells can be grown in spinner flasks which allows for an increased density of cells per ml of medium. MRC-5 cells were grown on Cytodex$^{TM}$ 1 (Amersham Biosciences) microcarriers and then infected with VZV-Ellen infected cells at a rate of approximately six non-infected cells to every infected cell. To further increase the virus titre, the cells were placed in medium without any FBS (Adam Burgener, University of Manitoba, personal communication). Two variations of two different methods of preparing cell free VZV were tried using this 50 ml culture: sonication of infected cells and PEG precipitation of the medium. The resulting titres are shown in Fig 9. The best titre of cell free virus was achieved when the infected cells, on the microcarriers, were shaken with glass beads (to disrupt the cells) and then sonicated. The titre was 44.17 PFU/ml for a total of 88.34 PFU (44.17 PFU/ml x 2 ml total virus preparation) from the 12.5 ml microcarrier culture (one quarter of the 50 ml culture) that was sonicated after shaking with glass beads. Considering that $8.52 \times 10^6$ cells had been used to seed the 50 ml culture, only 1 PFU per $2.4 \times 10^4$ cells seeded was achieved (one quarter of the culture had been used so $8.52 \times 10^6/4$ cells / 88.34 PFU).

To determine if enough cell free virus could be obtained to perform the temporal gene expression analysis, the amount of virus required was calculated and then compared to the expected yield from sonication. Ten μg of RNA is needed for hybridization to one array. In order to achieve a yield of 10 μg, an absolute minimum of a T-25 tissue culture flask of infected cells (~$1.2 \times 10^6$ MRC-5s) must be harvested for every array. Using a

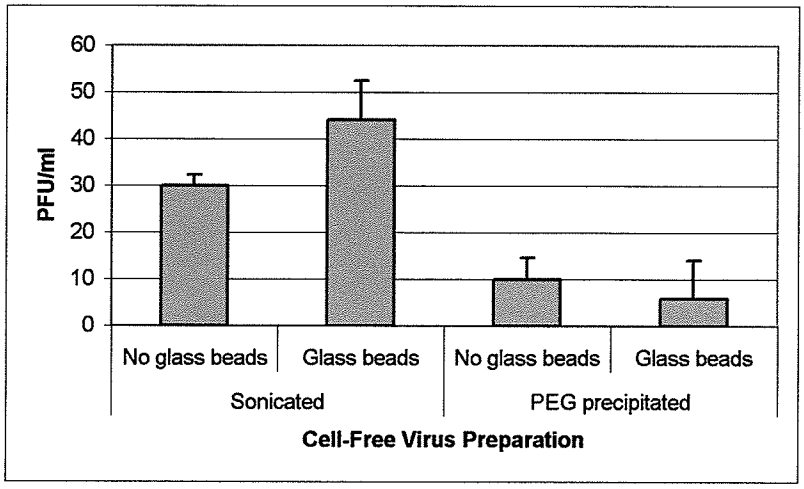



**Figure 9.** Cell free VZV titre using four different methods of virus preparation. A 50 ml 2 g/l Cytodex $^{TM}$ -1 (Amersham) microcarrier culture of MRC-5 cells was infected with VZV-Ellen infected MRC-5 cells (~1 infected cell to 6 non-infected cells) with high CPE. One day after infection, half of the culture was removed to prepare cell free VZV by sonication while the other half was left an additional two days after which virus was isolated from the medium by PEG precipitation. In each case, half of the aliquot (so one quarter of the original culture still on the microcarriers) was shaken with sterile glass beads while the other was shaken without beads. All four cell free virus preparations were stored at -80°C in PSGC storage buffer (a total of 2 ml each). On two separate occasions, an aliquot was thawed and plaque titrated (in triplicate). The results shown are the mean of two independent titrations and the error bar equals one standard deviation.

multiplicity of infection of five (as was used in the previous herpesvirus array papers, Stingley *et al.*, 2000 for example), $6x10^6$ plaque forming units (PFU) are required to infect each flask and thus for each array. A minimum of three time points are needed to classify immediate early, early and late transcripts, and approximately 4 replicates of each array need to be performed for statistical accuracy. This gives a total of 12 arrays (12 T-25 flasks) and thus requires $7.2x10^7$ PFU for every strain to be studied. From information given to me by Dr. Jennifer Moffat (Upstate Medical University), approximately 1000 PFU of cell free VZV is a good yield from a sonication of one T-175 flask of infected cells. Therefore in order to obtain the $7.2x10^7$ PFU needed, 72 000 T-175 flasks ($7.2x10^7$ PFU divided by 1000 PFU/T-175 flask) of infected cells must be harvested for every strain to be studied. This was simply not feasible. Another method of obtaining a synchronous infection was required.

### 3.2.2 Transfection of MeWo Cells with the VZV Genome

When it became clear that a synchronous infection could not be obtained with cell free virus, other methods were investigated one of which was transfection. Moriuchi *et al.* published a paper in 1994 describing the successful transfection of MeWo cells with purified VZV DNA. They found that when they co-transfected MeWo cells with a plasmid encoding the major viral transactivator ORF 62 (IE62) under the control of a CMV promoter, they obtained 60 to 80 fold more infectious virus than when the genome was transfected alone. They transfected 60 mm dishes with 0.5 μg each of viral genomic DNA and IE62 expressing plasmid. A 60 mm dish has approximately the same surface area as a T-25 flask so 0.5 μg of viral DNA could be used to transfect a T-25 flask. Going through the same calculations as above for cell free virus, 6 μg (0.5 μg/flask x 4

replicate flasks x 3 time points) of pure VZV DNA would be required for every strain to perform the temporal gene expression study. From our experience, approximately 3 μg of pure VZV genomic DNA can be obtained by isolating nucleocapsids from a T-75 flask of infected cells and extracting the genomic DNA. Therefore to obtain the 6 μg required for the temporal gene expression study, only about 2 T-75 flasks of infected cells would have to be harvested per strain. If 100 mm tissue culture dishes (~2 times the surface area of T-25 flask) were used, which would guarantee 10 μg of RNA), 1.3 μg of DNA would be needed per dish for a total of 16 μg for every strain or 5 T-75 flasks. This seemed to be very reasonable and achievable.

VZV IE62 was cloned into a plasmid behind a CMV promoter (pTracer$^{TM}$-CMV2 from Invitrogen) (Fig 15). The plasmid also encoded green fluorescent protein (GFP) to allow for easy determination of transfection efficiency by confocal microscopy. VZV-Ellen was grown in MRC-5s in T-175 flasks and viral genomic DNA harvested by isolating nucleocapsids and extracting the DNA. Initially, MRC-5 cells were transfected with 19 ng of the vector (pTracer), vector with IE62 insert (pTracerIE62), or purified VZV DNA plus pTracerIE62 using Lipofectamine 2000$^{TM}$ transfection reagent (Invitrogen). GFP fluorescence was not detected in any of the cells even after four days (data not shown).

Convinced that the transfection should work, it was repeated with MRC-5 and MeWo cells using increased amounts of DNA and commercial medium that was reported to increase transfection rates (Opti-MEM$^{TM}$, Invitrogen). MRC-5 and MeWo cells were transfected in 8-well chamber slides with viral and plasmid DNA. The 0.5 μg of genomic and plasmid DNA used by Moriuchi *et al* (1994) was scaled down by surface area to 19

ng per well (1x) (60 mm dishes have a surface area of 21 $cm^2$ and each well of a 8-well chamber slide has a surface area of 0.8 $cm^2$). Twice (2x, 38 ng) and three times (3x, 57 ng) the DNA was used as well. The total amount of DNA transfected per well was kept constant at 381 ng with sheared salmon sperm DNA (Invitrogen) (scaled down by surface area from the 10 μg used in Moriuchi *et al* 1994). When the cells were checked for presence of GFP two days post transfection, none was detected in any of the MRC-5 wells. The MeWo cells on the other hand were successfully transfected as indicated by the presence of GFP positive cells in both pTracer and pTracerIE62 + VZV-Ellen DNA wells (Fig 10). The transfection efficiency was very low and was best for cells transfected with only the vector. For cells transfected with 19 ng (1x) pTracer, approximately 10% of the cells were GFP positive, twice the amount of DNA (38 ng) resulted in approximately 15% of the cells being GFP positive, and when the amount of DNA was tripled to 57 ng, the maximum number of GFP positive cells were seen at about 20% (Fig 10b). When cells were transfected with pTracerIE62 and VZV-Ellen nucleocapsid DNA, the transfection efficiency dropped off and the maximum was with 3x DNA (57 ng of each) at only about 5% GFP positive cells (Fig 10d). No GFP positive cells were detected with 19 ng each (1x) and only a few cells were GFP positive (~1%) with 38 ng of plasmid and viral DNA (2x).

The transfection was repeated with a second reagent, Effectene™ (Qiagen), using only the vector and MeWo cells and the highest level of GFP positive cells was still only 20% (data not shown).

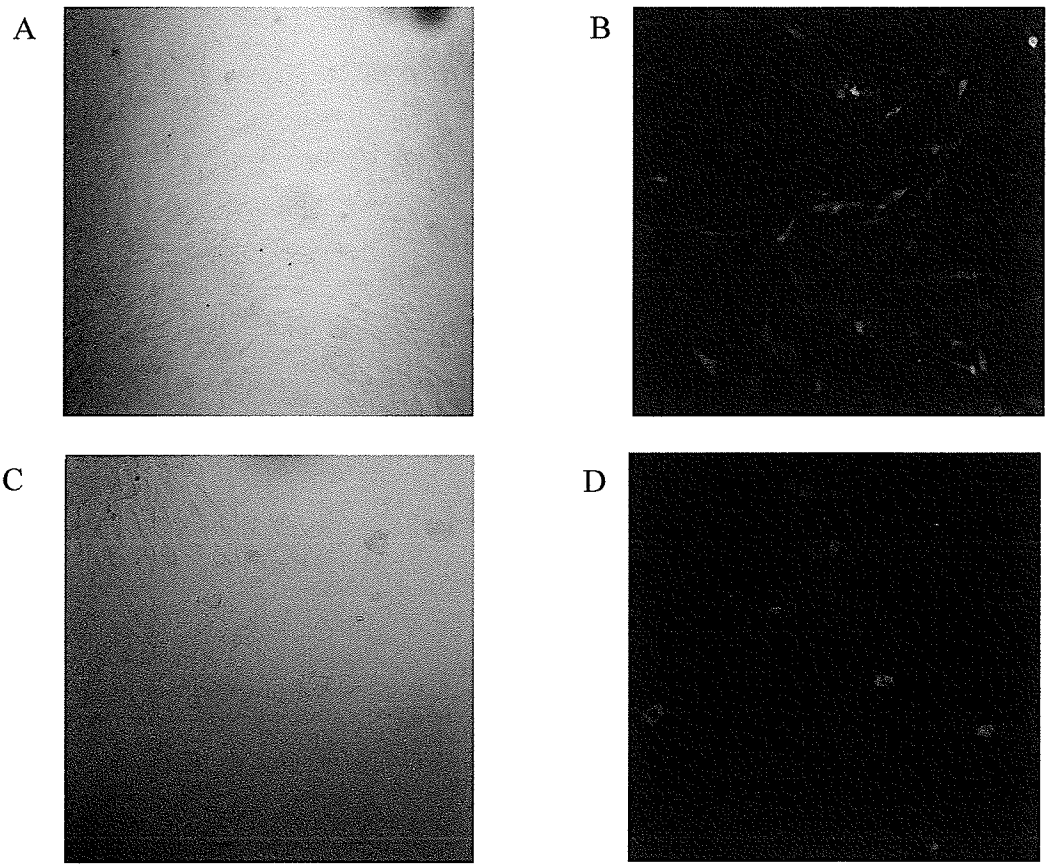



**Figure 10.** Transfection of MeWo cells. A and B: Cells transfected with 57 ng (3x) of pTracer$^{TM}$-CMV2. C and D: Cells transfected with 57 ng (3x) of plasmid bearing IE62 (pTracer$^{TM}$-IE62) and 57 ng (3x) of purified VZV-Ellen DNA. A and C: Transmitted. B and D: Detection of GFP expressing cells.

## 3.3 VZV SNP Arrays

### 3.3.1 Design

Documented SNPs in ORFs 54 and 62 were compiled from the literature, sequence information available in our lab, data from Dr. Charles Grose (Department of Microbiology and Pediatrics, University of Iowa College of Medicine), and the Stanford ORF Comparisons website (http://cmgm.stanford.edu/~jjcheng/VZV/) (Table 2). A total of 61 SNPs in these two ORFs were found and two more SNPs located in ORFs 38 and 68 were included (the *Pst*I and 3B3 mutations respectively). Short oligonucleotide probes (15-30 bases) with the polymorphic site near the centre were made using the coding strand sequence of these four ORFs and included all four possible bases at the polymorphic site. A "probe set" for each SNP then includes four probes differing only by one nucleotide (A, C, G or T) at the central polymorphic position. In the case of 19 SNPs, another SNP was located near enough that it was included in the sequence of the probe for the SNP in question. In these cases, the probe set consisted of four probes differing by the position in question (A, C, G or T) plus another four probes that included the documented neighbouring mutation and A, C, G or T in the site in question. For example, the t94641a SNP is included in the probe sequence for the 94632 SNP (and vice versa). Thus the probe set for 94632 includes 94632a (94641t), 94632c (94641t), 94632g (94641t), 94632t (94641t) and 94632a (94641a), 94632c (94641a), 94632g (94641a), 94632t (94641a) (Table 15). The probes were manufactured by Qiagen Operon and duplicate probes for 4 SNPs (*Pst*I, *Bgl*I, and *Sma*I restriction sites in ORFs 38, 54 and 62 respectively and the 3B3 site in ORF 68) were made by Sigma Genosys for a total of 67 probe sets and 368 probes. The probe sequences are listed in Table 15. After receiving

the probes and going over the sequences, it was determined that the probes for the SNP at position 105264 (ORF 62) were the noncoding sequence rather than the coding sequence and thus would not hybridize and were excluded from further study. The total number of sites studied was then 62 with 4 in duplicate for a total of 66 probe sets.

### 3.3.2 Array Validation with VZV-Ellen, Oka and BC

Ellen and Oka sequences are available in GenBank (accession numbers AY017047 and AY010906 for Ellen and AB097932 and AY016449 for Oka) and BC was sequenced through our facility (Tipples, 2002) and the identity of nearly all of the SNPS for these three VZV strains could be determined from these readily available sequences and from the literature. This information was used to determine the accuracy of the array. Labelled RNA from VZV-Ellen, Oka and BC amplicons were hybridized to four arrays each and the data collected and analyzed. To remove variation across the slide and between slides in order to compare the data, the net intensity (i.e. background corrected intensity) was expressed as the proportion of the total for the particular probe set (i.e. the net intensity for each probe was divided by the sum of all probes for that particular SNP for each replicate). For example, the net intensities for one replicate of probe set 69349 (*Pst*I) when hybridized with labelled RNA from VZV-Ellen were A: 8847.63; C: 5893.87; G: 8065.54; T: 14234.75. The sum of the intensity for the four probes of the set is 37041.79. Thus the fraction of the total signal for the set that each probe contributes is A: 0.2389 (= 8847.63 ÷ 37041.79), C: 0.1591 (= 5893.87 ÷ 37041.79), G: 0.2177 (= 8065.54 ÷ 37041.79), and T: 0.3843 (= 14234.75 ÷ 37041.79). This was repeated for each of the four replicate probe sets (i.e. replicate spots) on an array. The two probes for each probe set that had the highest proportion (and the highest

net intensity) were compared by one-tailed t-tests (unequal variances) (in this example, A and T). If the p-value was less than 0.05, then the difference between the two probes was considered significant and the base was called as the probe with the highest proportion of the signal for the set. The results of one replicate array for the *Pst*I, *Bgl*I, *Sma*I, and 3B3 SNPs are shown in Figures 11 - 14. For all 4 of these sites, the difference between the two highest probes was significant (p-value < 0.05 by one-tailed t-test) for the three VZV strains (Ellen, Oka and BC) and thus a base was assigned to these sites. The identity of these sites is known for these strains and the base assigned by this method was in fact the correct base for all 4 sites and all 3 strains.

The four replicate arrays for each strain were analyzed separately (n=4) and these results were used to determine the reproducibility and accuracy of both the array and the probe sets (Tables 4 and 5). The data for the replicate arrays was also combined and analyzed together (n=16) and the accuracy when the data was combined determined (Table 4). The call rate went up when the data was analyzed together from an average call rate of 90% to 97%. The accuracy of the calls made didn't change but since more calls were made, the accuracy of the array as a whole went up when the replicate arrays were combined from 87% to 93% on average. The error rate remained low at about 3-4% when the data was analyzed separately or together. As for the probe performance, the majority of the probes performed very well and were highly reproducible (average of 90%) and accurate (87% on average) (Table 5). A few probes performed poorly with low reproducibility and accuracy but only 8 out of 66 probe sets have averages lower than 75%. One probe set stood out as being particularly poor and that was the probe set for

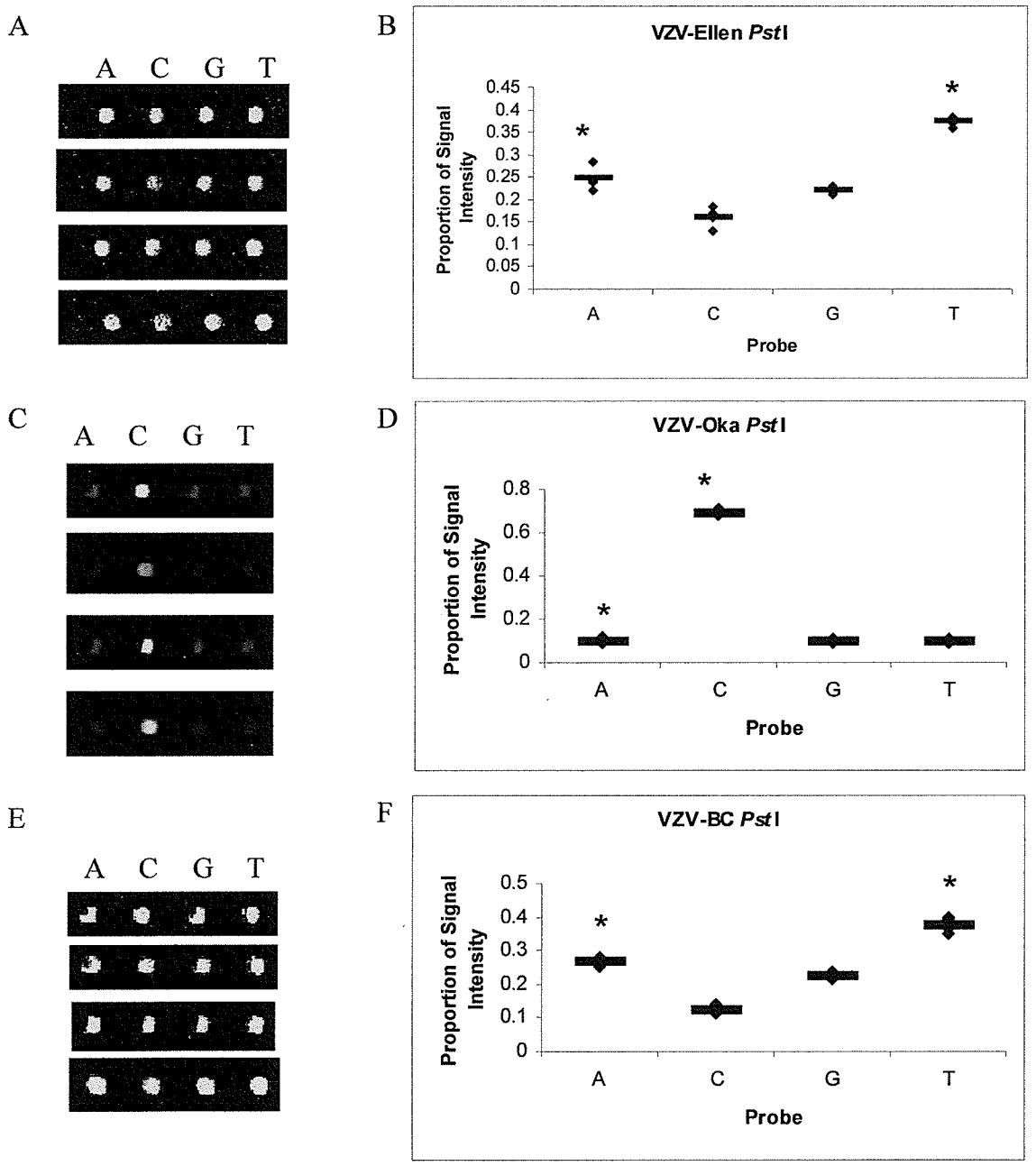



**Figure 11.** Hybridization results of *Pst*I probe set (69349) and VZV strains Ellen (A and B), Oka (C and D) and BC (E and F). A, C and E: Images of one of four replicate *Pst*I probe sets on one of the four replicated slides after hybridization with VZV strains Ellen (A), Oka (C) and BC (E). The images were pseudocoloured cyan for visual ease and converted to JPEGs. B, D and F: Graphs of proportion of signal intensity versus probe for all four replicate *Pst*I probe sets on one array for VZV strains Ellen (B), Oka, (D) and BC (F). Dumas is T (*Pst*I positive) and the documented mutation is C (*Pst*I negative). Thus by these results, VZV-Ellen and BC are *Pst*I positive and Oka is *Pst*I negative as expected. * p-value < 0.0004, 1-tailed t-test with unequal variances.

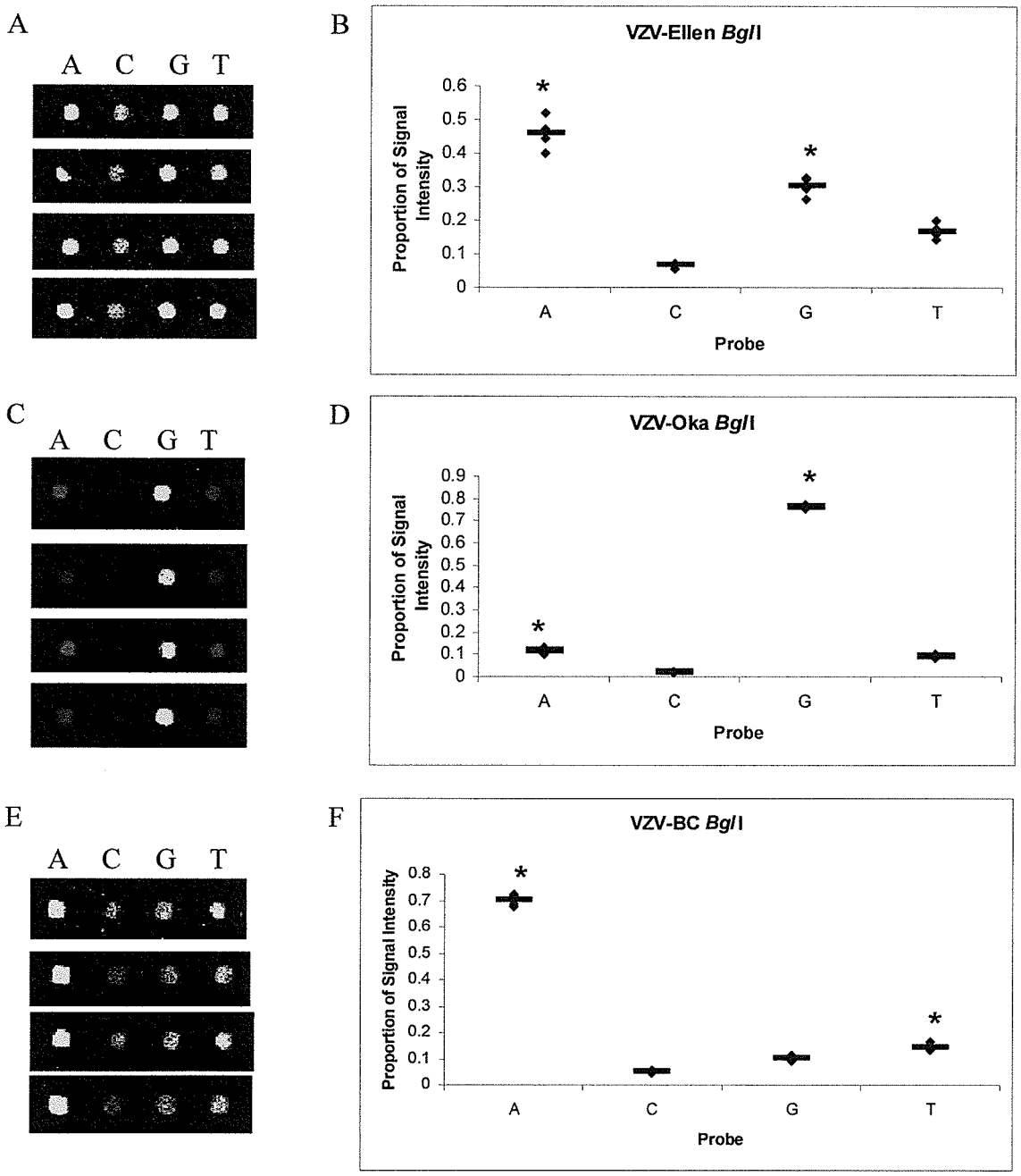



**Figure 12.** Hybridization results of *BglI* probe set (95241) and VZV strains Ellen (A and B), Oka (C and D) and BC (E and F). A, C and E: Images of one of four replicate *BglI* probe sets on one of the four replicated slides after hybridization with VZV strains Ellen (A), Oka (C) and BC (E). The images were pseudocoloured cyan for visual ease and converted to JPEGs. B, D and F: Graphs of proportion of signal intensity versus probe for all four replicate *BglI* probe sets on one array for VZV strains Ellen (B), Oka, (D) and BC (F). Dumas is A (*BglI* negative) and the documented mutation is G (*BglI* positive). Thus by these results, VZV-Ellen and BC are *BglI* negative and Oka is *BglI* positive as expected. * p-value < 0.002, 1-tailed t-test with unequal variances.

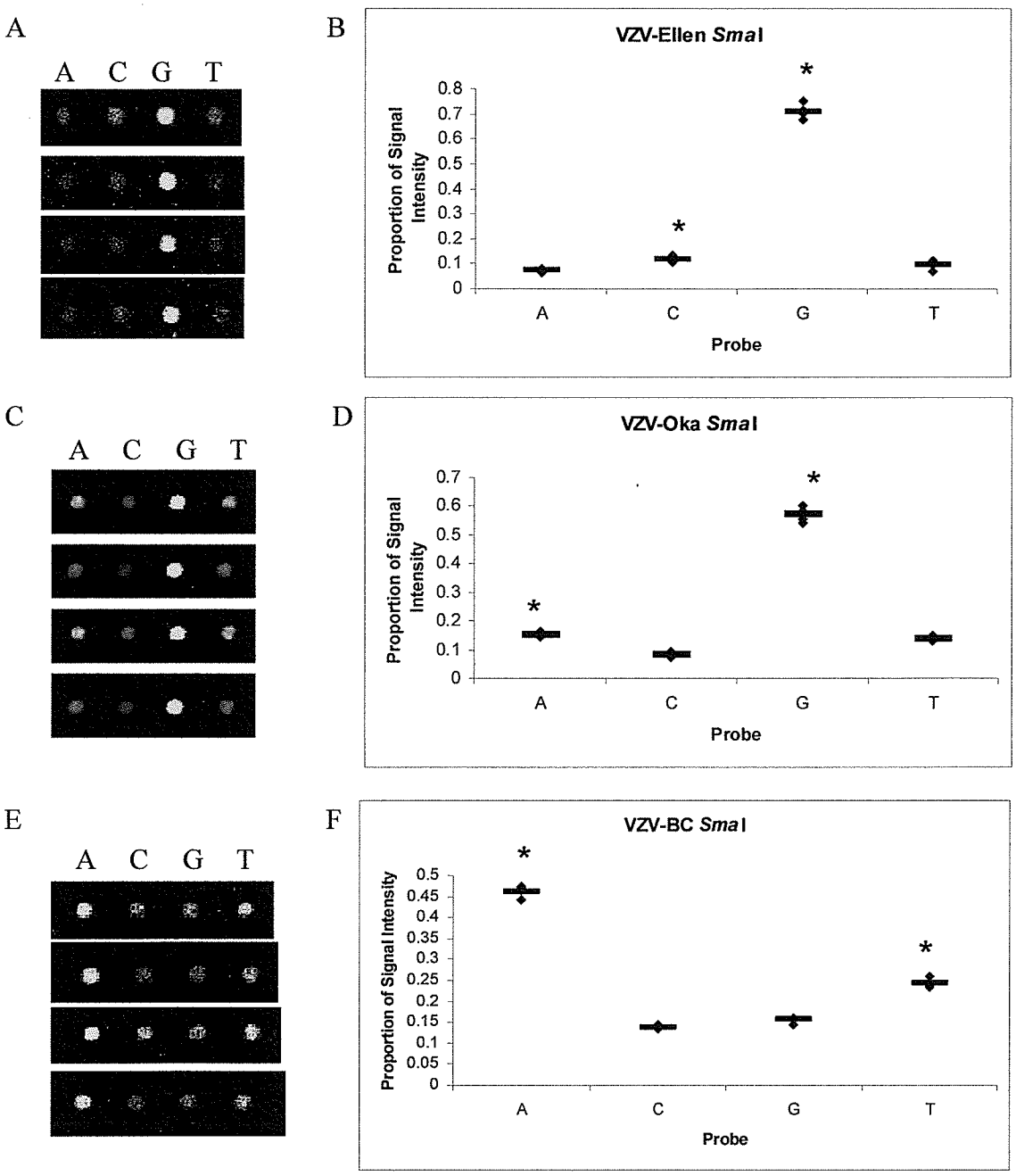



**Figure 13.** Hybridization results of *Sma*I probe set (106262) and VZV strains Ellen (A and B), Oka (C and D) and BC (E and F). A, C and E: Images of one of four replicate *Sma*I probe sets on one of the four replicated slides after hybridization with VZV strains Ellen (A), Oka (C) and BC (E). The images were pseudocoloured cyan for visual ease and converted to JPEGs. B, D and F: Graphs of proportion of signal intensity versus probe for all four replicate *Sma*I probe sets on one array for VZV strains Ellen (B), Oka, (D) and BC (F). Dumas is A (*Sma*I negative) and the documented mutation is G (*Sma*I positive). Thus by these results, VZV-Ellen and Oka are *Sma*I positive and BC is *Sma*I negative as expected. * p-value < 0.00001, 1-tailed t-test with unequal variances.

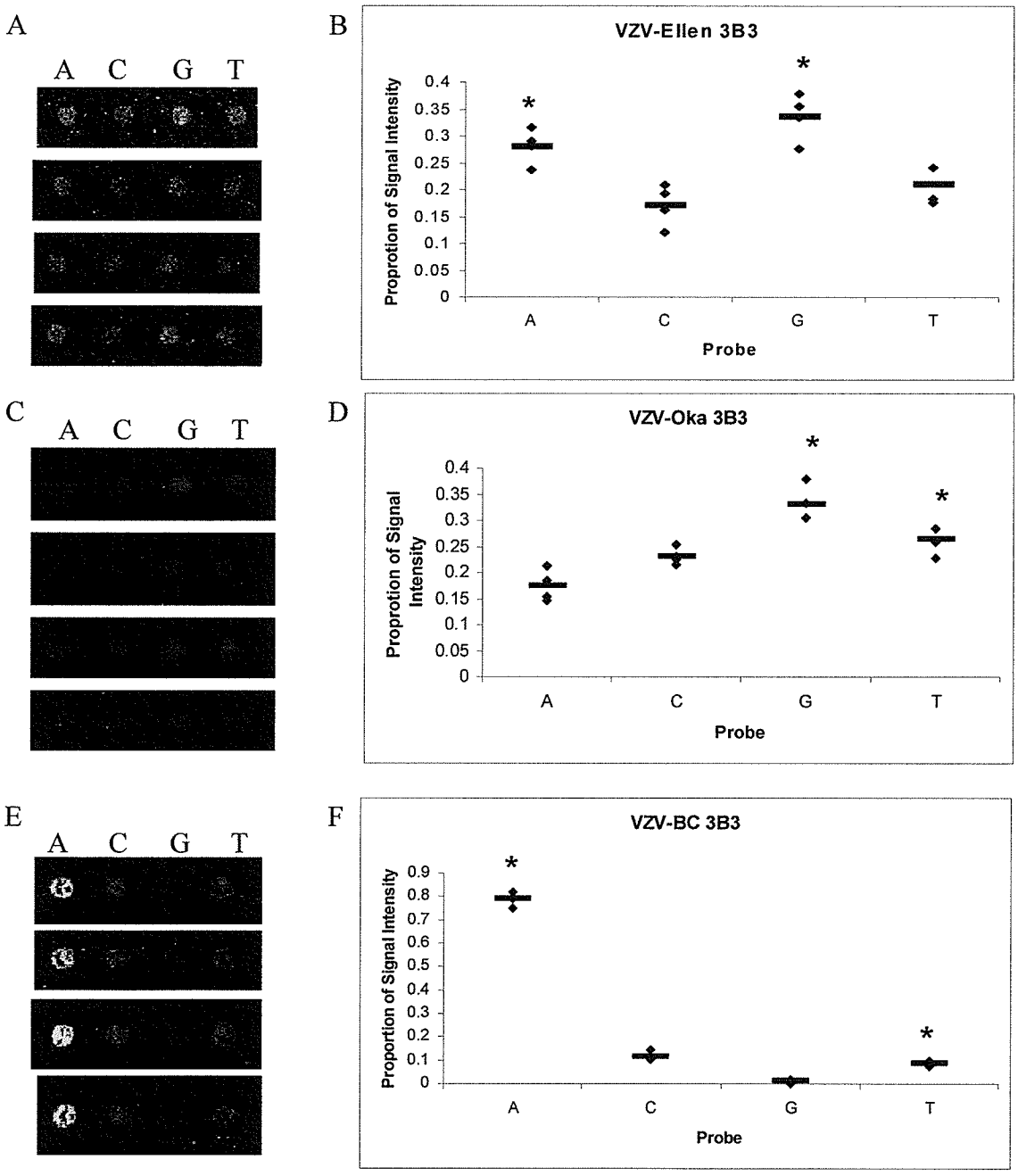



**Figure 14.** Hybridization results of 3B3 probe set (116255) and VZV strains Ellen (A and B), Oka (C and D) and BC (E and F). A, C and E: Images of one of four replicate 3B3 probe sets on one of the four replicated slides after hybridization with VZV strains Ellen (A), Oka (C) and BC (E). The images were pseudocoloured cyan for visual ease and converted to JPEGs. B, D and F: Graphs of proportion of signal intensity versus probe for all four replicate 3B3 probe sets on one array for VZV strains Ellen (B), Oka, (D) and BC (F). Dumas is G (3B3 positive) and the documented mutation is A (3B3 negative). Thus by these results, VZV-Ellen and Oka are 3B3 positive and BC is 3B3 negative as expected. * p-value < 0.05, 1-tailed t-test with unequal variances.

**Table 4.** Accuracy of the SNP array as determined by hybridization to VZV strains Ellen, Oka and BC (four replicate arrays each) and comparison of the results to sequence information available in GenBank, the literature, or in the case of VZV-BC, sequencing performed through our facility. The replicate arrays were analyzed separately (individual replicate arrays) and together (combined) and the results compared.

| | VZV-Ellen | | VZV-Oka | | VZV-BC | |
|---|---|---|---|---|---|---|
| | **Individual Replicate Arrays (n=4)** [1] | **Combined (n=16)** | **Individual Replicate Arrays (n=4)** | **Combined (n=16)** | **Individual Replicate Arrays (n=4)** | **Combined (n=16)** |
| **Total Probe Sets** | 66 | 66 | 66 | 66 | 66 | 66 |
| **# of sites known** | 59 | 59 | 66 | 66 | 66 | 66 |
| **# of calls ($p < 0.05$)** | 59.0 ± 2.8 | 63 | 60.3 ± 1.9 | 65 | 59.8 ± 1.0 | 64 |
| **# correct calls** | 50.0 ± 2.8 | 55 | 58.8 ± 1.5 | 63 | 57.3 ± 0.5 | 60 |
| **# calls wrong (not N)** | 2.0 ± 0.0 | 1 | 1.5 ± 0.6 | 2 | 2.5 ± 0.6 | 4 |
| **Call Rate**[2] | 89.4 ± 4.3% | 95.5% | 91.3 ± 2.9% | 98.5% | 90.5 ± 1.5% | 97.0% |
| **Accuracy of Calls Made**[3] | 84.7 ± 0.8% | 87.3% | 97.5 ± 0.9% | 96.9% | 95.8 ± 0.9% | 93.8% |
| **Accuracy of Array**[4] | 84.7 ± 4.8% | 93.2% | 89.0 ± 2.3% | 95.5% | 86.7 ± 0.8% | 90.9% |
| **Error Rate**[5] | 3.4 ± 0.2% | 1.6% | 2.5 ± 0.9% | 3.1% | 4.2 ± 0.9% | 6.3% |

[1]The data given is the average ± the standard deviation from the four replicate arrays that were analyzed separately.
[2]Call rate = (total # of calls / total # of probe sets) x 100
[3]Accuracy of calls made = (# of correct calls made by array / total # of calls made by array) x 100
[4]Accuracy of array = (# of correct calls made by array / total # of known sites) x 100
[5]Error Rate = (# of incorrect calls/# of calls made) x 100

**Table 5.** Accuracy and reproducibility of the SNP array probes as determined by hybridization to VZV strains Ellen, Oka and BC (four replicate arrays each) and comparison of the results to sequence information available in GenBank, the literature, or in the case of VZV-BC, sequencing performed through our facility. The replicate arrays were analyzed separately and the results used to determine the reproducibility and accuracy of the probe sets. Results are only shown for those probes that had reproducibility or accuracy less than 100%.

| SNP | Reproducibility[1] | | Accuracy[2] | |
|---|---|---|---|---|
| | Average | Range | Average | Range |
| 69349 (PstI) | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 69349 Genosys | 83.3% | 50 - 100% | 83.3% | 50 - 100% |
| 94641 | 75.0% | 25 - 100% | 62.5% | 25 - 100% |
| 105234 | 66.7% | 50 - 75% | 41.7% | 0 - 75% |
| 105312 | 83.3% | 50 - 100% | 83.3% | 50 - 100% |
| 105356 | 83.3% | 75 - 100% | 83.3% | 75 - 100% |
| 105406 | 58.3% | 25 - 75% | 58.3% | 25 - 75% |
| 105490 | 66.7% | 25 - 100% | 66.7% | 25 - 100% |
| 105512 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 105915 | 75.0% | 50 - 100% | 75.0% | 50 - 100% |
| 105919 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 105923 | 33.3% | 0 - 75% | 33.3% | 0 - 75% |
| 106247 | 83.3% | 50 - 100% | 83.3% | 50 - 100% |
| 106710 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 107026 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 107070 | 66.7% | 50 - 100% | 66.7% | 50 - 100% |
| 107329 | 33.3% | 0 - 75% | 33.3% | 0 - 75% |
| 107715 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 107797 | 83.3% | 75 - 100% | 83.3% | 75 - 100% |
| 108189 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 108551 | 50.0% | 50 - 50% | 16.7% | 0 - 50% |
| 108564 | 75.0% | 25 - 100% | 75.0% | 25 - 100% |
| 108618 | 83.3% | 50 - 100% | 83.3% | 50 - 100% |
| 108741 | 50.0% | 25 - 75% | 50.0% | 25 - 75% |
| 108838 | 91.7% | 75 - 100% | 91.7% | 75 - 100% |
| 109010 | 83.3% | 75 - 100% | 83.3% | 75 - 100% |
| 109044 | 100.0% | 100 - 100% | 0.0% | 0 - 0% |
| 116255 (3B3) | 50.1% | 0 - 100% | 58.3% | 25 - 100% |
| Average | 89.6% | 33.3 - 100.0% | 87.2% | 0.0 - 100.0% |

[1]Reproducibility= (# of identical calls made by a particular SNP across multiple arrays with the same sample / total # of arrays) x 100

[2]Accuracy = (# of correct calls made by a particular SNP across multiple arrays with the same sample / total # of arrays) x 100

the 109044 SNP. Although it was highly reproducible (100%), it was completely inaccurate for all 3 strains and assigned the incorrect base (A) for all three.

### 3.3.3 Results of VZV Strains

A total of 11 VZV-strains, including Ellen, Oka, BC and 8 clinical isolates, were analyzed with the SNP array. For each strain, four replicate arrays were hybridized and the data from the four arrays was combined and analyzed exactly as described for Ellen, Oka and BC. The resulting base assignments are listed in Table 6. The call rate remained high for the eight clinical isolates with an average of 93%. The identity of the *Pst*I, *Bgl*I, *Sma*I and 3B3 sites was known for all eight isolates. The correct base had been assigned for all the strains studied at these four sites. As well, there appeared to be a problem with the 3B3 probe set from Qiagen Operon but not the duplicate set from Sigma Genosys. For six of the isolates, no base assignment was made but the incorrect base was ranked highest for five of them. For the remaining two strains, the incorrect base was assigned (A which would indicate the loss of the 3B3 epitope).

### 3.3.4 Verification by Sequencing VZV ORF 62

To verify the SNP array results for the majority of the SNPs included in the array, ORF 62 was sequenced for all VZV strains studied by the array except VZV-BC which we had previously sequenced (those results are given above in section 3.3.2). Thus 54 of the 66 probe sets included on the array were verified by sequencing. The results are shown in Table 7 and the accuracy of the array is summarized in Table 8. The accuracy of the base assignments determined by hybridization to the array was very high with an average of 95.9% (95% confidence limits are 95.9 ± 1.3%) over all 10 strains sequenced.

**Table 6.** Identities of the nucleotide at 66 positions for VZV strains Ellen, Oka, BC and eight clinical isolates (numbers 6, 8, 9, 16, 20, 21, 31 and 55) as determined by hybridization of each strain to four replicate VZV SNP arrays. A base is only called if the p-value of a one-tailed t-test (unequal variances) performed on the two probes of the set with the most hybridization is less than 0.05 (N indicates a p-value > 0.05). The bases for VZV-Dumas and the documented mutations are also shown.

| | 69349 (PstI) | 69349 Genosys | 94167 | 94632* | 94641* | 95241 (BglI) | 95241 Genosys | 95300 | 95546 | 95601 | 105234 | 105310* | 105312* | 105356 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dumas | T (PstI+) | T (PstI+) | A | T | T | A (BglI-) | A (BglI-) | G | C | A | C | T | T | A |
| Mutant | C (PstI-) | C (PstI-) | G | C | A | G (BglI+) | G (BglI+) | A | T | C | T | C | C | G |
| Ellen | T | T | A | T | T | A | A | G | C | A | N | T | T | A |
| Oka | C | C | G | C 94641A | N | G | G | G | T | C | C | C 105312C | C 105310C | G |
| BC | T | T | A | T | T | A | A | G | C | A | A | T | T | A |
| 6 | T | T | N | T | T | A | A | G | C | A | A | T | T | A |
| 8 | T | T | N | T 94641A | A | G | G | G | T | A | A | T | T | A |
| 9 | T | T | N | T | T | A | A | G | C | A | A | T | T | A |
| 16 | T | T | G | T | T | A | A | G | C | A | C | T | T | A |
| 20 | T | T | G | T | T | A | A | G | C | A | A | T | T | A |
| 21 | T | T | G | T | T | A | A | G | C | A | A | T | T | A |
| 31 | T | T | A | T | T | A | A | G | C | A | A | T | T | A |
| 55 | T | T | A | T | A | G | G | A | T | A | A | T | T | A |

*These probes include a neighbouring SNP in their sequences. The identity of the neighbouring SNP (as determined by hybridization) is given below the base of the SNP in question only if it is the documented mutation.

| | 105371 | 105406* | 105413* | 105451 | 105490 | 105510* | 105512* | 105532 | 105544 | 105699* | 105705* | 105855 | 105894 | 105915* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dumas | T | A | A | T | A | A | T | A | T | A | A | A | A | T |
| Mutant | G | G | G | C | G | G | G | G | C | G | G | G | G | C |
| Ellen | T | N | A<br>105406G | T | A | A | T | A | T | A | A | G | A | T<br>105923A |
| Oka | T | A | A | C | A | A<br>105512G | G | A | C | A<br>105705G | G | A | A | T |
| BC | T | T<br>105413G | A | T | A | A | T | A | T | A | A | A | A | T |
| 6 | T | A | A | T | A | A | T | A | T | A | A | A | A | T |
| 8 | T | A | A | T | A | A | T | A | T | A | A | A | A | T |
| 9 | T | A | A | N | A | A | T | A | T | A | A | A | A | T |
| 16 | T | A | A | N | A | A | T | A | T | A | A | A | A | T |
| 20 | T | A | A | T | A | A | T | A | T | A | A | A | A | T |
| 21 | T | A | A | N | A | A | T | A | T | A | A | A | A | T |
| 31 | T | A | A | N | A | A | T | A | T | A | A | A | A | T |
| 55 | T | A | A | N | A | A | T | A | T | A | A | A | A | T |

| | 105919* | 105923* | 105964 | 106029 | 106247 | 106262 (SmaI) | 106262 Genosys | 106497 | 106569 | 106710 | 106905 | 107026 | 107070 | 107136 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dumas | A | C | A | A | T | A (SmaI-) | A (SmaI-) | A | G | T | T | T | A | A |
| Mutant | G | A | G | G | C | G (SmaI+) | G (SmaI+) | G | A | C | C | C | G | G |
| Ellen | A 105923A | A | A | A | C | G | G | A | G | T | T | C | A | A |
| Oka | A | C | A | A | T | G | G | A | G | T | T | T | A | A |
| BC | A | C | A | A | T | A | A | A | G | T | T | T | A | A |
| 6 | A | N | A | A | T | A | A | A | G | T | T | T | A | A |
| 8 | A | C | A | A | T | A | A | A | G | T | T | T | N | A |
| 9 | A | N | A | A | T | A | A | A | G | T | T | T | A | A |
| 16 | A | C | A | A | T | A | A | A | G | T | T | T | A | A |
| 20 | A | C | A | A | T | A | A | A | G | T | T | T | N | A |
| 21 | A | C | A | A | T | A | A | A | G | T | T | T | A | A |
| 31 | A | C | A | A | T | A | A | A | G | T | T | T | A | A |
| 55 | A | C | A | A | T | A | A | A | A | T | T | T | N | A |

| | 107165 | 107252 | 107307 | 107329 | 107586 | 107599* | 107607* | 107715 | 107797 | 108111 | 108189 | 108551* | 108564* | 108591 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dumas | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| Mutant | A | G | G | T | A | C | T | G | C | G | A | G | G | A |
| Ellen | A | G | A | C | G | T | G | A | T | G | G | G 108564G | G 108551G | G |
| Oka | A | G | G | C | G | T 107607T | T | G | T | G | G | A 108564G | A | G |
| BC | G | A | A | N | G | T | G | A | T | A | G | A 108564G | A | G |
| 6 | G | A | A | C | G | T | G | A | T | A | G | N | A | G |
| 8 | A | A | G | N | G | T | G | A | T | A | G | N | A | G |
| 9 | G | A | A | C | G | N | G | N | T | A | G | N | A | G |
| 16 | G | A | A | N | G | T | G | N | T | A | G | N | A | G |
| 20 | G | A | A | N | G | N | G | A | T | A | G | N | A | G |
| 21 | G | A | A | C | G | T 107607T | G | A | T | A | G | A 108564G | A | G |
| 31 | G | A | A | N | G | T | G | A | T | A | G | A | A | G |
| 55 | A | A | G | C | G | T | G | N | T | A | G | A | A | G |

| | 108618 | 108741* | 108747* | 108838 | 108951 | 109010 | 109044 | 109137 | 116255 (3B3) | 116255 Genosys | Call Rate |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Dumas | T | A | T | T | C | T | G | T | G (3B3+) | G (3B3+) | |
| Mutant | C | G | C | C | T | G | C | C | A (3B3-) | A (3B3-) | |
| Ellen | C | G 108747C | C 108741G | T | T | T | A | T | N | G | 95.5% |
| Oka | T | A 108747C | C | T | T | T | A | C | G | G | 98.5% |
| BC | T | N | T | T | C | T | A | T | A | A | 97.0% |
| 6 | T | N | T | T | C | T | A | T | N | G | 92.4% |
| 8 | T | A 108747C | C | T | T | T | A | T | N | G | 92.4% |
| 9 | T | N | T | T | C | T | A | T | N | G | 87.9% |
| 16 | T | N | T | T | C | T | A | T | N | G | 90.9% |
| 20 | T | G | T | T | C | T | A | T | A | G | 93.9% |
| 21 | T | G | T | T | C | T | A | T | A | G | 98.5% |
| 31 | T | A | T | T | C | T | A | T | N | G | 95.5% |
| 55 | T | A 108747C | C | T | T | T | A | T | N | G | 93.9% |

**Table 7.** IE62 sequencing results (53 positions) for VZV strains Ellen, Oka and eight clinical isolates (numbers 6, 8, 9, 16, 20, 21, 31 and 55). Letters in bold indicate differences from the bases assigned by hybridization to the SNP array.

| | 105234 | 105310 | 105312 | 105356 | 105371 | 105406 | 105413 | 105451 | 105490 | 105510 | 105512 | 105532 | 105544 | 105699 | 105705 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ellen | C | T | T | A | T | G | A | T | A | A | T | A | T | A | A |
| Oka | C | **T** | C | **A** | T | A | A | C | A | A | G | A | C | A | G |
| 6 | - | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 8 | **C** | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 9 | **C** | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 16 | C | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 20 | - | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 21 | - | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 31 | **C** | T | T | A | T | A | A | T | A | A | T | A | T | A | A |
| 55 | **C** | T | T | A | T | A | A | T | A | A | T | A | T | A | A |

- indicates that sequence could not be determined for that particular position.

| | 105855 | 105894 | 105915 | 105919 | 105923 | 105964 | 106029 | 106247 | 106262 | 106497 | 106569 | 106710 | 106905 | 107026 | 107070 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ellen | G | A | T | A | A | A | A | C | G | A | G | T | T | C | A |
| Oka | A | A | T | A | C | A | A | T | G | A | G | T | T | T | A |
| 6 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 8 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 9 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 16 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 20 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 21 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 31 | A | A | T | A | C | A | A | T | A | A | G | T | T | T | A |
| 55 | A | A | T | A | C | A | A | T | A | A | A | T | T | T | A |

| | 107136 | 107165 | 107252 | 107307 | 107329 | 107586 | 107599 | 107607 | 107715 | 107797 | 108111 | 108189 | 108551 | 108564 | 108591 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ellen | A | A | G | A | C | G | T | G | A | T | G | G | G | G | G |
| Oka | **G** | A | G | G | C | G | T | T | G | T | G | G | A | A | G |
| 6 | A | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| 8 | A | A | A | G | C | G | T | G | A | T | A | G | A | A | G |
| 9 | A | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| 16 | A | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| 20 | A | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| 21 | A | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| 31 | A | G | A | A | C | G | T | G | A | T | A | G | A | A | G |
| 55 | A | A | A | G | C | G | T | G | A | T | A | G | A | A | G |

| | 108618 | 108741 | 108747 | 108838 | 108951 | 109010 | 109044 | 109137 |
|---|---|---|---|---|---|---|---|---|
| Ellen | C | G | C | T | T | T | **C** | T |
| Oka | T | A | C | T | T | T | **C** | C |
| 6 | T | A | T | T | C | T | **C** | T |
| 8 | T | A | C | T | T | T | **C** | T |
| 9 | T | A | T | T | C | T | **C** | T |
| 16 | T | A | T | T | C | T | **C** | T |
| 20 | T | **A** | T | T | C | T | **C** | T |
| 21 | T | **A** | T | T | C | T | **C** | T |
| 31 | T | A | T | T | C | T | **C** | T |
| 55 | T | A | C | T | T | T | **C** | T |

**Table 8.** Accuracy of the SNP array as determined by sequencing of VZV strains Ellen, Oka and eight clinical isolates (numbers 6, 8, 9, 16, 20, 21, 31 and 55) and comparison to the results obtained by hybridization to the SNP array. The average of the ten strains and standard deviations are also given.

| | Ellen | Oka | 6 | 8 | 9 | 16 | 20 | 21 | 31 | 55 | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Total Probe Sets Sequenced** | 54 | 54 | 53 | 54 | 54 | 54 | 54 | 54 | 54 | 54 | |
| **# of calls ($p < 0.05$)** | 52 | 54 | 50 | 51 | 48 | 49 | 50 | 53 | 52 | 51 | 51.0 ± 1.8 |
| **# Correct Calls** | 51 | 50 | 49 | 49 | 46 | 48 | 47 | 50 | 50 | 49 | 48.9 ± 1.5 |
| **# calls Wrong (not N)** | 1 | 4 | 1 | 2 | 2 | 1 | 3 | 3 | 2 | 2 | 2.1 ± 1.0 |
| **Call Rate**[1] | 96.3% | 100.0% | 94.3% | 94.4% | 88.9% | 90.7% | 92.6% | 98.2% | 96.3% | 94.4% | 94.6 ± 3.3% |
| **Accuracy of Calls Made**[2] | 98.1% | 92.6% | 98.0% | 96.1% | 95.8% | 98.0% | 94.0% | 94.3% | 96.2% | 96.1% | 95.9 ± 1.8% |
| **Accuracy of Array**[3] | 94.4% | 92.6% | 92.5% | 90.7% | 85.2% | 88.9% | 87.0% | 92.6% | 92.6% | 90.7% | 90.7 ± 2.9% |
| **Error Rate**[4] | 1.9% | 7.4% | 2.0% | 3.9% | 4.2% | 2.0% | 6.0% | 5.7% | 3.9% | 3.9% | 4.1 ± 1.8% |

[1]Call rate = (total # of calls / total # of probe sets sequenced) x 100
[2]Accuracy of calls made = (# of correct calls made by array / total # of calls made by array) x 100
[3]Accuracy of array = (# of correct calls made by array / total # of sequenced sites) x 100
[4]Error Rate = (# of incorrect calls/# of calls made) x 100

The accuracy of the array itself is slightly lower due to the inability to assign bases at some positions and is 90.7 ± 2.1% (95% confidence intervals). The error rate was well within the acceptable range with an average of 4.1% (standard deviation of 1.8%) when the average was taken for all strains. Some strains had error rates over 5% (and thus lower accuracy) including VZV strains 20, 21 and Oka with error rates of 6.0, 5.7 and 7.4% respectively. Although these rates seem high, they only represent 3 or 4 wrong base assignments out of all calls made. In the case of VZV-Oka, 3 of these 4 erroneous base assignments at positions 105310, 105356, and 107136 were actually correct according to the literature and/or sequence information in GenBank.

As for the performance of the individual probes, they had an average accuracy of 90.6% (range of 0 to 100%) (Table 9). Again some probes were less accurate including the probe sets for positions 105234, 105451, 107329, 108551, 108741 and 109044 (accuracies of 22, 50, 60, 50, 50 and 0% average for all 10 strains sequenced). Again the 109044 probe set stood out as being highly inaccurate and unreliable. All 11 strains screened by the array (including VZV-BC) had an A assigned at that position by the array. In every case, that was incorrect. In fact neither Dumas nor documented mutants have an A at 109044; they are either G (Dumas) or C (mutation).

**Table 9.** Accuracy of the SNP array probes as determined by sequencing ORF 62 of VZV strains Ellen, Oka and eight clinical isolates and comparison to the results of hybridization to the VZV SNP array. Results are only shown for those probes that had reproducibility or accuracy less than 100%.

| SNP | Accuracy[1] |
|---|---|
| 105234[2] | 22.2% |
| 105310 | 90.0% |
| 105356 | 90.0% |
| 105406 | 90.0% |
| 105451 | 50.0% |
| 105923 | 80.0% |
| 107070 | 70.0% |
| 107136 | 90.0% |
| 107329 | 60.0% |
| 107715 | 70.0% |
| 107599 | 80.0% |
| 108551 | 50.0% |
| 108741 | 50.0% |
| 109044 | 0.0% |
| Average | 90.6% |
| Range | 0.0 - 100.0% |

[1]Accuracy = (# of correct calls made by a particular SNP across multiple samples / total # of samples) x 100

[2]Although good sequence was not obtained for VZV strains 20 and 21 for position 105234, it was assumed that the array results for these two strains were incorrect because the base that was assigned for both (A) is not the documented mutation nor the base in VZV-Dumas.

# 4.0 DISCUSSION

## 4.1 VZV Gene Expression Array

### 4.1.1 Array Design

Both PCR products and oligonucleotide probes had been used in published microarray studies of other herpesviruses (Chambers *et al.* 1999, Bresnahan and Shenk, 2000, Stingley *et al.* 2000, Jenner *et al.* 2001, Paulose-Murphy *et al.* 2001, Ahn *et al.* 2002, and Ebrahimi *et al.* 2003). Although PCR products have the advantage of the ability to be made in-house this is also a disadvantage in that generating enough purified product (at least 1 μg of each product was required for only one printing plate) for all the genes to be queried is very time consuming and labour intensive. As well, even though commercial software was used to design the primers, multiple products were generated with some of the primers tried here. In these cases, extra time and labour were required to either optimize the PCR to eliminate the extraneous products or, failing that, purification of the appropriate band by gel purification as in the case of the ORF 62 3' end PCR products. Oligonucleotide probes on the other hand, have many advantages and few disadvantages. The main disadvantage is the cost associated with having a commercial company synthesize the oligos but once purchased the amount of oligo received will last for thousands of arrays. The main advantage when it comes to herpesvirus arrays, is that oligo probes are strand specific. Many herpesviruses, VZV included, have overlapping ORFs with part of one ORF running on one strand overlapping with another ORF running on the opposite strand. Using oligo probes, the transcripts from overlapping ORFs can be distinguished from one another while this is not the case with PCR products. As well, probes for splice variants, such as ORFs S/L

and 42/45, can easily be engineered using oligonucleotides. For construction of the VZV gene array, both probe types (oligonucleotides 70 bases in length) were compared in a small-scale preliminary array. The oligonucleotide probes for the three viral ORFs tested appeared to be more sensitive than the PCR product probes. The PCR products assayed, on the other hand, resulted in very little hybridization to the viral probes. The calculated melting temperatures for the oligo and PCR product probes were similar but the $T_m$'s for the PCR products were slightly higher. This may not have been enough to explain the differences in hybridization but that combined with the fact that the amplicons are longer and may very well have more stable secondary structure the differences in hybridization intensities may have been expected. As well, the oligos are attached to the slide via a Schiff's base reaction between the amine modification at the 5' end of the probe and the aldehyde coating on the slide. This means that the length of the oligo probe remains free from the surface of the slide and thus available for hybridization. Conversely, the amplicons bind to the slide surface through ionic interactions between the negatively charged phosphate groups of the probes and the amine groups on the slide and then they are immobilized onto the slide surface by baking. Therefore, the length of the probe is physically bound to the surface and not as available for binding to free target unlike the oligo probes. There may have been other factors involved in the results (for example the probes were spotted in different buffers and onto different slides and slide chemistries, each optimized for either PCR products or oligos) but based on the visible results we chose to construct the array with oligo probes rather than PCR products. The complete array appeared to work very well in its ability to specifically detect viral transcripts and was sensitive down to 0.3 ng.

### 4.1.2 Hybridization with VZV-Oka cDNA

Hybridization of cDNA generated from mock and VZV-Oka infected MRC-5 cells demonstrated that nearly all viral transcripts could be detected with little cross-hybridization to cellular transcripts (Fig 8 a and b). Only two viral probes, for ORFs 48 and 64/69, showed detectable hybridization to cellular transcripts when hybridized to cDNA from mock infected cells (Fig 8b bar graph). None the less, the level of hybridization remained low, contributing less than 0.5% of the total signal from hybridization with viral cDNA while hybridization with viral cDNA to these two probes was considerably higher (Fig 8a bar graph).

The inclusion of probes to detect spliced and non-spliced ORF S/L transcripts resulted in the detection of 1.6 fold more of the spliced species (p-value of 0.012). Splicing of this transcript has been demonstrated for VZV-Oka purchased from American Type Culture Collection (ATCC) and through sequence comparison, it was found that the intronic region was conserved between Oka-ATCC and the Oka used here (Merck) as well as two other strains (Kemble *et al.*, 2000). To my knowledge this is the first study that suggests that the ORF S/L transcript is spliced in Oka-Merck as well although these results are only preliminary and will have to be confirmed by other methods.

At the time of writing this report, a paper was published describing the production and use of a PCR product microarray for all unique VZV ORFs except ORFs S/L, 9A and 14 (ORF 33.5 was excluded as well since it is located within the 3' end of ORF 33 on the same strand and therefore cannot be distinguished from ORF 33 transcripts) (Cohrs *et al.*, 2003). The authors cocultivated VZV-Ellen infected cells with non-infected cells, harvested RNA every day for four days and hybridized to their array. Over this period,

they were able to detect all transcripts for which they had probes and they found that the levels went up over time, peaking at three days post infection. They noted that the patterns of individual transcript abundance did not change; they all simply increased. When their data from two, three and four days post infection was combined, the four most abundant transcripts were ORFs 9, 64, 33 and 49. Only ORFs 9 and 64 ranked in the four highest in the array presented here. The remaining two were ranked 16 (ORF 33) and 6 (ORF 49). There was not much agreement in the rankings of the remainder of the VZV ORFs (Table 10); only 29 of the 68 ORFs with probes in both arrays were ranked within ten of each other. It is difficult to infer that these differences may be due to strain differences (VZV strain Ellen was used by Cohrs and Straus while Oka was used here) because there are too many confounding factors. First, the array platforms are different (PCR products versus oligos, nylon filter array versus glass slide, radiolabelling versus fluorescent labeling). Secondly the cell lines used are from different organisms. The cells used here are human lung fibroblasts which, along with human melanoma cells (MeWo), are the most commonly used cell lines for VZV culture while Cohrs and Straus used an African green monkey kidney cell line (BSC-1). It is entirely possible that transcript levels, with respect to differences between transcripts, could be dependent upon the host cell. Lastly, and perhaps most importantly, is that the infection is undefined in both situations due to the inability to use cell-free virus for infection. More work will have to be done to ensure that the levels of individual transcripts do not change significantly compared to the rest as a function of time or progress of infection. The study undertaken by Cohrs and Straus is an important first step but it will likely have to be repeated for each strain to be studied before comparisons can be made. At present no

**Table 10.** Rank of relative expression of VZV-Oka transcripts as determined by hybridization to five replicate gene arrays compared to the published rank for VZV-Ellen transcripts. Transcripts ranked within ten of each other are shown in bold.

| Probe | Gene function | Rank | Rank in Cohrs *et al.*, 2003 |
|---|---|---|---|
| 68 | Glycoprotein (gE) | 1 | 19 |
| **9** | **Abundant tegument phosphoprotein** | **2** | **1** |
| 41 | | 3 | 33 |
| **64/69** | | **4** | **2** |
| **63/70** | **Tegument protein** | **5** | **5** |
| **49** | | **6** | **4** |
| 33.5 | Assembly protein | 7 | nd* |
| 67 | Glycoprotein (gI) | 8 | 22 |
| **62/71** | **Transactivator, tegument protein** | **9** | **7** |
| **57** | **Cytoplasmic protein** | **10** | **6** |
| **48** | | **11** | **9** |
| 9A | Syncytia formation, virion protein | 12 | nd |
| 46 | | 13 | 66 |
| 47 | Protein kinase | 14 | 52 |
| 59 | Uracil-DNA glycosylase | 15 | 37 |
| 33 | Protease | 16 | 3 |
| 23 | | 17 | 57 |
| **50** | | **18** | **8** |
| **11** | | **19** | **10** |
| 32 | Phosphoprotein, probable substrate for ORF 47 kinase | 20 | 32 |
| 40 | Major nucleocapsid protein | 21 | 61 |
| **20** | | **22** | **15** |
| **10** | **Transactivator, tegument protein** | **23** | **17** |
| **29** | **Single stranded DNA binding protein** | **24** | **29** |
| 24 | | 25 | 11 |
| 58 | | 26 | 14 |
| **13** | **Thymidylate synthetase** | **27** | **34** |
| **53** | | **28** | **23** |
| **18** | **Ribonucleotide reductase, small subunit** | **29** | **36** |
| 61 | Transactivator, transrepressor | 30 | 18 |
| 31 | Glycoprotein (gB), fusogen | 31 | 43 |
| **4** | **Transactivator, tegument protein** | **32** | **31** |
| 36 | Thymidine kinase | 33 | 46 |
| 12 | | 34 | 63 |
| 38 | | 35 | 12 |
| **60** | **Glycoprotein (gL), chaperone for gH** | **36** | **41** |
| 44 | | 37 | 26 |
| 34 | | 38 | 49 |
| 3 | | 39 | 20 |
| **22** | | **40** | **40** |
| **27** | | **41** | **39** |
| **43** | | **42** | **44** |
| **25** | | **43** | **47** |
| 54 | | 44 | 28 |

| Probe | Gene function | Rank | Rank in Cohrs *et al.*, 2003 |
|---|---|---|---|
| 5 | Glycoprotein (gK) | 45 | 67 |
| **19** | **Ribonucleotide reductase, large subunit** | **46** | **54** |
| 66 | Protein kinase | 47 | 59 |
| S/L-I | Cytoplasmic protein; probe spans exon junction in Oka | 48 | nd |
| 42 | | 49 | 21 |
| **21** | | **50** | **45** |
| 15 | | 51 | 27 |
| 51 | Origin binding protein | 52 | 30 |
| 37 | Glycoprotein (gH) | 53 | 25 |
| **8** | **Deoxyuridine triphosphatase** | **54** | **58** |
| 42+45 | Probe spans 42 + 45 junction | 55 | nd |
| 65 | Virion protein | 56 | 42 |
| S/L | Cytoplasmic protein; probe spans intron in Oka | 57 | nd |
| 2 | | 58 | 16 |
| 26 | | 59 | 38 |
| **52** | | **60** | **64** |
| 6 | | 61 | 48 |
| **14** | **Glycoprotein (gC)** | **62** | **68** |
| 39 | | 63 | 50 |
| **16** | | **64** | **65** |
| 7 | | 65 | 35 |
| 30 | | 66 | 53 |
| 28 | DNA polymerase (large subunit) | 67 | 24 |
| **45** | | **68** | **62** |
| **35** | | **69** | **60** |
| 56 | | 70 | 51 |
| 17 | | 71 | 56 |
| 55 | | 72 | 55 |
| 1 | Membrane Protein | 73 | 13 |

*not determined. Probes for these transcripts were not included in the array.

such comparison can be made because it is not known if the pattern of transcript abundance seen here will change if RNA is extracted at a different point in the infection.

### 4.1.3 Cell Free VZV

It is a well known fact in VZV biology that VZV remains highly cell-associated in tissue culture. It was noted as early as 1953 that although VZV could be propagated in cell culture, substantial amounts of cell free virus were not produced (Weller, 1953). A standard method of obtaining cell free virus from infected tissue culture cells involves sonication of the cells but yields remain low at about $10^3$ PFU/ml (Jennifer Moffat, personal communication). The virus can be further concentrated by various methods including polyethylene glycol precipitation and storage buffers have been formulated to maintain infectivity during storage but virus prepared by sonication cannot be purified from the cellular debris without a loss in infectivity. Thus the virus is never truly "cell-free" nor would a titre determined by plaque titration be very accurate as more than one infectious, plaque-forming particle may be bound to the same bit of cellular debris.

When the VZV gene expression array project was first started, we were not aware of the amount of material required for the experiment and thus were not concerned with this well-known issue. Our approach was to scale up the infection using a microcarrier culture of cells. The titre achieved here was even 20 fold less than the $10^3$ PFU/ml expected at only 44.17 PFU/ml. As well, culturing of cells on microcarriers proved to be very difficult. Once the beads had been seeded with cells, we were unable to passage them onto more beads to facilitate scaling up the culture for infection purposes.

Once it became apparent that achieving a synchronous VZV infection could not be achieved by a traditional method such as cell-free virus, transfection of cells with

purified VZV genomic DNA was investigated. Moriuchi *et al.* (1994) successfully transfected MeWo melanoma cells with purified VZV genomic DNA using calcium phosphate as the transfection method. Their paper describing the transfection was published nearly a decade ago and a vast array of transfection reagents that are more reliable and easier to use than calcium phosphate have since become available. After investigating different transfection reagents from a variety of manufacturers, I came across Lipofectamine 2000™, a cationic lipid based reagent, from Invitrogen. This reagent was appealing because according to technical notes available from Invitrogen, it gave the highest transfection efficiency with MRC-5 cells of the Invitrogen line of transfection products and it had been used to successfully transfect Vero cells with the HSV-1 genome which at 150 kb is larger than the VZV genome. The Moriuchi *et al.* (1994) transfection experiment was repeated but with MRC-5s instead of MeWos and using Lipofectamine 2000™ rather than calcium phosphate transfection. The amount of DNA used by Moriuchi was kept the same but scaled down by surface area from the 60 mm dishes used in the paper to the 8-well chamber slides used here. Special attention was paid to the purity of both the plasmid DNA and viral DNA to be used in the transfection and anionic-exchange columns were used to purify the DNA. After transfecting the MRC-5s and checking for the production of GFP every day for four days, no fluorescence was seen (data not shown). Convinced that the transfection should have been successful, the transfection was repeated but with both MRC-5 and MeWo cell lines and with twice and thrice the amount of DNA used by Moriuchi (scaled down by surface area). Two days post transfection resulted in no GFP production in the MRC-5 cell line but some (20% at the most) in the MeWos. It is important to note that while the same

amount of DNA was used as in the Moriuchi paper (19 ng), this was substantially less than the amount recommended by the manufacturer (600ng). This is likely the reason for the lack of transfected MRC-5 cells (as measured by GFP fluorescence), for which the reagent was supposed to work very well, and the low rate of transfection in the MeWos. Had the quantity of pure viral genomic DNA readily available not been limiting, the transfection would have been attempted with the amounts recommended by the manufacturer. The viral genomic DNA was obtained by isolating viral nucleocapsids from infected cells, ethanol precipitating the DNA and then purifying by passage through an anion-exchange column and is a very involved procedure. Since ultracentrifugation on a glycerol gradient is required to pellet the nucleocapsids, the amount of infected cellular material processed is limited to only four T-175 flasks at a time. Three times the amount of DNA used by Moriuchi *et al.* (1.5 µg for a 60 mm dish) was about the maximum amount of pure viral genomic DNA that could be readily obtained. Another factor that may have contributed to low yields of viral DNA is the fact that the majority of the VZV capsids made in infected cells are aberrant and may contain incomplete viral genomes (Grose *et al.*, 1995).

Another transfection reagent, Effectene™ (Qiagen) was found that required much less DNA for transfecting and was on par with the amounts used by Moriuchi. The transfection was repeated using Effectene™ and MeWo cells since they had been successfully transfected. Three days post-transfection, the amount of GFP fluorescence detected in the cells was only approximately 20% using 75 ng of the vector without the IE62 insert (data not shown). It is unclear why the transfection efficiencies remained low especially when only the vector was used. The GFP gene was driven by a promoter for a

human housekeeping gene, EF-1α, so expression of GFP was likely not a problem.

A known reason for poor transfection is the presence of contaminants in the DNA. Special attention had been paid to ensuring that the DNA was of high purity by using anionic-exchange columns rather than standard silica gel for purification. As well, large amounts of dead cells were not seen so toxicity of the DNA (eg by contaminants such as bacterial endotoxins in the case of the plasmids) or their products did not appear to be an issue. The transfection efficiency was low even with the plasmid bearing GFP alone and decreased with the addition of IE62. This decrease is not unexpected as the IE62 gene almost doubled the size of the plasmid from 6 kb to 10 and thus might not be taken up as well. Perhaps the efficiency could have been increased by the addition of more DNA but although it was possible at the level of the 8-well chamber slide it would not have been easily performed at the larger scale required for the array experiment. There are cell lines that are transfected very well, such as 293T cells, but they were not tried because they are not known to be susceptible to VZV infection. Considering that the purpose of the transfection was to provide a means for starting a synchronous infection to study the life cycle of the virus, it did not seem appropriate to use a cell line not known to support productive VZV infection.

### 4.1.4 Summary of VZV Gene Array and Future Work

The VZV gene array itself looks very promising in its potential uses including studying the transcription profile of different strains. It was unfortunate that technical hurdles developed that were beyond the scope of this master's project. None the less, the analysis of VZV gene expression using the array developed here is a project continuing in the Viral Exanthemata department at the National Microbiology Laboratory. It may be of

interest to note that the transfection methods discussed here have been re-tried as well as other methods such as electroporation and calcium phosphate transfection and none have proven to provide the transfection efficiency desired for studying VZV transcription kinetics. To pursue this study, either another experimental set up or other methods of performing the array hybridizations will have to be investigated. In that regard, the MOI used to infect the cells could be decreased from the 5 used in the HSV-1 paper (Stingley *et al.,* 2000) and other herpesvirus papers by at least ten fold to 0.5 and perhaps as much as 100 fold. In a recent review written by members of the groups who developed the HSV-1 and HCMV gene expression array (Wagner *et al,* 2002) the effect that decreasing the MOI had on hybridization signal intensity and the resulting transcription profile was discussed. When arrays were hybridized with RNA generated from cells infected with HSV-1 or HCMV at an MOI of 5, 0.5 and 0.05 were compared, the signal intensities decreased however the transcripts were still detectable. In the case of the HSV-1 arrays, the pattern of transcripts seen at an MOI of 0.05 was not the same as that seen at the same time post infection with an MOI of 5 however the authors noted that the pattern seen with an MOI of 0.05 resembles the pattern with a higher MOI but at a shorter time post-infection. The experience with the HCMV array with differing MOIs, on the other hand, was that the transcription pattern seen with an MOI of 0.5 correlated very well with the pattern at an MOI of 0.05 ($r^2 = 0.9$). This possibility would be worth investigating with VZV. As well, there are different array protocols which allow for either the amplification of the extracted RNA or of the fluorescent signal itself and new slides that have a mirror coating on one side to enhance the signal actually read by the scanner. Thus if a decrease of 100 fold in the MOI and another decrease by 100 fold in the amount of material

(RNA) required for the array can be achieved then the amount of cell free virus required drops into the realm of achievability (from 72 000 T-175s to 7.2).

The VZV gene expression array developed here has a limitless number of potential uses. The first use that we had envisioned was to determine the transcription kinetics of VZV. Currently, it is assumed based on homology with HSV-1 and knowledge of the life cycles of herpesviruses, that transcription occurs in a tightly controlled kinetic fashion in VZV. This hypothesis could be ascertained with the gene expression array.

Secondly, because herpesviruses have such tightly coordinated transcription programs, knowledge of when a gene is transcribed gives insight into the function of the protein encoded. Thus in the case of VZV, in which many of its ORFs have been assigned putative functions based on homology with the better studied HSV-1, the knowledge of when a gene is transcribed may serve to confirm the putative function or perhaps aid in assigning function in the case of the six VZV ORFs that do not have homologs in HSV-1.

As well, if the transcription profiles of multiple strains of VZV, displaying different phenotypes (eg increased virulence in the case of VZV-MSP or BC or attenuation such as that in VZV-Oka vaccine) are determined, insights can be gained into what transcription processes, be they subtle or dramatic, are involved in the phenotypes. Characterization of many strains at the DNA sequence level has shown that differences between strains is often minor, often only differences in the number of repeats or single nucleotide changes. The mutations that have been noted appear to be random and are found in both coding and non-coding regions. Those mutations found in non-coding

regions may influence transcription of specific ORFs and these changes can be detected by the gene expression array. Used in this manner, the gene expression array may aid in determining genes which are important for virulence and therefore attenuation as well.

A fourth possible use is to use the gene array to study latency. There is currently no *in vitro* system for studying VZV latency but if such a system were developed the array could be used in determining which genes are expressed during latency and more importantly from a clinical perspective, which genes control the transition from latent to lytic cycles.

Another potential use of the gene array that would contribute greatly to knowledge on VZV pathogenesis would be to use it to study tissue tropism. During different periods of the virus's colonization of a host it is found in different tissues, for example, skin cells, T-cells and neuronal cells. The use of an animal model for VZV infection, such as the SCID-hu mouse model in which human skin is grafted onto a SCID mouse, combined with harvesting of specific tissues post infection and screening by the gene array could give valuable insights into viral processes in these different tissues.

In conclusion, the development of the VZV gene array described here provides the technology with which to investigate many aspects of the virus' behaviour. The knowledge gained from these experiments, once it is all assimilated, will shed some light onto VZV's pathology, be it virulent or attenuated.

## 4.2 VZV SNP Array

### 4.2.1 Array and Methodology

The VZV SNP array as designed appeared to perform very well. Although it was not always possible to visibly discern which probe of the set had the highest intensity (as

demonstrated in Figures 11 – 14 a, c and e), the trends are clear in the signal intensity values (Figures 11 – 14 b, d and f) and statistical differences are present as indicated by the high call rates (approximately 90% and up). The presence of visible hybridization to all four probes of a set may be due to overloading of the array with labeled sample. Four PCR amplicons, containing all the sites queried when combined, were *in vitro* transcribed and labeled separately and then mixed together and hybridized to four replicate arrays. In the case of two of the amplicons, they contained only one SNP each queried by the array (69349 and 116255). These amplicons were hybridized to the array in equal amounts as the other two amplicons for ORFs 54 and 62 which contained many more SNPs (seven and 53 respectively). There can be no doubt that for these two SNPs, corresponding to four probe sets, too much labeled sample was hybridized to the array. In addition, the four amplicons (400 ng each) were differentially amplified by the *in vitro* transcription step although they were all amplified separately and in parallel for each strain. The amplicon containing the 3B3 site (116255) always had the least amount of amplification at only about two to three fold. The amplicon containing the *Pst*I site (69349) was always the most amplified of the four amplicons at about ten to 12 fold. As seen in Figures 11 and 14 a, c and e, the signal intensity for the *Pst*I probe set appears to be greater than that of the 3B3 probe set. There is also visible hybridization to all four probes in the *Pst*I making a visual base assignment impossible while this is not the case for the 3B3 probe set. In general, the entire array had very high fluorescence signal, with areas that were often saturated, including the *Pst*I probe set. Thus it would be worth optimizing, before further SNP array analyses are done, the labeling and hybridization protocol with respect to the amount of the individual amplicons that are used for the

hybridization. As well, to increase the throughput of this system, the *in vitro* transcription protocol should be optimized so that individual amplicons of a given strain can be combined and processed together. The method used here was to handle each of the four amplicons for a given strain separately. Thus only two strains could be processed at a time to limit the number of tubes handled and therefore decrease the chance of contamination.

As for the design of the probes for the array, the probes were designed to be the same sequence of the coding strand of the ORFs that the SNPs were found in as outlined in the protocol followed here (Bodrossy, in press). Three of the four ORFs included on the array are on the reverse strand of the genome while the fourth is on the forward or top strand. This means that the majority (all but 1) of the probes have the complement of the published sequence while one is the same as the published genomic sequence. This makes interpretation a bit more difficult than it needs to be and does not allow for ready expansion of the array to include other SNPs. If this array is to be expanded to include all SNPs present in the VZV genome, the easiest method would be to make all the probes complementary to the published, forward strand of the genome. Thus overlapping PCR amplicons encompassing the entire genome could be employed rather than ORF and strand specific amplicons.

As well, the data analysis could be further optimized and perhaps automated. Although the method employed here of taking the fraction of the probe and then applying a t-test appeared to work well as shown by the low error rate and high accuracy of the base calls (averages of 4.1 and 95.9% for all strains sequenced), it only allows for one base to be assigned to a position which may not always be the case (for example, the

VZV-Oka vaccine strain has been noted to be heterogeneous at certain positions). In other words, by this analysis method in which statistical difference is required between the two probes showing the highest signal intensity for a base to be assigned, the inability to assign a base (denoted by the presence of an N) may be because the true identity of that position is in fact a mix of the two bases with the highest intensity. This would never be determined using the method presented here. The use of an analysis of variance (ANOVA) method of data analysis, which looks at the variability between all probes of the set would facilitate this type of analysis. As well, although the patterns of hybridization intensity remained constant in the replicate arrays, the actual values of signal intensity for a given probe varied between arrays. Therefore it was necessary to remove that variability, by taking the fraction the probe contributed to its set, in order to analyze replicate arrays. Another method is to model the variability in the data and then transform the data using the generated model, which also ensures that the data follow a normal distribution. A recent paper describing the genotyping of *Cryptosporidium parvum* with a SNP microarray uses this type of data analysis (Straub *et al.*, 2002). They modeled the error and variability in their data, fit the model to the data for each probe set and then performed ANOVA analysis followed by Tukey's multiple-comparison procedure to determine which probe(s) of the set displayed statistically different hybridization. This is the type of data analysis that I would recommend for further SNP array studies.

### 4.2.2 Probe Performance

As described in the Results section, some probes performed significantly poorer than the majority. Out of 364 probes on the array, it is to be expected that some will

perform more poorly than others. There are only three probe sets that I would note as perhaps being faulty, one of which definitely so, and those are 105234, 109044 and 116255 with average accuracies of 22, 0, and 18% respectively). According to the software used to design the probes, the 105234 probe set was rated "good" and had 82.4% GC, no significant hairpins or dimers but with a run or repeat of 6 nucleotides. Upon inspection of the probe sequence, there were only two bases (excluding the polymorphic base) that were not G or C (GGAGGGCGCGGGGGGAG, polymorphic position in red, VZV-Dumas sequence given). The 109044 probe on the other hand, was given a "poor" rating, had 100% GC, five potential hairpin bonds, eight potential dimer bonds and a run or repeat of two bases. The sequence, like that of the 105234 probe, was all Gs and Cs (excluding the polymorphic position) (GGCCGCGGCGGCCGC, polymorphic position in red, VZV-Dumas sequence given). Of the eleven VZV strains that were screened with SNP array, eight had the wrong base assigned at position 105234 (another assigned no base and two were correct) and all eleven were incorrect at position 109044. For both positions the incorrect base assigned was always A although this was never the VZV-Dumas base or documented mutation for those positions. These probes, with nearly 100% GC content, should not be used for further studies as they have been shown to be unreliable.

The 19 probe sets that had a neighbouring SNP included in the probe sequence performed in general very well. For the majority (17), the base assignments made by the probe sets were correct, even that of the neighbouring SNP included in the probe (for example 94632 in Table 6). In fact the presence of the extra four probes that differed by the neighbouring mutation served as a good verification of the results of the probe set for

the neighbouring SNP. For example, the probe with the highest (statistically significant) hybridization intensity with VZV-Oka at position 94632 is C at position 94632 and A at neighbouring SNP position 94641. Both base assignments are correct according to the literature. For the 94641 probe set, no base was assigned but based on the results of the 94632 probe set, it would be safe to assume that the nucleotide at this position is in fact A. Those two probe sets that did not have a correct base assignment for the majority of the strains studied (positions 108551 and 108741) actually failed to assign a base for most of the strains. In some of those cases, the two probes with the highest fluorescent intensity for the set were the two probes with the correct base for the position in question and only differed by the neighbouring site. In this case, the intensities were so close that they were not statistically significantly different and thus no base was assigned. It is interesting to note that the probe set for the neighbouring SNP (eg 108564) had the correct base assignment for that position and for the previous position for which a base had not been assigned. Although these 19 probe sets worked rather well (at least as well as the rest of the probe sets), it may be of value to have one probe made that differs by the position in question and contains a mix of the neighbouring SNP rather than having an entirely different probe made containing the neighbouring mutation. This could potentially simplify the analysis process and the interpretation of the results for these probes. The base assigned is then unambiguous, without the tagged on identity of the neighbouring site. As well, this type of probe may avoid the instance described above in which no base could be assigned because the neighbouring SNP was interfering.

### 4.2.3 Sequencing Verification

In general, the sequencing results agreed very well with the results of the array.

There were three instances in the VZV-Oka strain in which the base assigned by the array differed from the base determined by sequencing. This occurred at positions 105310, 105356 and 107136. For these three positions, the array assigned a C, G and A respectively while sequencing resulted in T, A and G at these three positions. Upon checking the literature and sequences available on GenBank (accession numbers AB097932 and AY016449) it was determined that both nucleotides were possible at those three positions. Further analysis of the sequence chromatogram at position 105310 revealed that although there was a T peak present (in fact there was a T in the position just upstream as well) there was a smaller but distinct C peak present as well. Perhaps the Oka-vaccine that we had was in fact a mixed population at those sites. The use of the ANOVA data analysis method described above that looks for statistical difference within the entire probe set rather than just the two highest as done here, may help to catch mixed populations and recognize them as such rather than just attempting to assign one base at each position.

### 4.2.4 Summary of VZV SNP Array

The SNP genotyping array described here performed very well and accurately assigned bases to the strains studied. With a few adjustments and improvements to both the protocol and the analysis method, this array promises to be a very accurate, reliable and fast method of screening hundreds of SNPs located throughout the VZV genome. Although there are certainly faster, and easier, methods of genotyping VZV, such as RFLP and real-time PCR, the main advantage of microarrays, and the reason for their use here, is their high-throughput. Hundreds of SNPs can be screened in the same assay with the intent to use the results to narrow in on the important SNPs which can then be

routinely screened with other, quicker and easier methods.

As well, the array could be used to screen many VZV strains and the resulting SNP profiles could be used to construct phylogenetic trees for example. This could also be useful for studying the evolution of VZV. By analyzing zoster isolates in elderly patients, which would have been caused by a virus picked up decades earlier on average, the SNP profiles of strains circulating at the time of acquisition can be determined. As well, certain SNPs may be elucidated that are more informative than others for differentiating VZV strains. These could be used to identify the source of an outbreak for example or determine if varicella in a vaccinated individual was caused by wild type virus or vaccine. Another application would be to screen strains with known phenotypes, for example VZV-MSP or BC with a more virulent phenotype and VZV-Oka, an attenuated strain, to see if certain SNPs are associated with the phenotype.

## 4.3 Summary and Conclusions

The methods developed here in the form of microarrays will provide the technology with which to study various aspects of VZV pathogenesis and virology. A powerful investigation tool would be to use both array types in conjunction and assimilate the results. By doing so, various VZV strains, with documented phenotypes, could be screened by the SNP array to determine their SNP profile or fingerprint. The transcription profile of these same strains could then be determined with the gene array and then it could be ascertained if the presence of a specific SNP correlates with a difference in transcription. This information could be used, in conjunction with phenotype information, to narrow in on genes or mutations that are responsible for the given phenotype. This methodology is a powerful tool with which to study VZV

virulence and attenuation and once perfected, promises to elucidate the molecular mechanisms by which VZV attenuation can be achieved. Once this has been determined, strains of VZV could be engineered that contain desired mutations vaccine development, gene therapy or other uses. In addition, insight into VZV pathogenesis can provide information with which to better treat and prevent both varicella and zoster.

## Appendix 1

**Table 11.** PCR primers for generating microarray probes.

| Pair | ORF | Primer 1 (5' to 3') | Primer 2 (5' to 3') | Product Size |
|---|---|---|---|---|
| 1 | 10 | TTCAGGAGGTCCAGGATTCG | TACACCCTCTACCCCAATGAC | 796 |
| 2 | 29 | TCTTTCAGGCAGGGAATTGG | TACAGGTCATATGGGTAAGGAC | 743 |
| 3 | 62 | GGGAGTGGGACCTTAACCTTC | ATGGATACGCCGCCGATG | 773 |
| 4 | 10 | ATTACCGCCGCCGAGAAC | TACACCCTCTACCCCAATGAC | 358 |
| 5 | 29 | CAGAGACTTGGAGGAGTTACAC | GCGTGTGGGCTTTAATTGG | 351 |
| 6 | 62 | CCATTAGATCCTGACCGTCCTC | TCCTTTGAAGGCTGCGAGAG | 359 |
| 7 | 62 | GACGAGGCGGCACATAGC | CACGAGAAAAGGAGGGGACTC | 704 |
| 8 | 62 | CCGCAGACGACAGAGAAC | CCGTTCCGCTTTTATTAACAAC | 351 |
| 9 | 62 | CCGCAGACGACAGAGAAC | CACGAGAAAAGGAGGGGACTC | 303 |

**Table 12.** PCR primers for generating a T7 RNA polymerase promoter on targeted regions. A T7 RNA polymerase promoter was included on the 5' on the reverse (with respect to coding direction) primer and is shown in bold.

| Primer Name | Sequence (5' to 3') (T7 promoter in blue) | Position | Product Length |
|---|---|---|---|
| VZV PstI F | CGGGTGAACCGTATTCTGAG | complement 69580 - 69599 | 350 bp |
| VZV PstI R T7 | **TAATACGACTCACTATAG**TTGAACAATCACGAACCGTT | 69250 - 69269 | |
| VZV ORF54 F | CGCTAATTAAATAAACGGAGAAGG | complement 96240 - 96263 | 2686 bp |
| VZV ORF54 R T7 | **TAATACGACTCACTATAG**TTTTACGCGCAGTGATACAACAG | 93578 - 93600 | |
| VZV IE62 F | AGGCTGCGAGAGCGTTTGGAAAAC | complement 109163 - 109186 | 4019 bp |
| VZV IE62 R T7 | **TAATACGACTCACTATAG**CAGTGGCGCTCACGAGAAAAGGAG | 105168 - 105191 | |
| VZV 3B3 F | ATAACGAATCCGGTCAGAGCATCC | 115877 - 115900 | 647 bp |
| VZV 3B3 R T7 | **TAATACGACTCACTATAG**TTTTCCGCGCAATCCACATCC | complement 116503 - 116523 | |

**Table 13.** Primers used for sequencing IE62.

| Primer Name | Sequence (5' to 3') | Position in ORF 62 cds |
|---|---|---|
| VZVIE62-354F | AGGATCTAATGGGCTCGCCTGTAA | Forward; 354 - 377 |
| VZVIE62-868F | CGGCCCGGTTGAGCAATTGTACCA | Forward; 868 - 891 |
| VZVIE62-1338F | CGAGGATTCGTAAGACCAAACGTC | Forward; 1338 - 1361 |
| VZVIE62-1867F | GGCCGCGGCCGTGGCTATGA | Forward; 1867 - 1886 |
| VZVIE62-2311F | AAACGGGGGATTTCGGCGTATTCC | Forward; 2311 - 2334 |
| VZVIE62-2819F | TTATTACTGTCGACCCGAGACCTG | Forward; 2819 - 2842 |
| VZVIE62-3313F | CTTTGTGGACGAGGCGGCACATAG | Forward; 3313 - 3336 |
| VZVIE62-3827F | TCATCCTTTGGGGTGAGCATCGTG | Forward; 3827 - 3850 |
| VZVIE62-108R | GCGCGTTGGGGTGTAGAGC | Reverse; 90 - 108 |
| VZVIE62-624R | GACCGCTGGTCTTCCCGTTGATCG | Reverse; 601 - 624 |
| VZVIE62-1112R | GACTCTTTGGCTTTTTCTCCACTG | Reverse; 1089 - 1112 |
| VZVIE62-1601R | CGGGTTCCGTTGAGGCCTCGTAAC | Reverse; 1578 - 1601 |
| VZVIE62-2104R | TCGTAAACGATCATCCGGTGGACA | Reverse; 2081 - 2104 |
| VZVIE62-2616R | TCGTAGATGATCAGAAGCCTCACA | Reverse; 2593-2616 |
| VZVIE62-3123R | AGGTTGGCAAACGCAGTCTCGATA | Reverse; 3100 - 3123 |
| VZVIE62-3608R | AACCGGGCGGCCACATTACTCTGG | Reverse; 3585 - 3608 |

## Appendix 2

**Table 14.** Oligonucleotide probes for VZV gene array.

| Name | Sequence (5' to 3') | Position in Dumas |
|---|---|---|
| ORF10 | AGAAACTACGGTAGTAGCATCGAAGCTATGATTTTAGCACCTCCGTCCCCATCCGAGATCCTGCCGGGGG | 13282-13351 |
| ORF29 | TCAGGGTCAAACATTTCAGGGTCAACTGTCCCTGGTCTTAAACGACCCCCCGAAGATGACGAACTCTTTG | 54349-54418 |
| ORF62 | GGGACAGCGAGGAGGGGGGCGGGGACGACGGGGACGCACCGGGGTCATCCTTTGGGGTGAGCATCGTGTC | 105418-105349 |
| ORFS/L | ATCTGGAGTACTACACCCAGTACATTGCATAACCTGTCCATTTGCATTTTCAGTTGCGCGGACGCCTTTC | 662-731 |
| ORFS/L-intron | CTCAACTACATGAAACTACTGTCCGGAAGGGGAAGTTGCGCGGACGCCTTTCTCCGGGATCGTGGCCTTG | 550-584 & 713-749 |
| ORF1 | GAAAGGCGTCCGCGCAACTGAAAATGCAAATGGACAGGTTATGCAACGTACTGGGTGTAGTACTCCAGAT | 731-662 |
| ORF2 | CGATGTTTATTACGAGTGCTTGCGGGGAGAACTGTAGACCTGCCAGGCGGAGGAACGTTACACATTACCT | 1251-1320 |
| ORF3 | TCGAGATAAACCATCTGAATCCCCAATGTATGTTGACCTACAGCCGTTTTGCAGCCTCCTAAATTCCCAA | 2139-2208 |
| ORF4 | ATTGGGCAAATGATGAAGCTATTGGCTCCACTCCAGGCGAGGACTCCACAACGTCTAGAACTGTGTATGT | 3963-4032 |
| ORF5 | CATACATTCGGATTCTGGCCTGGGTTGTTGTATGCACGCTCGCTATAGTAGAGCTTATATCTTATGTTAG | 4466-4535 |
| ORF6 | AATGCTCTATCGTACATTTTACAGCGAATACATACAGACACAACGGAAATACACGCAGTATCGGAGTATA | 7578-7647 |
| ORF7 | ACCGCTGAAAGTGAAGTGGTCATGTCTGATGTTGTTCTTTTGGAGAAAACTCTTGGGGTCGTTGCTAAAC | 9045-9114 |
| ORF8 | GGACTTACAACATCCGAGTATAAACGCTTATATTTTTGGAAGATCATCGAAAAGCCGATCAGGCATTATC | 9801-9870 |
| ORF9A | CTGTTCGGCTCGTGGAGTTTATATCGACGGATCAATGATCACCACCCTTTTCTTCTACGCATCCCTTTTG | 10752-10821 |
| ORF9 | ATAAAAATACGACCCCTCGCGTACATCAACCAAACGACTCCAGCGGATCGGAAGATGACTTTGAAGACAT | 11211-11280 |
| ORF11 | CATCCTTTTGGAAAGACCGCGATGATGCAATGCAAACTTTGGGGATTCGACCGACAACGGACGTTTTACC | 15787-15856 |
| ORF12 | GCGTCTGTTATTCCAGATGATTTGCTTAGACGACATTTAAAAACGGGTCCTGCGGTCTCAGCGGGCACAG | 17048-17117 |
| ORF13 | AGCTCGATACAATTTGCGAAATGAATTTCCTCTTTTAACTACAAAGCGTGTTTTTTGGAGGGCCGTCGTG | 18590-18659 |
| ORF14 | TGCTGTCGTTACCCCTGACGTCTATCCATTTCCCAACGTGTCTATAGGTATTATTGATGGACACATAGTA | 19785-19854 |
| ORF15 | TACGTGTATATATCGTAAAAATCGCCATAAGCAATTAAACACTTATACAGATTTAGGTTTTGCCGTTGTA | 21771-21840 |
| ORF16 | GTTCTTAAGTCCCAAAGACCTTAAAGAAAAATTAACCTCGATGCAATTATTTGCGAACATGGAATCTGTG | 22784-22853 |
| ORF17 | ATTTAGGAATTACATACCCTGAATTTTTGGTTGCCTTTGTTCGCTGTCAGACCGATTTGCATACAAGTGA | 24951-25020 |
| ORF18 | CAATGGTTGGAAGAAAAGGTACGGGACAACCCATCCGTTGCAGAAAAATATATACTAATGATTCTTATAG | 26019-26088 |
| ORF19 | CCAATTCTTTTTAAGACTGGCGGCCACCGTGACAACAGAAATCGTAAACCTGCCTAAAATCGCAACTCTT | 28336-28405 |
| ORF20 | ATTACAGGAGTAAATTGGAACTTGGCTAGACAGCGATTGTATCAATGGACCGGCGATTTTCGGGGACTTC | 29197-29266 |
| ORF21 | AGAAGCACTAAAGATTGCATATCCTACCGAATATAATCGCTCTAGCGTGTTTGTGGAACCGTCGTTTGTG | 32480-32549 |
| ORF22 | AATTTACATAAACCCGCTAAAAAGGTTCTCAAATTTAAACCAACTGTAGACGTGCCGGATAAAAC | 35169-35238 |
| ORF23 | CGTTGCTTTAATACGACTTTTAAACGCCAGTGGACCTTTGCAACCTGGTCACCGTGTGGACATCGCTGAT | 42983-43052 |

| Name | Sequence (5' to 3') | Position in Dumas |
|---|---|---|
| ORF24 | AGTTTAACGGCGAAGACAACGTATCCGAAGCCTTTCCGGTAGAGTATATTATGCGTTTAATGGCGAATTG | 43795-43864 |
| ORF25 | TTAATTACAGACACCGATGTTAATCTATTAAATATCGATGCACTGGAGTCAAAGTATTTTCCTGCTGATA | 44414-44483 |
| ORF26 | TTCCATTTAATGATGAGGGTCGATTCATAACCATTGTAAAAGCAGCAGGTTCCGAGGCCGTATATAAACA | 45887-45956 |
| ORF27 | ATGTTCAGCAGCTTAACAACATATACGAATATCGTGTGTTTCTTGCATCCATTTTGGCGCTATC | 46626-46695 |
| ORF28 | TCGATATTGAATGTAAATCAGGAGGATCTAATGAGCTGGCGTTTCCCGATGCAACACATCTGGAGGATCT | 49525-49594 |
| ORF30 | AGATTACTTTGTCTCAGCGTTTAACTTAAAAACCTGCCAATATCACCATGAGCTGTCGTTAACAACATAC | 54983-55052 |
| ORF31 | CTTCTAAAGCAACGTACGTACGAAATAACCACAAAGTTGAAGCCTTTAATGAGGATAAAAATCCACAGGA | 57459-57528 |
| ORF32 | TGGCGCCATATATTATTCCAGACCGAGAACAACCATCATCAGGGGATGAAGATTCTGACACCGATTAAAT | 60130-60199 |
| ORF33 | CCGCTTTGCCGCCTACTAGTAAAATACCAATAAACATCGATCATCGAAAAGACTGTGTCGTGGGTGAAGT | 61957-62026 |
| ORF33.5 | ATCCCAAAATACAACAAGCACCCCACATACAGATACAGTTGGACAAGATGTAAATGCAGTGGAGGCGAGT | 60402-60471 |
| ORF34 | TGGATTAGTTGCTCTTTGTTTAGACGCGCATACCCAGCTTAGTCGTGGAAGTTTAGGAAGAACCGTAATA | 62459-62528 |
| ORF35 | GGCTGTACGCATATCAATGTCGTTTGAAGTAAATTTGGGACGCCGACGCCCCGATTGTGTTTGTATTATA | 64511-64580 |
| ORF36 | TGTCTCCGTTCGTATTATCGTTAGAACAGACACCCCAGCATGCGGCACAAGAACTAAAAACTCTGCTACC | 65678-65747 |
| ORF37 | CTCTTTAGGAGGGGCGTTTCTGGCGGTAGTGGGGTTTGGTATTATCGGATGGATGTTATGTGGAAATTCC | 68494-68563 |
| ORF38 | GGTGTATATGCGTATGTTGGAGTTACCAGAAAGTTTGCAGGAGCTAGTGAGTGGATTATTCGACGGGACG | 69718-69787 |
| ORF39 | TTCACAGCCGGGTTACTTCTGACGATTCTACATCACTTTCGACTTATTATTATGTTATTGTGTGTCTACA | 70948-71017 |
| ORF40 | TTATATATGGGTGACCGTGACGCCGACATAGAGGCTATAATGTTTGATCACACACAATCGGATGTTGCTT | 75233-75302 |
| ORF41 | CATAAACGAGGGATGTCTGCTTTTATTGGCGCTGATTCCAACTTTGTTAGTCCAAGGAGCACACGACGGT | 76479-76548 |
| ORF42 | CATAGCGATGTCACAGCTTGTTCGTGTTATGTATTAAATAAGCCGGTTTTCATCACAATGGATGGAGCCA | 77729-77798 |
| ORF45 | CCGTGTTTGAGGGTATCAAGTCTCGCCTGGAACAGTGGTTTGGGGCAAATTACGTGGATCATGTAAAAGG | 81611-81680 |
| ORF42+45 | CAGCACAGCGGTGTTCGCGTCAAGTCACAACACAAACGGTATCCGAGGTCAAGATTTTAATCTTCTGTTT | 78006-78038 & 81537-81574 |
| ORF43 | GGGATTGATTTTATCATTGCCTGTTTATCTTAGCCCCGGTTTATTCTTTGATGCATTTAACGTTGTAGCG | 78481-78550 |
| ORF44 | TACTGGCTATGGAACCTAATTTAATAGGGCTATGTCCCGCCGGATGGCATGCTCGGATACTTGGCTCTGT | 80937-81006 |
| ORF46 | TGAGGAACAAACAAGTATTCAACAGATTTTAAATACAAACAGACGCTATATAGCACCCGATTTTATTCGC | 82973-83042 |
| ORF47 | CGACAACTTTGCCTCGTTGGAAATAACCACAGCAGTAATCGGAGACTATAGCCTAGTAACATTAAATACG | 83968-84037 |
| ORF48 | CTATTAACAATAGAATCGGAGACTCGTGGTCAGGGCGACAATGCCATTTGGACACTACTCAGACGAAATT | 85033-85102 |
| ORF49 | TTCTTTGAAGACTTTGACTTTGATGAGAATGTAACAGAGGACGCCGATAAATCCACACAACGCCGCCCAC | 86331-86400 |
| ORF50 | ATTAATTTTGGCGCATTATATACACGTGTTAATAGGACCGTCCCTGGGAACGCTCGTGGCCTGTGCTACG | 86947-87016 |
| ORF51 | TGGATCGTTCTGCATGTCGTACCTACCGCCAGTTGTATAACCTGCTTATGAGCCAGCGCCATTCGTTCTC | 90147-90216 |
| ORF52 | ACTGGAGTCTTCAGGAGCAAAATTCGGAAACAAAAGCAATACCGGCATCTTATCTGACATACGGACACAT | 92477-92546 |
| ORF53 | GTTGATAACAACCCCTTTACAAAAACCCGCGTATTGCCGCATTATATGTTTGTTAAGTTACTGAACCCTC | 93109-93178 |
| ORF54 | GGTAAACAATAATCAAGGCCAGGAAATTTCTGTATCATATTCAAATGCGTCAATCTCATTACTAGTTGCG | 94203-94272 |
| ORF55 | CATTATTACCGGATCCACCAGGGTTGCTGCCCAAAATGTTCATGCTAAGTTATCAACGGCTTATGCGAGT | 96322-96391 |

| Name | Sequence (5' to 3') | Position in Dumas |
|---|---|---|
| ORF56 | GGGTTCTTCACTGGACTTTACCGGATCATGAACAAACACTCTACGCATTTACGGGTGGGTCAAGATCAAT | 98770-98839 |
| ORF57 | CGCGAGTTAATTTTGTCGGGACACAACAACGTCGTATTACAGACATACACTGGTAAATGGTCAGACTGCC | 99485-99554 |
| ORF58 | CGTTGATGCAAATGAAAAATTTCCGATTGGGCACGCGGGCTGTATTGAGAAAACCAAAGACGACTATGTA | 100063-100132 |
| ORF59 | TAAGCTAGACCCCCCAACCTTCAAGTTAAATAACATGTCACATCATTACGACACGGAAACGTTCACCCCC | 101007-101076 |
| ORF60 | TTTGGAGCCAAACCGCTATCGGATGTGAGTTTGATTATAACGGAACCGTGCGTGTCATCGGTATATGAGG | 101487-101556 |
| ORF61 | TGTAGAGCCTACAAAAATGTTGATATTAAACTATATAGGTATTTACGGGCGTGATGAGGCGGGATTAAAA | 103604-103673 |
| ORF63/70 | TGCATAGGAGCGCACTGGAATGTGACGTATCTGATGATGGTGGTGAAGACGATAGCGACGATGATGGGTC | 111020-111089 & 118808-118877 |
| ORF64/69 | TTGGATTGTCGCCCGATAATAAATGCATCCGCGCCGAATTATATGACCGCCCCGGGGGAATTTGTCACAG | 111974-112043 & 117854-117923 |
| ORF65 | AATTCTGTTATATTACGCTAATAATTGTCTTTGTATTTGCTATGACGGGAGCGGCCTTTGCCTTGGGATA | 112357-112426 |
| ORF66 | CAATTATGACGTTAACATGGATATACAGTCTTTTAATATATTTGATGGTGTACACGAAACTGAAGCCGAA | 113192-113261 |
| ORF67 | GGATATGTCCACGTTACCCGAAAAGTCCCTTAATGATCCTCCAGAAAATCTACTTATAATTATTCCTATA | 115263-115332 |
| ORF68 | GTCAATACGGTGACGTGTTTAAAGGAGATCTTAATCCAAAACCCCAAGGCCAAAGACTCATTGAGGTGTC | 116262-116331 |
| hu18SrRNA | GAAAGCATTTGCCAAGAATGTTTTCATTAATCAAGAACGAAAGTCGGAGGTTCGAAGACGATCAGATACC | 995-1064 |
| huGAPDH | TGGGTGTGAACCATGAGAAGTATGACAACAGCCTCAAGATCATCAGCAATGCCTCCTGCACCACCAACTG | 473-542 |
| huß-actin | CCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCAGCCATGTACGTTGCTATCCA | 414-483 |

## Appendix 3

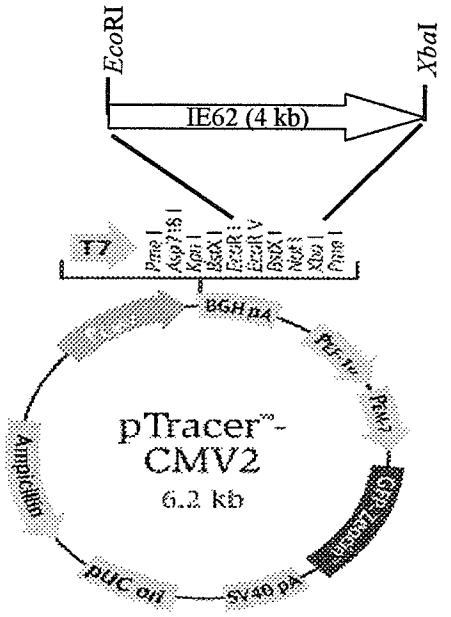



**Figure 15.** IE62 cloning strategy. The schematic depicts the strategy used to clone the 4 kb VZV gene IE62 into the vector using *Eco*RI and *Xba*I restriction sites. Vector map available at http://www.invitrogen.com/content/sfs/vectors/ptracercmv2_map.pdf

## Appendix 4

**Table 15.** Oligonucleotide probes for VZV SNP genotyping array.

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 69349a | GAGACGATATATACAGCAGTTGTTGCGGTA |
| 69349c(mutant) | GAGACGATATATACCGCAGTTGTTGCGGTA |
| 69349g | GAGACGATATATACGGCAGTTGTTGCGGTA |
| 69349t(DumasPstI) | GAGACGATATATACTGCAGTTGTTGCGGTA |
| 94167a(Dumas) | GGCCACCCGATTAGGATTCTTGGTAA |
| 94167c | GGCCACCCGATTCGGATTCTTGGTAA |
| 94167g(mutant) | GGCCACCCGATTGGGATTCTTGGTAA |
| 94167t | GGCCACCCGATTTGGATTCTTGGTAA |
| 94632a | AAACTTCGGTGGTAAACAGTATCAACGG |
| 94632a(94641a) | AAACATCGGTGGTAAACAGTATCAACGG |
| 94632c(mutant) | AAACTTCGGTGGTCAACAGTATCAACGG |
| 94632c(94641a) | AAACATCGGTGGTCAACAGTATCAACGG |
| 94632g | AAACTTCGGTGGTGAACAGTATCAACGG |
| 94632g(94641a) | AAACATCGGTGGTGAACAGTATCAACGG |
| 94632t(Dumas) | AAACTTCGGTGGTTAACAGTATCAACGG |
| 94632t(94641a) | AAACATCGGTGGTTAACAGTATCAACGG |
| 94641a(mutant) | AGTGCGTTTAAAACATCGGTGGTTAACAG |
| 94641a(94632c) | AGTGCGTTTAAAACATCGGTGGTCAACAG |
| 94641c | AGTGCGTTTAAAACCTCGGTGGTTAACAG |
| 94641c(94632c) | AGTGCGTTTAAAACCTCGGTGGTCAACAG |
| 94641g | AGTGCGTTTAAAACGTCGGTGGTTAACAG |
| 94641g(94632c) | AGTGCGTTTAAAACGTCGGTGGTCAACAG |
| 94641t(Dumas) | AGTGCGTTTAAAACTTCGGTGGTTAACAG |
| 94641t(94632c) | AGTGCGTTTAAAACTTCGGTGGTCAACAG |
| 95241a(Dumas) | GGCCAGAGGAGCGGCGGTT |
| 95241c | GGCCAGAGGCGCGGCGGTT |
| 95241g(mutantBglI) | GGCCAGAGGGGCGGCGGTT |
| 95241t | GGCCAGAGGTGCGGCGGTT |
| 95300a(mutant) | AAATACGTATATGAGTCAAGCGGCAGAA |
| 95300c | AAATACGTATATGCGTCAAGCGGCAGAA |
| 95300g(Dumas) | AAATACGTATATGGGTCAAGCGGCAGAA |
| 95300t | AAATACGTATATGTGTCAAGCGGCAGAA |
| 95546a | GGAACGTTATGGGATATGTGGAGAATCA |
| 95546c(Dumas) | GGAACGTTATGGGCTATGTGGAGAATCA |
| 95546g | GGAACGTTATGGGGTATGTGGAGAATCA |
| 95546t(mutant) | GGAACGTTATGGGTTATGTGGAGAATCA |
| 95601a(Dumas) | AGCAGATTGGGAACACCACATCCAAC |
| 95601c(mutant) | AGCAGATTGGGACCACCACATCCAAC |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 95601g | AGCAGATTGGGAGCACCACATCCAAC |
| 95601t | AGCAGATTGGGATCACCACATCCAAC |
| 105234a | GGAGGGCGAGGGGGGAG |
| 105234c(Dumas) | GGAGGGCGCGGGGGGAG |
| 105234g | GGAGGGCGGGGGGGGAG |
| 105234t(mutant) | GGAGGGCGTGGGGGGAG |
| 105264a | TCCTCGTCCTCAGAGGACGA |
| 105264c(Dumas) | TCCTCGTCCTCCGAGGACGA |
| 105264g(mutant) | TCCTCGTCCTCGGAGGACGA |
| 105264t | TCCTCGTCCTCTGAGGACGA |
| 105310a | GGTGGGTTAGCGCCCGGC |
| 105310a(105312c) | GGTGGGCTAGCGCCCGGC |
| 105310c(mutant) | GGTGGGTTCGCGCCCGGC |
| 105310c(105312c) | GGTGGGCTCGCGCCCGGC |
| 105310g | GGTGGGTTGGCGCCCGGC |
| 105310g(105312c) | GGTGGGCTGGCGCCCGGC |
| 105310t(Dumas) | GGTGGGTTTGCGCCCGGC |
| 105310t(105312c) | GGTGGGCTTGCGCCCGGC |
| 105312a | CGGGTGGGATTGCGCCCG |
| 105312a(105310c) | CGGGTGGGATCGCGCCCG |
| 105312c(mutant) | CGGGTGGGCTTGCGCCCG |
| 105312c(105310c) | CGGGTGGGCTCGCGCCCG |
| 105312g | CGGGTGGGGTTGCGCCCG |
| 105312g(105310c) | CGGGTGGGGTCGCGCCCG |
| 105312t(Dumas) | CGGGTGGGTTTGCGCCCG |
| 105312t(105310c) | CGGGTGGGTTCGCGCCCG |
| 105356a(Dumas) | TGGGGTGAGCATCGTGTCGGT |
| 105356c | TGGGGTGAGCCTCGTGTCGGT |
| 105356g(mutant) | TGGGGTGAGCGTCGTGTCGGT |
| 105356t | TGGGGTGAGCTTCGTGTCGGT |
| 105371a | ACCGGGGTCAACCTTTGGGGTG |
| 105371c | ACCGGGGTCACCCTTTGGGGTG |
| 105371g(mutant) | ACCGGGGTCAGCCTTTGGGGTG |
| 105371t(Dumas) | ACCGGGGTCATCCTTTGGGGTG |
| 105406a(Dumas) | CAGCGAGGAGGGGGGCGG |
| 105406a(105413g) | CGGCGAGGAGGGGGGCGG |
| 105406c | CAGCGAGGCGGGGGGCGG |
| 105406c(105413g) | CGGCGAGGCGGGGGGCGG |
| 105406g(mutant) | CAGCGAGGGGGGGGGCGG |
| 105406g(105413g) | CGGCGAGGGGGGGGGCGG |
| 105406t | CAGCGAGGTGGGGGGCGG |
| 105406t(105413g) | CGGCGAGGTGGGGGGCGG |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 105413a(Dumas) | GCGTGGGACAGCGAGGAGG |
| 105413a(105406g) | GCGTGGGACAGCGAGGGGG |
| 105413c | GCGTGGGACCGCGAGGAGG |
| 105413c(105406g) | GCGTGGGACCGCGAGGGGG |
| 105413g(mutant) | GCGTGGGACGGCGAGGAGG |
| 105413g(105406g) | GCGTGGGACGGCGAGGGGG |
| 105413t | GCGTGGGACTGCGAGGAGG |
| 105413t(105406g) | GCGTGGGACTGCGAGGGGG |
| 105451a | CGCTTTGGAGCAGGACGATTGGG |
| 105451c(mutant) | CGCTTTGGAGCCGGACGATTGGG |
| 105451g | CGCTTTGGAGCGGGACGATTGGG |
| 105451t(Dumas) | CGCTTTGGAGCTGGACGATTGGG |
| 105490a(Dumas) | GGGCGACAGCGGCGCA |
| 105490c | GGGCGACCGCGGCGCA |
| 105490g(mutant) | GGGCGACGGCGGCGCA |
| 105490t | GGGCGACTGCGGCGCA |
| 105510a(Dumas) | AACGGCCGTCAGAAACCCAGG |
| 105510a(105512g) | AACGGCCGGCAGAAACCCAGG |
| 105510c | AACGGCCGTCCGAAACCCAGG |
| 105510c(105512g) | AACGGCCGGCCGAAACCCAGG |
| 105510g(mutant) | AACGGCCGTCGGAAACCCAGG |
| 105510g(105512g) | AACGGCCGGCGGAAACCCAGG |
| 105510t | AACGGCCGTCTGAAACCCAGG |
| 105510t(105512g) | AACGGCCGGCTGAAACCCAGG |
| 105512a | GGAACGGCCGACAGAAACCCAG |
| 105512a(105510g) | GGAACGGCCGACGGAAACCCAG |
| 105512c | GGAACGGCCGCCAGAAACCCAG |
| 105512c(105510g) | GGAACGGCCGCCGGAAACCCAG |
| 105512g(mutant) | GGAACGGCCGGCAGAAACCCAG |
| 105512g(105510g) | GGAACGGCCGGCGGAAACCCAG |
| 105512t(Dumas) | GGAACGGCCGTCAGAAACCCAG |
| 105512t(105510g) | GGAACGGCCGTCGGAAACCCAG |
| 105532a(Dumas) | TCGGACACCACGGGGGCTC |
| 105532c | TCGGACACCCCGGGGGCTC |
| 105532g(mutant) | TCGGACACCGCGGGGGCTC |
| 105532t | TCGGACACCTCGGGGGCTC |
| 105544a | ACACGGTTCGAGTTCGGACACC |
| 105544c(mutant) | ACACGGTTCGCGTTCGGACACC |
| 105544g | ACACGGTTCGGGTTCGGACACC |
| 105544t(Dumas) | ACACGGTTCGTGTTCGGACACC |
| 105699a(Dumas) | CCGCAGAACCACTCGTGCTTCC |
| 105699a(105705g) | CCGCGGAACCACTCGTGCTTCC |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 105699c | CCGCAGAACCCCTCGTGCTTCC |
| 105699c(105705g) | CCGCGGAACCCCTCGTGCTTCC |
| 105699g(mutant) | CCGCAGAACCGCTCGTGCTTCC |
| 105699g(105705g) | CCGCGGAACCGCTCGTGCTTCC |
| 105699t | CCGCAGAACCTCTCGTGCTTCC |
| 105699t(105705g) | CCGCGGAACCTCTCGTGCTTCC |
| 105705a(Dumas) | CCCTGCCGCAGAACCACTCG |
| 105705a(105699g) | CCCTGCCGCAGAACCGCTCG |
| 105705c | CCCTGCCGCCGAACCACTCG |
| 105705c(105699g) | CCCTGCCGCCGAACCGCTCG |
| 105705g(mutant) | CCCTGCCGCGGAACCACTCG |
| 105705g(105699g) | CCCTGCCGCGGAACCGCTCG |
| 105705t | CCCTGCCGCTGAACCACTCG |
| 105705t(105699g) | CCCTGCCGCTGAACCGCTCG |
| 105855a(Dumas) | CCACCGCGCAGCAAACCGA |
| 105855c | CCACCGCGCCGCAAACCGA |
| 105855g(mutant) | CCACCGCGCGGCAAACCGA |
| 105855t | CCACCGCGCTGCAAACCGA |
| 105894a(Dumas) | GCTCGGGGCAGCCGACTTTG |
| 105894c | GCTCGGGGCCGCCGACTTTG |
| 105894g(mutant) | GCTCGGGGCGGCCGACTTTG |
| 105894t | GCTCGGGGCTGCCGACTTTG |
| 105915a | GCGACAGTCACGGGGTCTGG |
| 105915a(105919g) | GCGACGGTCACGGGGTCTGG |
| 105915a(105923a) | GAGACAGTCACGGGGTCTGG |
| 105915a(19g&23a) | GAGACGGTCACGGGGTCTGG |
| 105915c(mutant) | GCGACAGTCCCGGGGTCTGG |
| 105915c(105919g) | GCGACGGTCCCGGGGTCTGG |
| 105915c(105923a) | GAGACAGTCCCGGGGTCTGG |
| 105915c(19g&23a) | GAGACGGTCCCGGGGTCTGG |
| 105915g | GCGACAGTCGCGGGGTCTGG |
| 105915g(105919g) | GCGACGGTCGCGGGGTCTGG |
| 105915g(105923a) | GAGACAGTCGCGGGGTCTGG |
| 105915g(19g&23a) | GAGACGGTCGCGGGGTCTGG |
| 105915t(Dumas) | GCGACAGTCTCGGGGTCTGG |
| 105915t(105919g) | GCGACGGTCTCGGGGTCTGG |
| 105915t(105923a) | GAGACAGTCTCGGGGTCTGG |
| 105915t(19g&23a) | GAGACGGTCTCGGGGTCTGG |
| 105919a(Dumas) | CGCGCGACAGTCTCGGGG |
| 105919a(105915c) | CGCGCGACAGTCCCGGGG |
| 105919a(105923a) | CGCGAGACAGTCTCGGGG |
| 105919a(15c&23a) | CGCGAGACAGTCCCGGGG |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 105919c | CGCGCGACCGTCTCGGGG |
| 105919c(105915c) | CGCGCGACCGTCCCGGGG |
| 105919c(105923a) | CGCGAGACCGTCTCGGGG |
| 105919c(15c&23a) | CGCGAGACCGTCCCGGGG |
| 105919g(mutant) | CGCGCGACGGTCTCGGGG |
| 105919g(105915c) | CGCGCGACGGTCCCGGGG |
| 105919g(105923a) | CGCGAGACGGTCTCGGGG |
| 105919g(15c&23a) | CGCGAGACGGTCCCGGGG |
| 105919t | CGCGCGACTGTCTCGGGG |
| 105919t(105915c) | CGCGCGACTGTCCCGGGG |
| 105919t(105923a) | CGCGAGACTGTCTCGGGG |
| 105919t(15c&23a) | CGCGAGACTGTCCCGGGG |
| 105923a(mutant) | GACCTCGCGAGACAGTCTCG |
| 105923a(105915c) | GACCTCGCGAGACAGTCCCG |
| 105923a(105919g) | GACCTCGCGAGACGGTCTCG |
| 105923a(15c&19g) | GACCTCGCGAGACGGTCCCG |
| 105923c(Dumas) | GACCTCGCGCGACAGTCTCG |
| 105923c(105915c) | GACCTCGCGCGACAGTCCCG |
| 105923c(105919g) | GACCTCGCGCGACGGTCTCG |
| 105923c(15c&19g) | GACCTCGCGCGACGGTCCCG |
| 105923g | GACCTCGCGGGACAGTCTCG |
| 105923g(105915c) | GACCTCGCGGGACAGTCCCG |
| 105923g(105919g) | GACCTCGCGGGACGGTCTCG |
| 105923g(15c&19g) | GACCTCGCGGGACGGTCCCG |
| 105923t | GACCTCGCGTGACAGTCTCG |
| 105923t(105915c) | GACCTCGCGTGACAGTCCCG |
| 105923t(105919g) | GACCTCGCGTGACGGTCTCG |
| 105923t(15c&19g) | GACCTCGCGTGACGGTCCCG |
| 105964a(Dumas) | TGAATACCGGCAGTACGTGCTGC |
| 105964c | TGAATACCGGCCGTACGTGCTGC |
| 105964g(mutant) | TGAATACCGGCGGTACGTGCTGC |
| 105964t | TGAATACCGGCTGTACGTGCTGC |
| 106029a(Dumas) | GGAACGTCGCATACACCGTGTG |
| 106029c | GGAACGTCGCCTACACCGTGTG |
| 106029g(mutant) | GGAACGTCGCGTACACCGTGTG |
| 106029t | GGAACGTCGCTTACACCGTGTG |
| 106247a | GATGGACCCGCTATGTCTCAGTATCA |
| 106247c(mutant) | GATGGACCCGCTCTGTCTCAGTATCA |
| 106247g | GATGGACCCGCTGTGTCTCAGTATCA |
| 106247t(Dumas) | GATGGACCCGCTTTGTCTCAGTATCA |
| 106262a(Dumas) | GAGGTGGCCCAGGGATGGACC |
| 106262c | GAGGTGGCCCCGGGATGGACC |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 106262g(mutantSmaI) | GAGGTGGCCCGGGGATGGACC |
| 106262t | GAGGTGGCCCTGGGATGGACC |
| 106497a(Dumas) | CCCGAGGGGGACTGTCTGTGG |
| 106497c | CCCGAGGGGGCCTGTCTGTGG |
| 106497g(mutant) | CCCGAGGGGGGCTGTCTGTGG |
| 106497t | CCCGAGGGGGTCTGTCTGTGG |
| 106569a(mutant) | TGCCCGGAGAAGACATCAACGG |
| 106569c | TGCCCGGAGACGACATCAACGG |
| 106569g(Dumas) | TGCCCGGAGAGGACATCAACGG |
| 106569t | TGCCCGGAGATGACATCAACGG |
| 106710a | ATCGCCGCACGGCGTCC |
| 106710c(mutant) | ATCGCCGCCCGGCGTCC |
| 106710g | ATCGCCGCGCGGCGTCC |
| 106710t(Dumas) | ATCGCCGCTCGGCGTCC |
| 106905a | CGATGCCACTAGACGGACCGG |
| 106905c(mutant) | CGATGCCACTCGACGGACCGG |
| 106905g | CGATGCCACTGGACGGACCGG |
| 106905t(Dumas) | CGATGCCACTTGACGGACCGG |
| 107026a | GACACCCGACGCGGACGC |
| 107026c(mutant) | GACACCCGCCGCGGACGC |
| 107026g | GACACCCGGCGCGGACGC |
| 107026t(Dumas) | GACACCCGTCGCGGACGC |
| 107070a(Dumas) | AGTCCCAGCCAGTCGAGAGCAG |
| 107070c | AGTCCCAGCCCGTCGAGAGCAG |
| 107070g(mutant) | AGTCCCAGCCGGTCGAGAGCAG |
| 107070t | AGTCCCAGCCTGTCGAGAGCAG |
| 107136a(Dumas) | GCCGCCGCACGCTCTCTT |
| 107136c | GCCGCCGCCCGCTCTCTT |
| 107136g(mutant) | GCCGCCGCGCGCTCTCTT |
| 107136t | GCCGCCGCTCGCTCTCTT |
| 107165a(mutant) | GCCACACAGACTCCCGACC |
| 107165c | GCCACACAGCCTCCCGACC |
| 107165g(Dumas) | GCCACACAGGCTCCCGACC |
| 107165t | GCCACACAGTCTCCCGACC |
| 107252a(Dumas) | GAGCCTTTGCCAGCATGGCATAC |
| 107252c | GAGCCTTTGCCCGCATGGCATAC |
| 107252g(mutant) | GAGCCTTTGCCGGCATGGCATAC |
| 107252t | GAGCCTTTGCCTGCATGGCATAC |
| 107307a(Dumas) | TGAGCCGTCGATACGACCGGG |
| 107307c | TGAGCCGTCGCTACGACCGGG |
| 107307g(mutant) | TGAGCCGTCGGTACGACCGGG |
| 107307t | TGAGCCGTCGTTACGACCGGG |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 107329a | GGCCGAGGCCGTGGC |
| 107329c(Dumas) | GGCCGCGGCCGTGGC |
| 107329g | GGCCGGGGCCGTGGC |
| 107329t(mutant) | GGCCGTGGCCGTGGC |
| 107586a(mutant) | CGCTTTACGTACCGGAGTTGGG |
| 107586c | CGCTTTACGTCCCGGAGTTGGG |
| 107586g(Dumas) | CGCTTTACGTGCCGGAGTTGGG |
| 107586t | CGCTTTACGTTCCGGAGTTGGG |
| 107599a | CGGAACCCGAGCCGCTTTAC |
| 107599a(107607t) | CTGAACCCGAGCCGCTTTAC |
| 107599c(mutant) | CGGAACCCGCGCCGCTTTAC |
| 107599c(107607t) | CTGAACCCGCGCCGCTTTAC |
| 107599g | CGGAACCCGGGCCGCTTTAC |
| 107599g(107607t) | CTGAACCCGGGCCGCTTTAC |
| 107599t(Dumas) | CGGAACCCGTGCCGCTTTAC |
| 107599t(107607t) | CTGAACCCGTGCCGCTTTAC |
| 107607a | GGCCTCAACAGAACCCGTGC |
| 107607a(107599c) | GGCCTCAACAGAACCCGCGC |
| 107607c | GGCCTCAACCGAACCCGTGC |
| 107607c(107599c) | GGCCTCAACCGAACCCGCGC |
| 107607g(Dumas) | GGCCTCAACGGAACCCGTGC |
| 107607g(107599c) | GGCCTCAACGGAACCCGCGC |
| 107607t(mutant) | GGCCTCAACTGAACCCGTGC |
| 107607t(107599c) | GGCCTCAACTGAACCCGCGC |
| 107715a(Dumas) | CGGATCGGCACCCCTCCCAT |
| 107715c | CGGATCGGCCCCCCTCCCAT |
| 107715g(mutant) | CGGATCGGCGCCCCTCCCAT |
| 107715t | CGGATCGGCTCCCCTCCCAT |
| 107797a | ATCATTTTCCCAGCCGCGATCCAG |
| 107797c(mutant) | ATCATTTTCCCCGCCGCGATCCAG |
| 107797g | ATCATTTTCCCGGCCGCGATCCAG |
| 107797t(Dumas) | ATCATTTTCCCTGCCGCGATCCAG |
| 108111a(Dumas) | GGATGACAGCCCAGTGGAGAAAAAGC |
| 108111c | GGATGACAGCCCCGTGGAGAAAAAGC |
| 108111g(mutant) | GGATGACAGCCCGGTGGAGAAAAAGC |
| 108111t | GGATGACAGCCCTGTGGAGAAAAAGC |
| 108189a(mutant) | TGGACGTTCATCGGGGGGC |
| 108189c | TGGACGTTCCTCGGGGGGC |
| 108189g(Dumas) | TGGACGTTCGTCGGGGGGC |
| 108189t | TGGACGTTCTTCGGGGGGC |
| 108551a(Dumas) | AACGACGAGAGGAAATTCACCAGGTTT |
| 108551a(108564g) | GACGACGAGAGGAAATTCACCAGGTTT |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 108551c | AACGACGAGAGGACATTCACCAGGTTT |
| 108551c(108564g) | GACGACGAGAGGACATTCACCAGGTTT |
| 108551g(mutant) | AACGACGAGAGGAGATTCACCAGGTTT |
| 108551g(108564g) | GACGACGAGAGGAGATTCACCAGGTTT |
| 108551t | AACGACGAGAGGATATTCACCAGGTTT |
| 108551t(108564g) | GACGACGAGAGGATATTCACCAGGTTT |
| 108564a(Dumas) | TATACCCCGACAAACGACGAGAGGAA |
| 108564a(108551g) | TATACCCCGACAAACGACGAGAGGAG |
| 108564c | TATACCCCGACACACGACGAGAGGAA |
| 108564c(108551g) | TATACCCCGACACACGACGAGAGGAG |
| 108564g(mutant) | TATACCCCGACAGACGACGAGAGGAA |
| 108564g(108551g) | TATACCCCGACAGACGACGAGAGGAG |
| 108564t | TATACCCCGACATACGACGAGAGGAA |
| 108564t(108551g) | TATACCCCGACATACGACGAGAGGAG |
| 108591a(mutant) | AACGATCAACGAGAAGACCAGCG |
| 108591c | AACGATCAACGCGAAGACCAGCG |
| 108591g(Dumas) | AACGATCAACGGGAAGACCAGCG |
| 108591t | AACGATCAACGTGAAGACCAGCG |
| 108618a | TCTCCTGTAGGAGACGCCGC |
| 108618c(mutant) | TCTCCTGTCGGAGACGCCGC |
| 108618g | TCTCCTGTGGGAGACGCCGC |
| 108618t(Dumas) | TCTCCTGTTGGAGACGCCGC |
| 108741a(Dumas) | CTCCCTTGAGACAGTCTCTCTCGGAA |
| 108741a(108747c) | CTCCCTCGAGACAGTCTCTCTCGGAA |
| 108741c | CTCCCTTGAGACCGTCTCTCTCGGAA |
| 108741c(108747c) | CTCCCTCGAGACCGTCTCTCTCGGAA |
| 108741g(mutant) | CTCCCTTGAGACGGTCTCTCTCGGAA |
| 108741g(108747c) | CTCCCTCGAGACGGTCTCTCTCGGAA |
| 108741t | CTCCCTTGAGACTGTCTCTCTCGGAA |
| 108741t(108747c) | CTCCCTCGAGACTGTCTCTCTCGGAA |
| 108747a | CGATCGCTCCCTAGAGACAGTCTCT |
| 108747a(108741g) | CGATCGCTCCCTAGAGACGGTCTCT |
| 108747c(mutant) | CGATCGCTCCCTCGAGACAGTCTCT |
| 108747c(108741g) | CGATCGCTCCCTCGAGACGGTCTCT |
| 108747g | CGATCGCTCCCTGGAGACAGTCTCT |
| 108747g(108741g) | CGATCGCTCCCTGGAGACGGTCTCT |
| 108747t(Dumas) | CGATCGCTCCCTTGAGACAGTCTCT |
| 108747t(108741g) | CGATCGCTCCCTTGAGACGGTCTCT |
| 108838a | TCAGGATCTAAAGGGCTCGCCTGT |
| 108838c(mutant) | TCAGGATCTAACGGGCTCGCCTGT |
| 108838g | TCAGGATCTAAGGGGCTCGCCTGT |
| 108838t(Dumas) | TCAGGATCTAATGGGCTCGCCTGT |

| Oligo Name | Sequence (5' to 3') |
|---|---|
| 108951a | GGACGGGAACAAGAGATCGTTTCA |
| 108951c(Dumas) | GGACGGGAACACGAGATCGTTTCA |
| 108951g | GGACGGGAACAGGAGATCGTTTCA |
| 108951t(mutant) | GGACGGGAACATGAGATCGTTTCA |
| 109010a | GGTGGTCACCACGAGTCAGCCT |
| 109010c | GGTGGTCACCCCGAGTCAGCCT |
| 109010g(mutant) | GGTGGTCACCGCGAGTCAGCCT |
| 109010t(Dumas) | GGTGGTCACCTCGAGTCAGCCT |
| 109044a | GGCCGCAGCGGCCGC |
| 109044c(mutant) | GGCCGCCGCGGCCGC |
| 109044g(Dumas) | GGCCGCGGCGGCCGC |
| 109044t | GGCCGCTGCGGCCGC |
| 109137a | AATTCACCCCAGAGCGATGGATACG |
| 109137c(mutant) | AATTCACCCCAGCGCGATGGATACG |
| 109137g | AATTCACCCCAGGGCGATGGATACG |
| 109137t(Dumas) | AATTCACCCCAGTGCGATGGATACG |
| 116255a(mutant3B3) | AAATTGTAAATGTGAACCAACGTCAATACG |
| 116255c | AAATTGTAAATGTGCACCAACGTCAATACG |
| 116255g(Dumas) | AAATTGTAAATGTGGACCAACGTCAATACG |
| 116255t | AAATTGTAAATGTGTACCAACGTCAATACG |

## REFERENCES


Ahn JW, KL Powell, P Kellam, and DG Alber (2002) Gammaherpesvirus lytic gene expression as characterized by DNA array. *Journal of Virology* 76:6244-6256.

Applebaum E, MH Rachelson and VB Dolgopol (1953) Varicella encephalitis. *American Journal of Medicine*15:523.

Argaw T, JI Cohen, M Klutch, K Lekstrom, T Yoshikawa, Y Asano, and PR Krause (2000) Nucleotide sequences that distinguish Oka vaccine from parental Oka and other varicella-zoster virus isolates. *The Journal of Infectious Diseases* 181:1153-1157.

Arvin AM (2001) Varicella-zoster virus. In: Fields Virology 4th edition, DM Knipe and PM Howley (ed.), Lippincott Williams & Wilkins, pp 2731-2767.

Barrett-Muir W, K Hawrami, J Clarke, and J Breuer (2001) Investigation of varicella-zoster virus variation by heteroduplex mobility assay. *Archives of Virology Supplement* 17:17-25.

Bodrossy L (In press) Diagnostic oligonucleotide microarrays for microbiology. In: Microarrays and Bioinformatics for Beginners, E.Blalock (ed.), Kluwer Academic Publishers, New York.

Bresnahan WA, and T Shenk (2000) A subset of viral transcripts packaged within human cytomegalovirus particles. *Science* 288:2373-2376.

Brunell PA, CF Geiser, V Novelli, S Lipton, and S Narkewicz (1987) Varicella-like illness caused by live varicella vaccine with acute leukemia. *Pediatrics* 79:922-927.

Burgoon CF Jr, JS Burgoon, and GD Baldridge (1957) The natural history of herpes zoster. *Journal of the American Medical Asoociation* 164:265-269.

Chambers J, A Angulo, D Amaratunga, H Guo, Y Jiang, JS Wan, A Bittner, K Frueh, MR Jackson, PA Peterson, MG Erlander, and P Ghazal (1999) DNA microarrays of the complex human cytomegalovirus genome: Profiling kinetic class with drug sensitivity of viral gene expression. *Journal of Virology* 73:5757-5766.

Choo PW, JG Donahue, JE Manson, and R Platt (1995) The epidemiology of varicella and its complications. *Journal of Infectious Diseases* 172:706-712.

Cohen JI and SE Straus (2001) Varicella-zoster virus and its replication. In: Fields Virology 4th edition, DM Knipe and PM Howley (ed.), Lippincott Williams & Wilkins, pp 2707-2730.

Cohrs RJ, MP Hurley and DH Gilden (2003) Array analysis of viral gene transcription during lytic infection of cells in tissue culture with varicella-zoster virus. *Journal of Virology* 77:11718-11732.

Cole NL and C Grose (2003) Membrane fusion mediated by herpesvirus glycoproteins: the paradigm of varicella-zoster virus. *Reviews in Medical Virology* 13:207-222.

Cronin MT, RV Fucini, SM Kim, RS Masino, RM Wespi, and CG Miyada (1996) Cystic fibrosis mutation detection by hybridization to light-generated DNA probe arrays. *Human Mutation* 7:244-255.

Davison AJ and JE Scott (1986) The complete DNA sequence of varicella-zoster virus. *Journal of General Virol*ogy 67:1759-1816.

de Moragas JM and RR Kierland (1957) The outcome of patients with herpes zoster. *AMA Archives of Dermatology* 75:193-196.

DeRisi J, L Penland, P.O. Brown, ML Bittner, PS Meltzer, M Ray, Y Chen, YA Su, and JM Trent (1996) Use of a cDNA microarray to analyse gene expression. *Nature Genetics* 14:457-460.

Ebrahimi B, BM Dutia, KL Roberts, JJ Garcia-Ramirez, P Dickinson, JP Stewart, P Ghazal, DJ Roy, and AA Nash (2003) Transcriptome profile of murine gammaherpesvirus-68 lytic infection. *Journal of General Virology* 84:99-109.

Faga B, W Maury, DA Bruckner, and C Grose (2001) Identification and mapping of single nucleotide polymorphisms in the varicella-zoster virus genome. *Virology* 280:1-6.

Fairley CK, and E Miller (1996) Varicella-zoster virus epidemiology – a changing scene? *Journal of Infectious Diseases* 174:S314-319.

Gershon AA, DL Sherman, Z Zhu, CA Gabel, RT Ambron, MD Gershon (1994) Intracellular transport of newly synthesized varicella-zoster virus: Final envelopment in the trans-Golgi network. *Journal of Virology* 68:6372-6390.

Gilden DH, AN Dueland, ME Devlin, R Mahalingam, and R Cohrs (1992) Varicella-zoster virus reactivation without rash. *Journal of Infectious Dis*eases 166:S30-S34.

Gilden DH, BR Beinlich, EM Rubinstien, E Stommel, R Swenson, D Rubinstein, and R Mahalingam (1994a) Varicella-zoster virus myelitis: an expanding spectrum. *Neurology* 44:1818-1823.

Gilden DH, RR Wright, SA Schneck, JM Gwaltney Jr, R Mahalingam (1994b) Zoster sine herpete, a clinical variant. *Annals of Neurology* 35:530-533.

Gomi Y, T Imagawa, M Takahashi, and K Yamanishi (2000) Oka varicella vaccine is distinguishable from its parental virus in DNA sequence of open reading frame 62 and its transactivation activity. *Journal of Medical Virology* 61:497-503.

Gomi Y, T Imagawa, M Takahashi, and K Yamanishi (2001) Comparison of DNA sequence and transactivation activity of open reading frame 62 of Oka varicella vaccine and its parental viruses. *Archives of Virology Supplement* 17:49-56.

Gomi Y, H Sunamachi, Y Mori, K Nagaike, M Takahashi, and K Yamanishi (2002) Comparison of the complete DNA sequences of the Oka varicella vaccine and its parental virus. *Journal of Virology* 76:11447-59.

Grose C, and PA Brunel (1978) Varicella-zoster virus: isolation and propagation in human melanoma cells at 36 and 32 degrees C. *Infection and Immunity* 19:199-203.

Grose C, DM Perrotta, PA Brunell, and GC Smith (1979) Cell-free varicella-zoster virus in cultured human melanoma cells. *Journal of General Virology* 43:15-27.

Grose C, WE Friedrichs, and GC Smith (1983) Purification and molecular anatomy of the varicella-zoster virion. *Biken Journal* 26:1-15.

Grose C, R Harson R and S Beck (1995) Computer modeling of prototypic and aberrant nucleocapsids of varicella-zoster virus. *Virology* 214:321-9.

Guess HA, DD Broughton, LJ Melton 3rd and LT Kurland (1986) Population-based studies of varicella complications. *Pediatrics* 78:723-727.

Hacia JG, LC Brody, MS Chee, SPA Fodor, and FS Collins (1996) Detection of heterozygous mutations in *BRCA1* using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nature Genetics* 14:441-447.

Harper DR, N Mathieu, and J Mullarkey (1998) High-titre, cryostable cell-free varicella zoster virus. *Archives of Virology* 143:1163-1170.

Harson R, C Grose (1995) Egress of varicella-zoster virus from the melanoma cell: A tropism for the melanocyte. *Journal of Virology* 69:4994-5010.

Hasseman J (2002a) Aminoallyl labeling of RNA for microarrays. The Institute for Genomic Research Standard Operating Procedure M004:1-8.

Hasseman J (2002b) Microarray labelled probe hybridization. The Institute for Genomic Research Standard Operating Procedure M005:1-8.

Hayakawa Y, S Torigoe, K Shiraki, K Yamanishi, and M Takahashi (1984) Biological and biophysical markers of live varicella vaccine strain (Oka): Identification of clinical isolates from vaccine recipients. *Journal of Infectious Diseases* 149:956-963.

Hondo R, Y Yogo, M Yoshida, A Fujima, and S Itoh (1989) Distribution of varicella-zoster virus strains carrying a PstI-site-less mutation in Japan and DNA change responsible for the mutation. *Japanese Journal of Experimental Medicine* 59:233-237.

Husson-van Vliet J, G Colinet, F Yane, and P Lemoine (1987) A simplified plaque assay for varicella vaccine. *Journal of Virological Methods* 18:113-120.

Jenner RG, M Mar Albà, C Boshoff, and P Kellam (2001) Kaposi's sarcoma-associated herpesvirus latent and lytic gene expression as revealed by DNA arrays. *Journal of Virology* 75:891-902.

Kemble GW, P Annunziato, O Lungu, RE Winter, T Cha, SJ Silverstein and RR Spaete (2000) Open reading frame S/L of varicella-zoster virus encodes a cytoplasmic protein expressed in infected cells. *Journal of Virology* 74:11311-11321.

Kenyon TK, JI Cohen and C Grose (2002) Phosphorylation by the varicella-zoster virus ORF47 protein serine kinase determines whether endocytosed viral gE traffics to the trans Golgi network or recycles to the cell membrane. *Journal of Virology* 76:10980-10993.

Kinchington PR, P Ling, M Pensiero, B Moss, WT Ruyechan, and J Hay (1990) The glycoprotein products of varicella-zoster virus gene 14 and their defective accumulation in a vaccine strain (oka). *Journal of Virology* 64:4540-4548.

LaRussa P, O Lungu, I Hardy, A Gershon, SP Steinberg, and S Silverstein (1992) Restriction fragment length polymorphism of polymerase chain reaction products from vaccine and wild-type varicella-zoster virus isolates. *Journal of Virology* 66:1016-1020.

LaRussa P (2000) Clinical manifestations of varicella. In: Varicella-Zoster Virus, AM Arvin and AA Gershon (ed.), Cambridge University Press, pp 206-219.

Loparev VN, K McCaustland, BP Holloway, PR Krause, M Takayama and DS Schmid (2000) Rapid genotyping of varicella-zoster virus vaccine and wild-type strains with fluorophore-labeled hybridization probes. *Journal of Clinical Microbiology* 38:4315-4319.

Martin JH, DE Dohner, WJ Wellinghoff and LD Gelb (1982) Restriction endonuclease analysis of varicella-zoster vaccine virus and wild-type DNAs. *Journal of Medical Virology* 9:69-76.

McKendrick MW, JI McGill, JE White, and JJ Wood (1986) Oral acyclovir in acute herpes zoster. *British Medical Journal* 293:1529-1532.

Meier JL and SE Straus (1993) Varicella-zoster virus DNA polymerase and major DNA binding protein genes have overlapping divergent promoters. *Journal of Virology* 67:7573-7581.

Moffat JF, L Zerboni, PR Kinchington, C Grose, H Kaneshima, and AM Arvin (1998) Attenuation of the vaccine Oka strain of varicella-zoster virus and role of glycoprotein C in alphaherpesvirus virulence demonstrated in the SCID-hu mouse. *Journal of Virology* 72:965-974.

Moriuchi M, H Moriuchi, SE Straus and JI Cohen (1994) Varicella-Zoster virus (VZV) virion-associated transactivator open reading frame 62 protein enhances the infectivity of VZV DNA. *Virology* 200:297-300.

Munoz L, J Balmana, R Martino, A Sureda, N Rabella, and S Brunet (1998) Abdominal pain as the initial symptom of visceral varicella zoster infection in hematopoietic stem cell transplant recipients. *Medicina Clinica* 111:19-22.

National Advisory Committee on Immunization (NACI) (2002) NACI update to statement on varicella vaccine. *Canada Communicable Disease Report* 28:ACS-3.

Ng L-K, P Sawatzky, IE Martin, and S Booth (2002) Characterization of ciprofloxacin resistance in *Neisseria gonorrhoeae* isolates in Canada. *Sexually Transmitted Diseases* 29:780-788.

Nikkels AF, C Debrus, C Sadzot-Delvaux, J Piette, B Rentier and GE Pierard (1995) Localization of varicella-zoster virus nucleic acids and proteins in human skin. *Neurology* 45:S47-S49.

Oxman MN (2000) Clinical manifestations of herpes zoster. In: Varicella-Zoster Virus, AM Arvin and AA Gershon (ed.), Cambridge University Press, pp 246-275.

Paulose-Murphy M, N-K Ha, C Xiang, Y Chen, L Gillim, R Yarchoan, P Meltzer, M Bittner, J Trent, and S Zeichner. *Journal of Virology* 75:4843-4853.

Peters ACB, J Versteeg, J Lindenman and AM Botts (1978) Varicella and acute cerebellar ataxia. *Archives of Neurology* 35:769.

Peterson CL, L Mascola, SM Chao, JM Lieberman, EL Arcinue, DA Blumberg, KS Kim, A Kovacs, VK Wong and PA Brunell (1996) Children hospitalized for varicella: a prevaccine review. *Journal of Pediatrics* 129:529-536.

Ragozzino MW, LJ Melton 3rd, LT Kurland, CP Chu, and HO Perry (1982a) Population-based study of herpes zoster and its sequelae. *Medicine*. 61:310-316.

Ragozzino MW, LJ Melton 3rd, LT Kurland, CP Chu, HO Perry (1982b) Risk of cancer after herpes zoster: a population-based study. *New England Journal of Medicine* 307:393-397.

Ramsay, G (1997) DNA chips: state-of-the-art. *Nature Biotechnology* 16:40-44.

Rogers RS III and JP Tindall (1972) Herpes zoster in children. *Archives of Dermatology* 106:204-207.

Roizman B and PE Pellet (2001) The family *Herpesviridae*: a brief introduction. In: Fields Virology 4th edition, DM Knipe and PM Howley (ed.), Lippincott Williams & Wilkins, pp 2381-2407.

Ross A, E Lencher, and G Reitman (1962) Modification of chickenpox in family contacts by administration of gamma globulin. *New England Journal of Medicine* 267:369-376.

Sadzot-Delvaux C, L Baudoux, P Defechereux, J Piette and B Rentier (1999) Overview of the replication cycle of varicella-zoster virus. *Contributions to Microbiology* 3:21-42 .s

Santos RA, JA Padilla, C Hatfield and C Grose (1998) Antigenic variation of varicella zoster virus Fc receptor gE: loss of a major B cell epitope in the ectodomain. *Virology* 249:21-31.

Santos RA, CC Hatfield, NL Cole, JA Padilla, JF Moffat, AM Arvin, WT Ruyechan, J Hay and C Grose (2000) Varicella-zoster virus gE escape mutant VZV-MSP exhibits an accelerated cell-to-cell spread phenotype in both infected cell cultures and SCID-hu mice. *Virology* 275:306-317.

Schena M, D Shalon, RW Davis, and PO Brown (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470.

Schena M (1996) Genome analysis with gene expression microarrays. *Bioessays* 18:427-431.

Schmader KE (2000) Epidemiology of herpes zoster. In: Varicella-Zoster Virus, AM Arvin and AA Gershon (ed.), Cambridge University Press, pp 220-245.

Seward J, K Galil and M Wharton (2000) Epidemiology of varicella. In: Varicella-Zoster Virus, AM Arvin and AA Gershon (ed.), Cambridge University Press, pp 187-205.

Stanford ORF Comparisons web page. http://cmgm.stanford.edu/~jjcheng/VZV/

Stingley SW, JJ Garcia Ramirez, SA Aguilar, K Simmen, RM Sandri-Goldin, P Ghazal, and EK Wagner (2000) Global analysis of herpes simplex virus type 1 transcription using an oligonucleotide-based DNA microarray. *Journal of Virology* 74:9916-9927.

Straub TM, DS Daly, S Wunshel, PA Rochelle, R DeLeon and DP Chandler (2002) Genotyping *Cryptosporidium parvum* with an *hsp70* single-nucleotide polymorphism microarray. *Applied and Environmental Microbiology* 68:1817-1826.

Straus SE, JM Ostrove, G Inchauspe, JM Felser, A Freifeld, KD Croen, and MH Sawyer (1988) NIH conference. Varicella-zoster virus infections. Biology, natural history, treatment, and prevention. *Annals of Internal Medicine* 108:221-237. Erratum in: *Annals of Internal Medicine* 1988 109:438-439.

Takahashi M, T Otsuka, Y Okuno, Y Asano, T Yazaki and S Isomura (1974) Live vaccine used to prevent the spread of varicella in children in hospital. *Lancet* 2:1288-1290.

Takayama M, N Takayama, N Inoue and Y Kameoka (1996) Application of long PCR method of identification of variations in nucleotide sequences among varicella-zoster virus isolates. *Journal of Clinical Microbiology* 34:2869-2874.

Taylor-Robinson D (1959) Chickenpox and herpes zoster. III. Tissue culture studies. *British Journal of Experimental Pathology* 40:521-532.

Tipples GA, G Peters, S Tyler, JS Hiebert, G Parkyn, W Jackson, T Wagenaar and C Grose (2002) Sequencing of the first North American VZV genome: VZV-MSP. 27th International Herpesvirus Workshop, Cairns, Australia.

Tipples GA, GM Stephens, C Sherlock, M Bowler, B Hoy, D Cook and C Grose (2002) New variant of varicella-zoster virus. *Emerging Infectious Diseases* 12:1504-1505

Tipples GA, D Safronetz and M Gray (2003) A real-time PCR assay for the detection of varicella-zoster virus DNA and differentiation of vaccine, wild-type and control strains. *Journal of Virological Methods* 113:113-116.

Tyzzer EE (1906) The histology of the skin lesions in varicella. *Journal of Medical Research* 14:361-392.

Wagenaar TR, VT Chow, C Buranathai, P Thawatsupha and C Grose (2003) The out of Africa model of varicella-zoster virus evolution: single nucleotide polymorphisms

and private alleles distinguish Asian clades from European/North American clades. *Vaccine* 21:1072-1081.

Wagner EK, JJ Garcia Ramirez, SW Stingley, SA Aguilar, L Buehler, GB Devi-Rao and P Ghazal (2002) Practical approaches to long oligonucleotide-based DNA microarray: Lessons from herpesviruses. *Progress in Nucleic Acid Research and Molecular Biology* 71:445-491.

Weigle KA and C Grose (1983) Common expression of varicella-zoster viral glycoprotein antigens *in vitro* and in chickenpox and zoster vesicles. *Journal of Infectious Diseases* 148:630-638.

Weller TH. (1953) Serial propagation *in vitro* of agents producing inclusion bodies derived from varicella and herpes zoster. *Proceedings of the Society for Experimental Biology and Medicine* 83:340–346.

Weller TH, HM Witton and EJ Bell (1958) The etiologic agents of varicella and herpes zoster. Isolation, propagation, and cultural characteristics *in vitro*. *Journal of Experimental Medicine* 108:843-868.

Wharton M, L Fehrs, SL Cochi, *et al* (1990) Health impact of varicella in the 1980s. In Abstracts of the 30th Interscience Conference on Antimicrobial Agents and Chemotherapy, Atlanta.

Wharton M (1996) The epidemiology of varicella-zoster virus infections. *Infectious Disease Clinics of North America* 10:571-581.

Wilkinson M (1991) Purification of RNA. In: Essential Molecular Biology: A Practical Approach, Volume 1, T.A. Brown (ed), IRL Press at Oxford University Press, UK, pp 69-72.

Wood MJ (1991) Herpes zoster and pain. *Scandinavian Journal of Infectious Diseases* 78:S53-S61.