Improving the Yeast two-hybrid system with permutated fusion proteins: The Varicella zoster virus protein interaction network

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)
Fakultät für Chemie und Biowissenschaften
Karlsruher Institut für Technologie (KIT) - Universitätsbereich
vorgelegte
DISSERTATION
von
Diplom-Biologe Thorsten Stellberger
aus
Rauenberg

Dekan: Prof. Dr. Stefan Bräse
Referent: Prof. Dr. Jonathan Sleeman
Korreferent: Prof. Dr. Jörg Kämper
Tag der mündlichen Prüfung: 20.10.2010

Die vorliegende Arbeit wurde in der Zeit von Januar 2007 bis September 2010 in der Arbeitsgruppe von PD Dr. Peter Uetz im Institut für Toxikologie und Genetik des Karlsruher Instituts für Technologie (KIT), Campus Nord, angefertigt.

Zusammenfassung

Die vorliegende Studie beschreibt den ersten Versuch zur Erstellung eines vergleichenden Protein-Protein Interaktionsnetzwerks mit dem Yeast two-hybrid (Y2H) System. Viele Studien bestätigen, dass Interaktionsnetzwerke aus proteomweiten Screens unvollständig sind. Dies stützt sich auf die Beobachtung, dass Interaktionsdaten, die mit unterschiedlichen Methoden erhoben wurden, nur geringe Überlappungen zeigen. Dies betrifft auch unterschiedliche Systeme innerhalb einer Methode, wie verschiedene Y2H-Systeme. Dadurch wurde die Frage aufgeworfen, welche Rolle strukturelle Unterschiede, insbesondere sterische Bedingungen im Testsystem spielen, verursacht durch die Orientierung der verwendeten Fusionsdomänen. In dieser Arbeit untersuche ich deren Einfluss auf die Detektierbarkeit von Protein-Protein Interaktionen.

Zunächst habe ich ein Vektorsystem entwickelt, welches die Y2H-Testdomänen an den C-Terminus und nicht an den N-terminus fusioniert, wie es traditionell gemacht wird. Die Ausgangsvektoren pGBKT7g und pGADT7g habe ich zunächst entsprechend umgebaut und wieder für die Hochdurchsatz-Klonierung kompatibel gemacht (Gateway®-Klonierung, Invitrogen, Karlsruhe). Nach der Konversion konnten beide Vektorsysteme kombiniert werden, wodurch vier verschiedene Bait-Prey Kombinationen getestet werden können, mit N-N-, N-C-, C-N- und C-C-terminalen Testdomänen.

Eine Bibliothek aus Gateway® Eingangsvektoren von Varizella Zoster Virus (VZV) wurde in beide Vektorsysteme hineinrekombiniert und zusätzlich zur NN-Topologie auf binäre Proteininteraktionen getestet. Dadurch konnten etwa doppelt so viele Interaktionen identifiziert und die Rate an falsch-negativen Interaktionen gesenkt werden. Ähnliche Ergebnisse wurden bei der Validierung mittels eines humanen Referenz-Sets erzielt. Deshalb empfehle ich, dieses System in zukünftigen Studien routinemäßig anzuwenden.

Zusätzlich zum VZV-Interaktom habe ich ein Subnetzwerk von DNA-Enkapsidations Proteinen analysiert und eine besonders interessante Interaktion näher charakterisiert. Das essentielle, vom offenen Leserahmen Nummer 25 (ORF25) codierte Protein zeigt viele Interaktionen mit DNA-Verpackungsproteinen wie auch mit den meisten anderen viralen Proteinen.

Mithilfe von Peptid-Arrays konnte ich drei Sequenzbereiche Identifizieren, welche die Selbstinteraktion des Proteins vermitteln. Diese Erkenntnisse können in Kombination mit 3-D Strukturen und neuen Methoden des virtuellen *drug-design* bei der Entwicklung antiviraler Therapeutika verwendet werden.

Summary

The study at hand is the first comprehensive attempt to perform a high-throughput protein-protein interaction network with the Yeast two-hybrid (Y2H) system. Reports of incompleteness and low overlaps between protein-protein interaction datasets derived by different methods and variants of the Yeast two-hybrid system, raised the question about the impact of sterical circumstances caused by the fusion tags added to the proteins in order to detect a possible interaction. First, I developed a Y2H vector system with the fusion tags C-terminally fused to the test-constructs. Additionally, the vectors had to be suitable for high-throughput cloning of test libraries. As parental vectors I have used the pGBKT7g bait and pGADT7g prey vector, derivatives of the Clontech MatchMaker™ Y2H system that were previously modified for high-throughput cloning by the Gateway® recombination cloning technology. After the conversion, both vector systems, the traditional N-terminally- and the new C-terminally tagging vector system could be combined to screen four different tag-topologies of bait- and prey-fusion tags. A Gateway® entry-vector library of Varicella zoster virus (VZV), which was recently screened with the progenitor Y2H vectors, was recombinated into both vector systems and additional Y2H screens were performed to gain a complete "combinatorial" network of VZV (A screen repeated with four bait-prey combinations N-N, N-C, C-N and C-C bait-prey fusion tag orientation, respectively). The permutations of N- and C-terminal Y2H vectors achieved an extensive increase of the coverage of this interactome screen, and thus significantly reduce the rate of undetected interactions. Similar results were determined by screening a human reference set. Accordingly, I recommend that future interaction screening projects should use such vector combinations on a routine basis.

In the second part of this study I generated a sub-network of VZV DNA-packaging proteins. Thereupon, I synthesized peptide-arrays to characterize the self-interaction of one of those proteins, encoded by the essential VZV open reading frame number 25 (ORF25). I could identify three interacting peptides within the 156 amino acid protein, which contribute to its self-interaction. These findings provide a basis for modern drug-design in order to identify and develop new antiviral compounds.

Contents

List of Abbreviations

3-AT	3-amino-1,2,4-triazole
5-FOA	5-fluoroorotic acid
AA	amino acid
AB	Antibody
AD	activation domain
AHT	anhydrotetracycline
Amp	ampicillin
APS	ammonium persulfate
att	attachment site
bp	base pair(s)
BSA	bovine serum albumin
CPL	characteristic path length
(H, M) CMV	Cytomegalovirus (prefix H: human, M: murine)
Co-IP	Co-Immunoprecipitation
DBD	DNA binding domain
DIW	deionized water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	1,4-dithiothreitol
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
FBLD	fragment-based lead discovery
FL	Full-length
Fmoc	9-fluorenyl-methoxycarbonyl
GAL4	gene encoding the yeast transcription activator protein Gal4
Gen	gentamycin
GST	glutathione S-transferase
HA	hemagglutinin
HRP	horseradish peroxidase

HSV	Herpes simplex virus
HTP	high-throughput
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
kb	kilo bases
kDa	kilodaltons
KSHV	Kaposi's sarcoma associated herpesvirus
l	liter
LB	lysogeny broth
LC	literature-curated
MAT	mating type locus
µl	microliter
µM	micromolar
M	molar mass
mM	millimolar
nt	nucleotide(s)
OD	optical density
o/n	overnight
ORF	open reading frame
ORFeome	the totality of open reading frames
PBS	Phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
PPI(s)	protein-protein interaction(s)
RT	room temperature
SAP	shrimp alkaline phosphatase
SD	synthetic defined
SDS	sodium dodecyl sulphate
PAGE	polyacrylamide gel electrophoresis
SPPS	solid phase peptide synthesis
TBS	Tris-buffered saline
TEMED	N,N,N',N'-tetramethylethylenediamine

Tris	tris-(hydroxymethyl)-aminomethane
VZV	Varicella zoster virus
w/o	without
WT	wild type
Y2H	Yeast two-hybrid
YPD	yeast extract peptone dextrose

List of Figures

List of Tables

1 Introduction

1.1 Protein-Protein Interactions

Every biological system is composed of a large number of components (e.g. DNA, proteins, lipids and sugars) which mostly function in complex interaction networks, where one component is able to affect a wide range of other components. Protein-protein interactions (PPIs) affect virtually all processes in the cell being connected via an extensive network of non-covalent interactions which are constantly forming and dissociating. PPIs are an essential aspect in practically all biological processes, like generation and maintaining of macromolecular structures, cell signaling, dynamic regulation, and metabolic pathways. Functional relationships of proteins can be revealed by either direct protein-protein interactions or their co-occurrence within protein-complexes. This may allow a first prediction of the function of uncharacterized or hypothetical proteins, predicted computationally after whole genome sequencing of any possible organism (Sharan et al., 2007; Uetz and Finley, 2005; Walhout and Vidal, 2001). In addition to this important aspect in basic research, Protein-protein interactions have been recognized to be important drug targets, which can be targeted by small molecules, binding to the protein contact surfaces (Arrell and Terzic, 2010; Pujol et al., 2010; Xie et al., 2009). Mapping and characterization of PPI networks is therefore a crucial duty in the post-genomic era as they provide the basis for a global understanding of the cellular proteome, and for this reason are one major goal in systems biology (Auerbach et al., 2002).

1.2 From single interactions to the interactome

Protein interactions have been first characterized individually, but this so called "reductionist" approach lacks information about time, space and the context in which an interaction occurs *in vivo* (Chautard et al., 2009b). A novel approach has been developed in the last decade, focusing on the building of protein-protein interaction maps. This approach is part of an upcoming field, called systems biology which can be defined as “the study of an organism, viewed as an integrated and interacting

network of genes, proteins and biochemical reactions which give rise to life" (http://www.systemsbiology.org/).

An interactome comprises the whole set of molecular physical interactions between biological entities in cells and organisms and it is the basis of understanding how gene functions and regulations are integrated at the whole organism level. The exploration of entire protein interactomes (the totality of protein-protein interactions of an organism) was based on the development of technologies like the Yeast two-hybrid method and mass spectrometry which allowed the investigation of PPIs on a high-throughput (HTP) level. In the past two decades, these technologies have generated the data for our understanding of protein interactomes and are crucial to exploit their therapeutic potential for the emerging "post-antibiotic era" (Alanis, 2005).

1.2.1 The Yeast two-hybrid system

The Yeast two-hybrid (Y2H) method, which was originally developed by Stanley Fields, is a genetic method to detect binary protein-protein interactions that exploits the modularity of eukaryotic transcription factors and the well-established genetic engineering of the yeast, *Saccharomyces cerevisiae*, to monitor PPIs (Fields and Song, 1989). A bait protein is fused to the DNA binding domain (DBD) and a prey protein is fused to the activation domain (AD) of a transcriptional activator. The term "two-hybrid" is based on these two chimeric proteins. The bait and prey fusions are co-expressed in yeast and upon physical interaction between the bait and prey protein, the functional transcription factor (TF) is reconstituted. This results in the activation of a reporter gene, which allows either growth under selective conditions or produces a color or fluorescence signal (auxotrophic yeast strain, lacZ or GFP reporter gene). Figure 1 shows the basic principle of the assay. Hence, a bait protein can be screened against prey libraries expressing all encoded or expressed proteins, derived either from genomic or cDNA library, in the organism or tissue of interest. The Yeast two-hybrid system can be used for the detection of virtually any protein-protein interaction, independent of the function of the corresponding proteins.

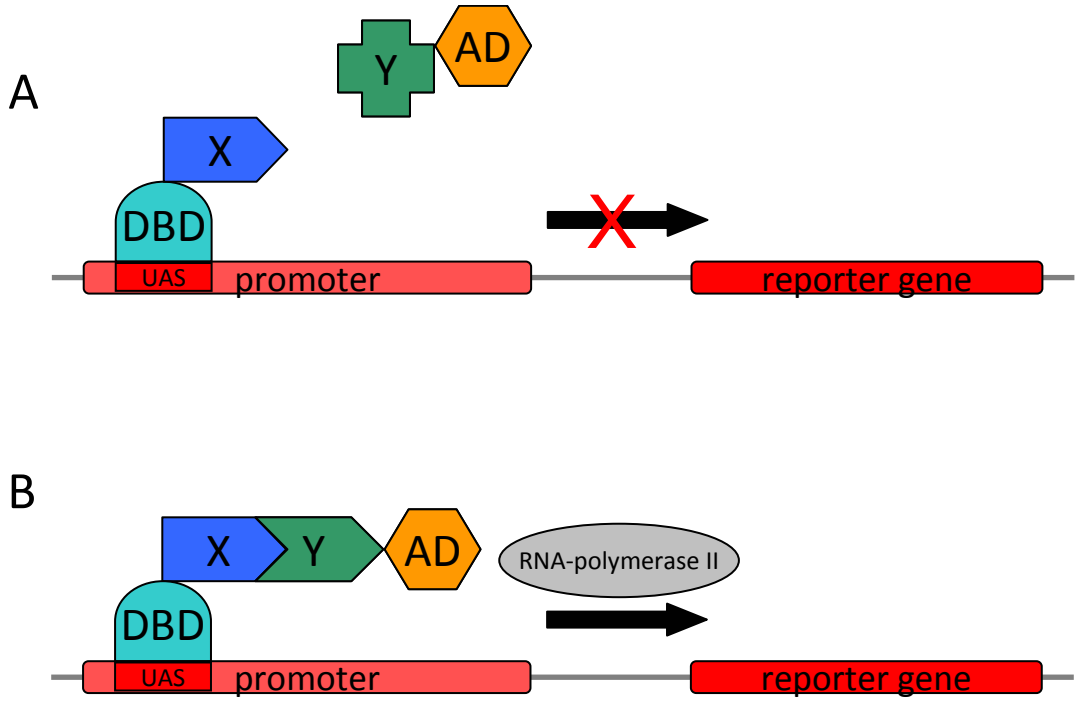


Figure 1: Yeast two-hybrid principle.
A) The protein of interest X is fused to the DNA binding domain (DBD), representing the so called "bait" construct. The potential interacting protein Y is fused to the activation domain (AD) and is called "prey". If both fusion proteins are co-expressed in the same yeast cell, they are translocated in the nucleus. If both constructs do not interact, no reporter gene expression is driven. **B)** The bait, i.e. the DBD-X fusion protein, binds the upstream activator sequence (UAS) of the promoter. Upon interaction of bait and prey, the AD-Y fusion protein, recruits the AD and thus reconstitutes a functional transcription factor. This is in turn leading to the recruitment of cellular RNA polymerase II and subsequent transcription of the reporter gene.

1.2.2 Protein-interactome studies performed with the Y2H system

The first protein-interactome to be mapped was that of the bacteriophage T7 (Bartel et al., 1996), followed by the first complex, free-living organism, *Saccharomyces cerevisiae*, which was published by Peter Uetz and coworkers in 2000 (Uetz et al., 2000a) and shortly after by Takashi Ito (Ito et al., 2001). Both studies were merged into one high confidence interactome in 2008 by Yu and colleagues (Yu et al., 2008). The efforts on the yeast interactome triggered al lot of projects, often still ongoing, that aim the mapping of genome-wide protein interactomes, for example of bacteria (*Helicobacter pylori*, *Campylobacter jejuni*, *Treponema pallidum* (Parrish et al., 2007; Rain et al., 2001; Rajagopala et al., work in progress; Rajagopala et al., 2007; Titz et al., 2008) as well as *Streptococcus pneumoniae* and *Escherichia coli* (Rajagopala et al., work in progress) of plants like *Arabidopsis thaliana* (Hackbusch et al., 2005) and Cotton (*Gossypium hirsutum*) (Zhang et al., 2010b) and animals, e.g. the malaria

parasite *Plasmodium falciparum* (LaCount et al., 2005), *Caenorhabditis elegans* (Li et al., 2004; Simonis et al., 2009) and *Drosophila melanogaster* (Giot et al., 2003).
Additionally, a number of focused interactome sub-networks have been generated, based on human diseases such as Huntington's disease (Goehler et al., 2004) and human inherited ataxias (Lim et al., 2006) or viral-host pathogen networks like the ones for Hepatitis-C-Virus (de Chassey et al., 2008) and the herpesviruses VZV and KSHV (Uetz et al., 2006) and EBV (Calderwood et al., 2007) as well as intraviral protein interaction screens of herpesviruses (Calderwood et al., 2007; Fossum et al., 2009; Stellberger et al., 2010). Huge efforts are being made on the mapping of the human interactome. In 2005, two publications have described first subsets of the human interactome (Rual et al., 2005; Stelzl et al., 2005), with larger subsets being in the screening process.
The reasons for the common use of the Y2H are the relatively low costs and the convenient use. Thus, a large amount of the data so far generated from protein interaction studies have come from Y2H screening. For example, 6,124 protein interactions are listed for Saccharomyces cerevisiae in the MINT database (Ceol et al., 2010) and about 6,721 for humans, creating huge protein interaction networks.

1.2.3 Biological databases

The large numbers of protein interaction studies, derived either from small- or large-scale studies have yielded hundreds of thousands of interactions. These interactions are collected in biomolecular interaction databases that allow the interactions to be assembled and be provided to the scientific community. The most important databases are listed in the following:

name	address	reference	# interactions listed
DIP	http://dip.doe-mbi.ucla.edu	(Salwinski et al., 2004)	70,411
IntAct	www.ebi.ac.uk/intact	(Aranda et al., 2010)	227,866
MINT	http://mint.bio.uniroma2.it/mint	(Ceol et al., 2010)	86,506
MPact	http://mips.gsf.de/genre/proj/mpact	(Pagel et al., 2005)	n.A.
MatrixDB	http://matrixdb.ibcp.fr/	(Chautard et al., 2009a)	1,836 PPIs and 119 protein-glycosaminoglycan interactions.
MPIDB	http://www.jcvi.org/mpidb	(Goll et al., 2008)	24,295 microbial PPIs
BioGRID	http://www.thebiogrid.org/	(Breitkreutz et al., 2008)	177,804
InnateDB	www.innatedb.com	(Lynn et al., 2008)	115,000+
BIND	www.blueprint.org	(Bader et al., 2003)	n/a

The data of all above listed databases except of BIND can be addressed via a single search interface at the IMEx consortium (http://www.imexconsortium.org/) (Orchard et al., 2007). Approximately half of the interaction data available on those databases are coming from Y2H assays, followed by a combination of affinity purification followed by mass spectrometry (AP-MS) (Bruckner et al., 2009).

1.2.4 Y2H screening methods

Two basic screening approaches can be distinguished: the matrix-based (or array-based) and the library-screen.

1.2.4.1 Matrix-based Y2H screens

In a matrix-screen, the possible combinations between open reading frames (ORFs) are systematically examined by performing direct mating of a set of baits versus a set of preys expressed in opposite yeast mating types, e.g. mating type *a* for baits and mating type α for preys .The defined position of each prey protein in a matrix allows rapid identification of interacting preys without sequencing, but screens are not restricted to a limited set of full-length ORFs, proteins can be split or divided into defined domains to decrease the rate of false-negative interactions as well as to map the interactions. Today, matrix-based screens are used mostly for smaller and medium size clone collections in combination with automation and cloning techniques that allow for reliable and fast interaction screening. However, the capacity of matrix-based screens is limited by the size of the clone set to be tested. For example, a small proteome that encodes for 1,000 proteins requires at least $1{,}000^2$ (one million) individual pairwise tests in one comprehensive screen. For large genomes such as the human one, $23{,}000^2$ (over half a billion) pairwise tests would be necessary to test all possible combinations. Nevertheless, the human interactome is being mapped by a matrix-based screening strategy (http://ccsb.dfci.harvard.edu/web/www/ccsb/groups/ Interactomegroup.html).

Genome-wide screens face three main issues: efficiency (hands-on time), specificity (detecting false-positives), and sensitivity (avoiding false-negatives). The need to make large-scale matrix-screens more efficient, led to the development of pooling strategies which can drastically reduce the number of individual Y2H tests while keeping the resulting sequencing efforts reasonable while conserving the advantages of a matrix-based screen (Jin et al., 2007; Jin et al., 2006; Xin et al., 2009; Zhong et

al., 2003). “Smart” pooling and arrangements of prey as well as bait clones can help to speed up the screening procedure drastically, resulting in interaction detection with nearly the same sensitivity and specificity as one-on-one Yeast two-hybrid screens.

1.2.4.2 Library screens

The classical cDNA-library screen is searching for pairwise interactions between a defined protein of interest (bait) and potential interaction partners (preys), present in cDNA libraries or sub-pools of libraries. A defined bait or bait-pool is mated against the whole prey pool represented by the used cDNA library and plated onto the appropriate readout medium where positive interactants are selected. Alternatively, prey plasmid libraries can be transformed directly into the haploid bait reporter strain.
In contrast to the matrix-based strategy this approach requires identification of the interacting prey (and bait if pooled) by colony PCR analysis and subsequent sequencing, making such screens more expensive. However, this procedure is also prone to produce false-negatives due to subsets of preys being underrepresented in the pool while others may be overrepresented which is almost impossible to monitor w/o adequate efforts. One possibility to control the composition of prey libraries is pooling a yeast prey matrix, thereby achieving a degree of normalization (minimizing under- or overrepresentation of preys in a pool).
Since most library screens use cDNA or even random genomic libraries, false-positives may result from fragments that do not fold properly or that expose protein sequences that are not exposed *in vivo*. On the other hand, libraries may contain cDNA fragments in addition to full-length ORFs, thus largely covering a transcriptome and reducing the rate of false-negatives. However, inherent to this type of library screening, the rate of wrongly identified proteins (called false-positives) is increased. Clearly, both library- and matrix-screens do have advantages and disadvantages that should be considered depending on the project that is planned.

1.2.5 Limitations of the Yeast two-hybrid system

1.2.5.1 False-negative interactions

Yeast two-hybrid screens do not generate complete protein interactomes. As for any other detection method, it is almost impossible to detect all physiologically occurring interactants of every screened bait protein. Those false-negative interactions may be traced back to steric hindrance due to the two fusion tags, preventing physical interaction by covering interaction sites or preventing subsequent transcriptional activation. Another source for false-negatives is instability of proteins due to un- or improper folding in the yeast cell or due to harboring a PEST sequence (a peptide sequence which is rich in proline (P), glutamic acid (E), serine (S), and threonine (T)). These sequences are associated with proteins that have a short intracellular half-life and regarded as a signal peptide for protein degradation (Rogers et al., 1986). An additional factor that affects protein stability is the nature of the N-terminal amino acid residue, according to the N-end rule which was originally described in *S. cerevisiae* (Bachmair et al., 1986). This should play a minor role in the Yeast two-hybrid as the constructs bear a stabilizing Methionine at their N-termini. Apart from that it is difficult to monitor potentially cleaving events in the yeast cell in a HTP-system. But the fact that the Yeast two-hybrid System works with fusion proteins can also lead to missed interactions. The standard vector systems generate N-terminal fusion tags. If the interacting domain of a protein is at, or near its N-terminus, the fusion of the DNA-binding or activation domain may prevent an interaction. Other error sources may be the failure of nuclear localization, the absence of a prey protein from a library or an improper post-translational modification of a bait- or prey-construct, like a phosphorylation that is indispensable for an interaction. It was estimated that the false-negative rate in array-based Yeast two-hybrid screens lies at about 75 % which means that three quarters of all interactions may be missed (Rajagopala et al., 2007). The assay sensitivity, which is the fraction of all physical interactions that take place in a given organism, would be in reverse 25 %. For Y2H library screens, the assay sensitivity was estimated to be somewhat lower, about 20 % based on the Uetz *S. cerevisiae* screen from 2000 (Yu et al., 2008). This large number of undetected PPIs is a major handicap in the understanding of biological processes at the whole level but still bears a lot of information that can be exploited for the understanding of

cellular processes. The challenge is the reduction of false-negatives. For example, screening of related species can reveal different subsets of the interactome in regard of orthologous processes involving orthologous proteins. For example, comparison of sub-networks of bacterial motility derived from *Treponema pallidum* and *Campylobacter jejuni* with estimated false-negative rates of 76 % and 77 %. By combining both datasets, 33 % of all known flagellar interactions could be recovered, decreasing the false negative rate to 67 % (Rajagopala et al., 2007). When protein domains and fragments are used, this number can be further reduced. This study was facing the challenging task to recover more than 50 % of all interactions using the Y2H method, which could be achieved for a reference set of human PPIs (Chen et al., 2010).

1.2.5.2 False-positives

Like other methods, the Yeast two-hybrid system has the potential to detect interactions which do not naturally occur in the investigated organism or between the investigated organisms, respectively. False-positive interactions can be divided into technical and biological artifacts.

Technical false-negatives are for example Yeast two-hybrid interactions that are not based on the assembly of two-hybrid proteins (that is, the reporter gene(s) are activated without a protein-protein interaction between bait and prey). These kinds of false-positives appear due to bait proteins that act as transcriptional transactivators like the AD fusion tag of the prey. They can be in a large part avoided by proper pretesting of bait strains for their autoactivation properties. Second, some bait or prey proteins may affect general colony viability and hence enhance the ability of a cell to grow under selective conditions or activate the reporter gene. When protein fragments are investigated, it is also possible that protein regions are exposed which do not occur naturally and may be of sticky nature. Those ‘sticky’ proteins that presumably bind non-specifically can be identified by incorporating a cut-off rate at the level of the promiscuity of PPIs (Albers et al., 2005). If only a small number of interacting partners is allowed and promiscuous proteins are filtered-out, the false-positive rate decreases, but at the cost of an increased number of false-negatives. However, so-called ‘hubs’, highly connected proteins that have many binding partners may be excluded. Mutations or other random events of unexpected nature must be accepted as well. A number of procedures have been developed to identify

or avoid false-positives, including the utilization of multiple reporters, independent methods of specificity testing, or simply repeating assays to make sure a result is reproducible (Koegl and Uetz, 2007; Serebriiskii et al., 2000; Serebriiskii and Golemis, 2001).

Biological false-positives are two-hybrid interactions without physiological relevance. This means that two proteins do physically interact but this interaction never occurs *in vivo* as they are never co-expressed in time or in space. Examples may include paralogs that are expressed in different tissues or at different developmental stages. The problem is that the “false-positive” nature can rarely be proven, as there might be conditions under which these proteins do interact with a biological purpose. Overall, hardly any false-positive can be explained mechanistically (although many may simply interact non-physiologically in a living cell as well). While it often remains difficult to prove the biological significance of an interaction, many studies have attempted to validate them by independent methods. Finding an interaction by several methods may certainly increase the probability that it is biologically significant. Uetz and coworkers evaluated the all the Y2H interactions found in a proteome-wide screen of Kaposi's Sarcoma Herpes virus (KSHV) by Co-Immunoprecipitation (Co-IP) and found that about 50 % of them can be confirmed (Uetz et al., 2006). Similarly, when subsets of the large-scale human Y2H interactome were evaluated, 78 and 65 % of them could be verified by independent methods (Rual et al., 2005; Stelzl et al., 2005). Finally, integrating external datasets such as literature-curated interactions or homologous interacting proteins can enhance the reliability of an interaction network.

1.2.6 Experimental strategy for matrix-based Y2H-screens

Yeast two-hybrid assays are carried out in living yeast cells although in theory any other cell could be used. Actually, mammalian two-hybrid variations have been developed in the last years, like Split β-galactosidase, β-lactamase protein-fragment complementation, Bimolecular fluorescence complementation (BiFC), Luciferase complementation, Split TEV assay and Resonance energy transfer systems (FRET and BRET) (Lievens et al., 2009). The advantage of yeast is still its convenient handling with no need for high-end screening platforms. Another advantage compared to *in vitro* assays is that the yeast cell provides an *in vivo* situation.

The proteins of interest are expressed as plasmid-encoded recombinant fusion proteins. The bait protein is fused to the DNA-binding domain (DBD) of the yeast GAL4 transcription factor, the prey protein to the transactivation domain (AD) of GAL4. A physical contact of the bait and prey protein simulates the native GAL4 transcription factor. Other fusion proteins can be used, too and have been established in other systems. For example, instead of the GAL4 domains, the bacterial transcription factor LexA has been used. In general, any protein that can be split and reconstituted to form an active protein can be used (Drees, 1999). For high-throughput screens our laboratory routinely uses the HIS3 auxotrophic marker. It encodes the essential enzyme imidazoleglycerole-phosphate dehydratase (IGPD) which catalyses the sixth step of Histidine biosynthesis. Hence, yeast growth on minimal medium that lacks Histidine can be used to indicate an interacting protein pair. Non-interacting pairs cannot support growth on selective medium. This reporter system is very simple and easy to use because the presence of yeast colonies indicates an interaction.

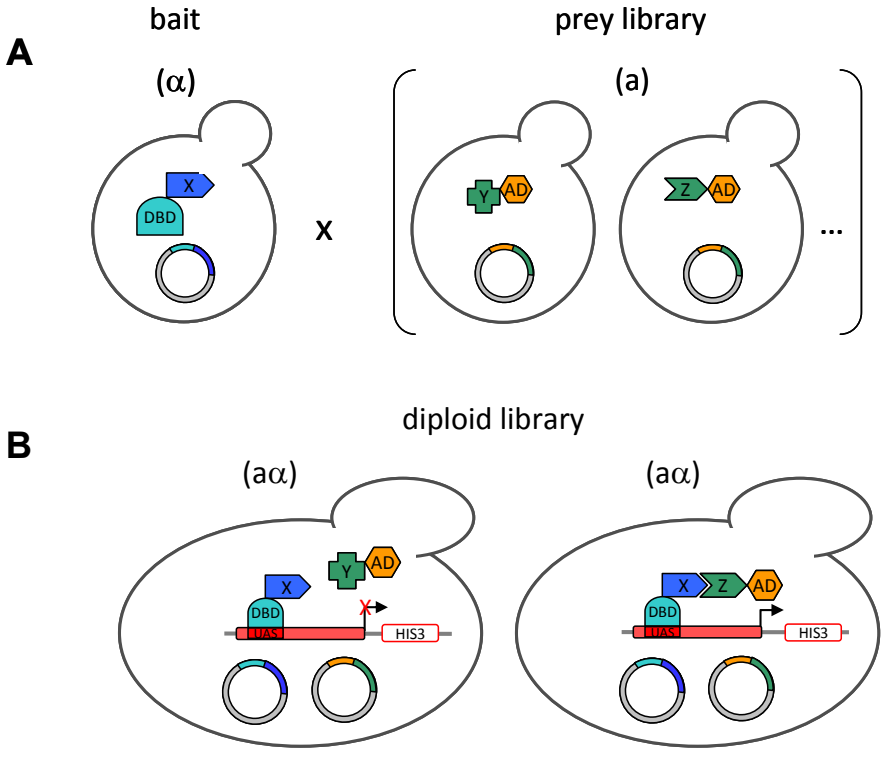


Figure 2: Mating strategy used for Y2H screens.
A) Haploid yeast cells of mating type a are transformed with a bait plasmid and those of mating type α with prey plasmids. A single bait strain is mated with a prey library. **B**) Resulting diploids (aα) carry the genetic material of mated haploids. Interacting fusion proteins activate expression of the HIS3 reporter gene which assures survival on minimal medium that lacks Histidine (diploid on the right); diploids with non-interacting fusions cannot grow (diploid on the left).

Before the binary tests are carried out, the bait and prey plasmids must be transfered into the same yeast cell. This is conveniently done by mating. The bait and prey plasmids are separately transformed into haploid yeast cells of the opposite mating types *a* and α (Figure 2A). Mating results in diploid yeast cells that carry the genetic material of both haploid test strains including the bait and prey plasmids (Figure 2B). HIS3 is often used as reporter gene. Alternatively, other reporters have been introduced. LEU2 and URA3 allow selection on readout medium that lacks Leucine or Uracil. Auxotrophic markers are not the only ones that can be used. The ADE2 reporter system changes colony color from red to white on adenine starvation medium when diploids express interacting proteins. Beta-galactosidase (*lacZ*) or GFP (green fluorescent protein) are used as colorimetric or fluorescence reporter as well. Finally, transcription-independent two-hybrid systems have been developed. The Split-Ubiquitin system for instance, is based on the cleavage of the interacting fusion proteins by the proteasome (Johnsson and Varshavsky, 1994) which takes place in the cytoplasm. In a matrix-based screen the preys are arrayed on defined positions of the test plate. For a high-throughput application the preys can be arranged in a 384 format on a single test plate (Uetz et al., 2000b).

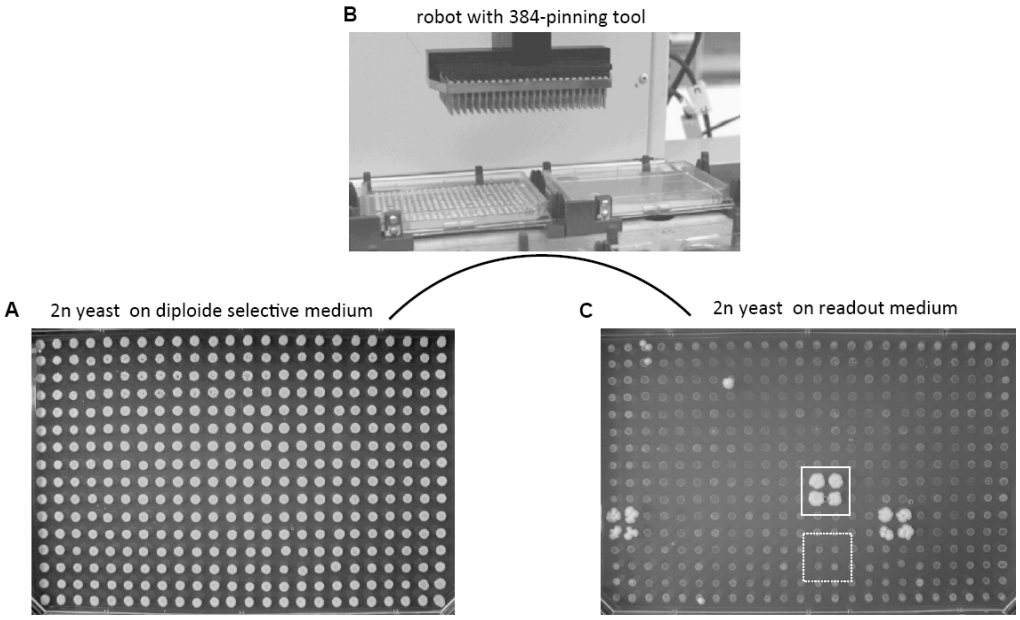


Figure 3: Scheme of a matrix-based Y2H-screen.
A) Prey array mated against single bait on diploid selective agar medium containing 96 individual preys. Single preys are replicated as quadruplicates to check interaction reproducibility. **B**) 384-pinning tool of replication robot during pinning step of diploids onto readout medium. **C**) Diploids (2n yeast) on readout medium that lacks Histidine. Diploids were grown on selective medium for one week at 30 °C. Activation of the HIS3 reporter leads to growth on minimal medium indicating a pairwise interaction (see quadruplicate position in white square). Non-interacting pairs do not support growth on minimal medium (for example the quadruplicate in dashed square). (Rajagopala and Uetz, 2009)

Preys may be organized as individual colonies, yet it is recommended to screen duplicate or quadruplicate copies to ensure reproducibility (Figure 3C). The whole array of haploid preys is usually mated against a single bait of the opposite mating type. Thus each potential interaction pair is tested one-on-one (see Figure 3A). For high-throughput analysis a replication robot is used with a 96- or 384-pin tool (Figure 3B) to replicate the test-arrays between the different selective steps.

1.3 Varicella zoster virus as model for combinatorial Y2H-screening

1.3.1 VZV - clinical aspects

Varicella zoster virus (VZV), also called HHV-3 (Human herpesvirus 3), is an alphaherpesvirus of the genus varicellovirus. It is one of nine human pathogenic herpesvirus species. Commonly, it is the causative agent of Varicella (also called chickenpox) in children and both herpes zoster (shingles) or postherpetic neuralgia in adults (Sampathkumar et al., 2009; Takahashi et al., 1974). Primary VZV infection results in chickenpox, a vesicular rash accompanied by fever, which may rarely result in complications like encephalitis or pneumonia (Kleinschmidt-DeMasters et al., 1996). After the primary infection, VZV remains dormant in the trigeminal and dorsal root ganglia of the infected person, so called virus latency (Steiner et al., 2007). In about ten to 20 percent of all cases, VZV is being released from the dormant state. The secondary infection causes herpes zoster. Serious complications of herpes zoster include VZV vasculopathy, VZV myelopathy, postherpetic neuralgia, herpes ophthalmicus, zoster multiplex, or zoster sine herpete (Gilden et al., 2009), see Figure 4 for an overview.

1.3.2 General introduction into the *Herpesviridae* family

VZV belongs to the *Herpesviridae*, which is a large family of enveloped double stranded DNA viruses with a broad host spectrum ranging from mammals to birds and reptiles. More than 100 different species of herpesviruses have been identified, including 9 human pathogenic viruses. Common for all herpesviruses are the ability to persist within a host in a latent state after primary infection. During latency only a few viral genes are expressed, limiting the hosts opportunity to establish an immune response directed against specific viral antigens. The latent virus can reactivate at

later time points and lead to secondary infections which sometimes are of different nature than the primary infection.

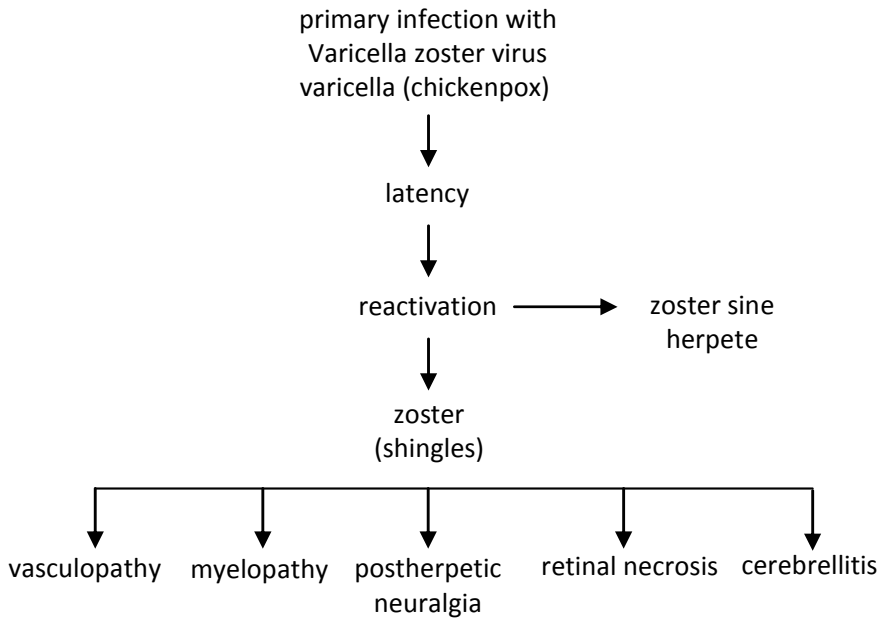


Figure 4: Diseases from secondary VZV infections.
Neurologic diseases produced by reactivation of Varicella zoster virus (Gilden et al., 2009).

1.3.3 *Herpesviridae* - subfamilies and phylogeny

The Herpesviridae family is divided into three subfamilies: *Alpha*-, *Beta*- and *Gammaherpesvirinae* (Fauquet, 2005).

The three subfamilies were initially separated based on biological differences such as cell-tropism and growth properties in cell-culture. *Alphaherpesvirinae* are neurotrophic and replicate efficiently in cell culture, whereas *Betaherpesvirinae* have a narrower cell tropism in culture and infection *in vivo* may result in enlargement of the infected cells (Cytomegaly). *Gammaherpesvirinae* replicate poorly in culture and are oncogenic lymphotropic viruses. With the advances in genetics, herpesviruses are now divided into subfamilies based on genomic differences. While most of the biological differences accurately predicted the subfamily association, some viruses have been moved into a different subfamily after their genome was fully sequenced. This was for example the case for Marek's disease virus (MDV), which was initially thought to be closely related to Epstein Barr virus (EBV), a gammaherpesvirus, due to its ability to infect lymphocytes in addition to its oncogenicity (Osterrieder et al., 2006). But genetic analysis revealed that the virus had more in common with alphaherpesviruses (Buckmaster et al., 1988; Cebrian et al., 1982). The evolutionary

divergence of the three subfamilies has been predicted to have taken place around 400 million years ago, which is about the same time as prehistoric animals first started to step onto land (Daeschler et al., 2006; McGeoch and Gatherer, 2005).

1.3.4 Virion structure and genetic organization

A major characteristic of herpesviruses is the architecture of the virion. The virion size varies between 150 and 200 nm and is composed of four distinct components: envelope, tegument, capsid and the core (Figure 6A). An electron micrograph of VZV is shown in Figure 5.

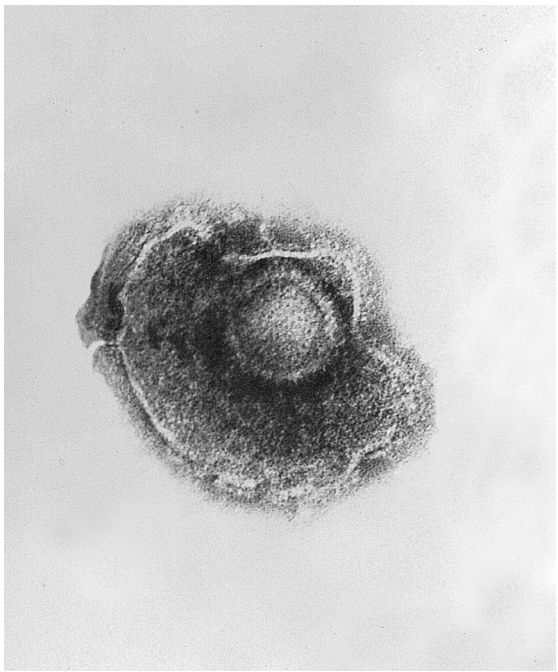


Figure 5: VZV virion.
Electron micrograph of a VZV virion. Source: Dr. Erskine Palmer; B.G. Partin, Centers for Disease Control and Prevention (CDC), Public Health Image Library (PHIL), #1878.

The core consists of the viral genome, which is believed to be packaged in a toroidal shape (Furlong et al., 1972; Perdue et al., 1976). While all herpesviral genomes are made up of double stranded DNA, the genome size of different species varies from 125 kbp (VZV) up to 230 kbp (HCMV) (Chee et al., 1990; Davison and Scott, 1986). The genome size also reflects the protein coding potential of different species with VZV encoding at least 74 ORFs, while HCMV encodes approximately 200 ORFs. 46 of the VZV ORFs encode proteins essential for virus replication (Figure 7). In addition, the genetic organization of the genomes differs between species. The VZV genome consists of two covalently linked segments called unique long (U_L) and unique short (U_s), each flanked by sequences of inverted repeats (IR_S and IR_L)

(Figure 6B). The U_L and U_S segments occur in four isoforms, depending on the orientation of the U_L and U_S region (Figure 6C). The KSHV genome in contrast is made up of only one segment flanked by multiple uniform repeats (Lagunoff and Ganem, 1997). The inner nucleoprotein core is surrounded by an icosahedral (T=16) capsid shell of 162 capsomeres (Major Capsid Protein), encoded by the HSV-1 UL19 orthologs. Therefrom, 12 are pentavalent capsomeres located at the vertices of the capsid, while the remaining 150 are hexavalent capsomeres (Mettenleiter, 2002).

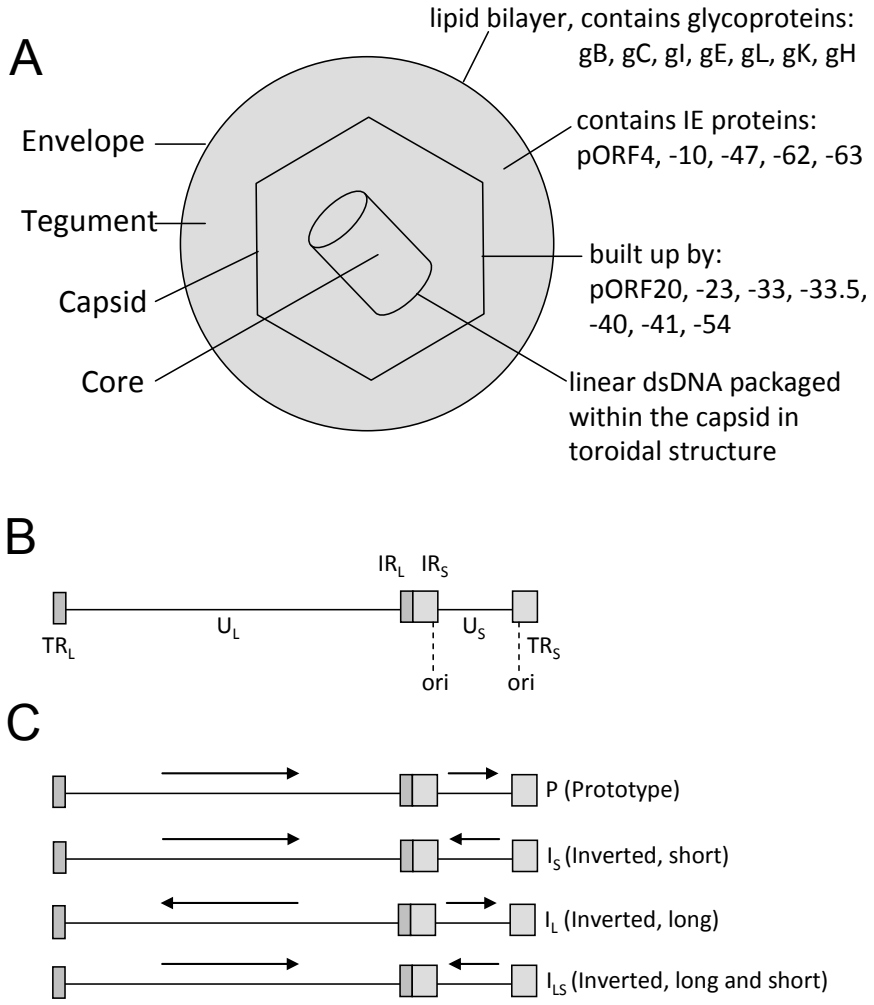


Figure 6: Schematic image of the VZV virion.
A) Major elements of VZV are indicated left. Important components of each element are shown at the right. **B**) General structure of the VZV genome. The genome consists of 124,884 bp (Davison and Scott, 1986) can be divided into the unique long (U_L) and unique short (U_S) region, which are flanked by terminal repeats long and short (TR_L, TR_S) and internal repeat long and short (IR_L, IR_S). The origins of DNA replication (ori) are located in the IR_S and TR_S. **C**) Isomeric forms of Herpesvirus genomes: four different isoforms with respect to the orientation of the U_L and U_S.sequences do occur (Morse et al., 1977). The P and I_S isomers make up more than 95 % of the packaged VZV DNA (Rahaus, 2006).

Surrounding the capsid is a proteinaceous, loosely structured layer called the tegument containing IE-proteins (immediate-early proteins) which are needed for the first steps of host cell infection. While the composition and structure of the capsid is

quite conserved throughout the *herpesviridae*, the composition of the tegument has a higher degree of variation between different species. There is, however, a set of at least five tegument proteins which are believed to be conserved between the three subfamilies. The tegument is enclosed within a lipid bilayer, the envelope, which contains at least seven conserved virus-encoded glycoproteins (Figure 6) (McGeoch et al., 2006).

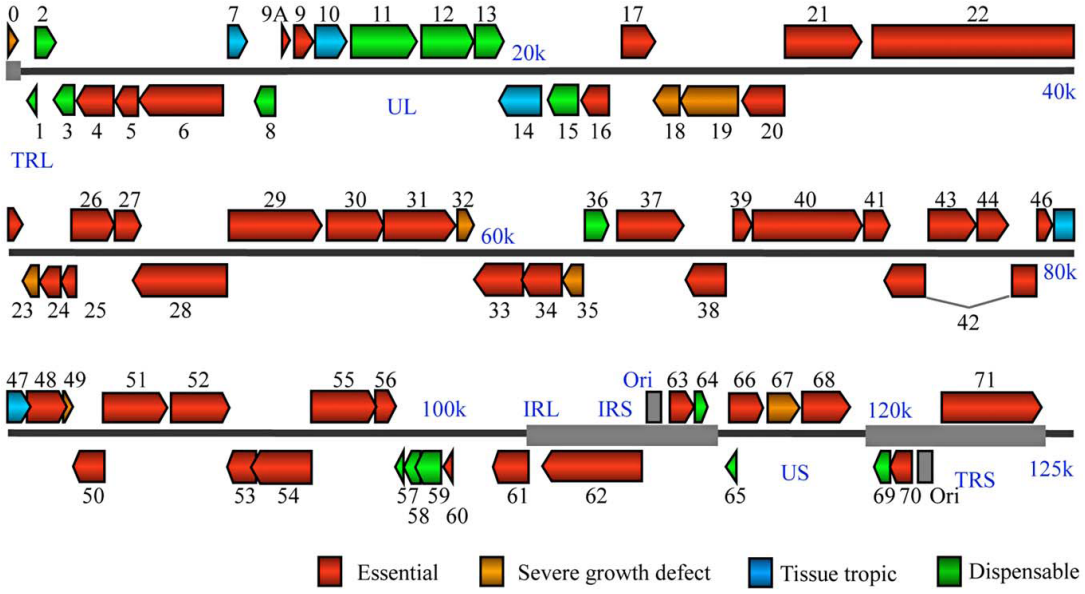


Figure 7: VZV ORFs and their orientation in the genome.
The 74 open reading frames of VZV have recently been investigated by a systematic deletion mutagenesis study by Zhang and coworkers (Zhang et al., 2010a). The colour scheme reflects the effect of the deletion of the respective ORF on virus replication. This Figure is taken from the original publication (doi:10.1371/journal.ppat. 1000971.g002.).

1.3.5 Herpesviral life-cycle

Herpesviruses share a common replication cycle. An infection is initiated when one or more of the glycoproteins protruding from the viral envelope attach to specific surface receptors on the host cell. The ubiquitously expressed glycosaminoglycan heparan sulfate is a common receptor for initial cell attachment of most herpesviruses, including VZV. Additionally, different herpesvirus species attach to different surface receptors, which is an important factor in determining the tropism of a virus.

VZV mediates the entry into the host cell via interaction of glycoprotein E (gE, Figure 6) and Insulin degrading receptor (IDE) (Li et al., 2006), while HSV-1 enters the host cell via interaction of gD and HVEM (Herpesvirus Entry Mediator) (Montgomery et al.,

1996). In addition to HVEM, HSV-1 can also attach to the cellular surface proteins nectin-1 and nectin-2 (Geraghty et al., 1998; Warner et al., 1998). Other cellular receptors for herpesviruses include intergrins in the case of HCMV (Feire et al., 2004), complement receptor 2 (CR2) for EBV (Yefenof et al., 1976), as well as DC-SIGN, the Cystine transporter xCT and Intergrin α3β1 for KSHV (Akula et al., 2002; Kaleeba and Berger, 2006; Rappocciolo et al., 2006).

Several theories on the entry mechanisms of the viral capsid into the host cell have been proposed. The most accepted theory suggests that the interaction between the viral glycoprotein and cellular receptor brings the viral envelope into close proximity of the cell membrane, which results in the fusion between the two membranes (Roizman et al., 2005). Viral tegument proteins and the capsid are subsequently released into the host-cell, thus initiating the infectious cycle. When released into the cytoplasm, the capsid is transported to the nucleus along microtubules, which is mediated via the motor proteins Dynein and Dynactin (Dohner et al., 2002; Naranatt et al., 2005). At the nucleus the viral genome is transported through a nuclear pore, leaving behind the empty capsid. Studies using temperature sensitive mutants of HSV-1 have indicated that the large tegument protein (ORF40 for VZV) is involved in this process (Batterson et al., 1983). After the entry of the viral genome transcription of the viral genes occurs in a cascade like manner where the immediate early (IE) genes are expressed first, followed by the early (E) genes and finally the late (L) genes (Figure 8b-d). While the IE genes mostly encode transcriptional activators necessary for proper expression of E and L genes, E genes encode genes involved in the replication of the viral genome. Late genes are expressed after the replication of the viral genome is initiated, encoding structural proteins required or building up new viral particles (Roizman, 1996). Production of new virus particles occurs within the nucleus of the infected cell, in specific replication compartments (Sourvinos and Everett, 2002; Taylor et al., 2003). These compartments are thought to contain the structural proteins which make up the capsid, the proteins necessary for replicating the viral DNA, in addition to other proteins necessary for proper production of new viral particles. One of these proteins is the viral scaffolding protein (ORF33.5 in VZV), which forms a scaffold for the capsid proteins to assemble around (Singer et al., 2005). Herpesviral DNA is replicated in a rolling circle mechanism, resulting in a long concatemeric DNA molecule where several viral genomes are organized in a head-to-tail fashion.

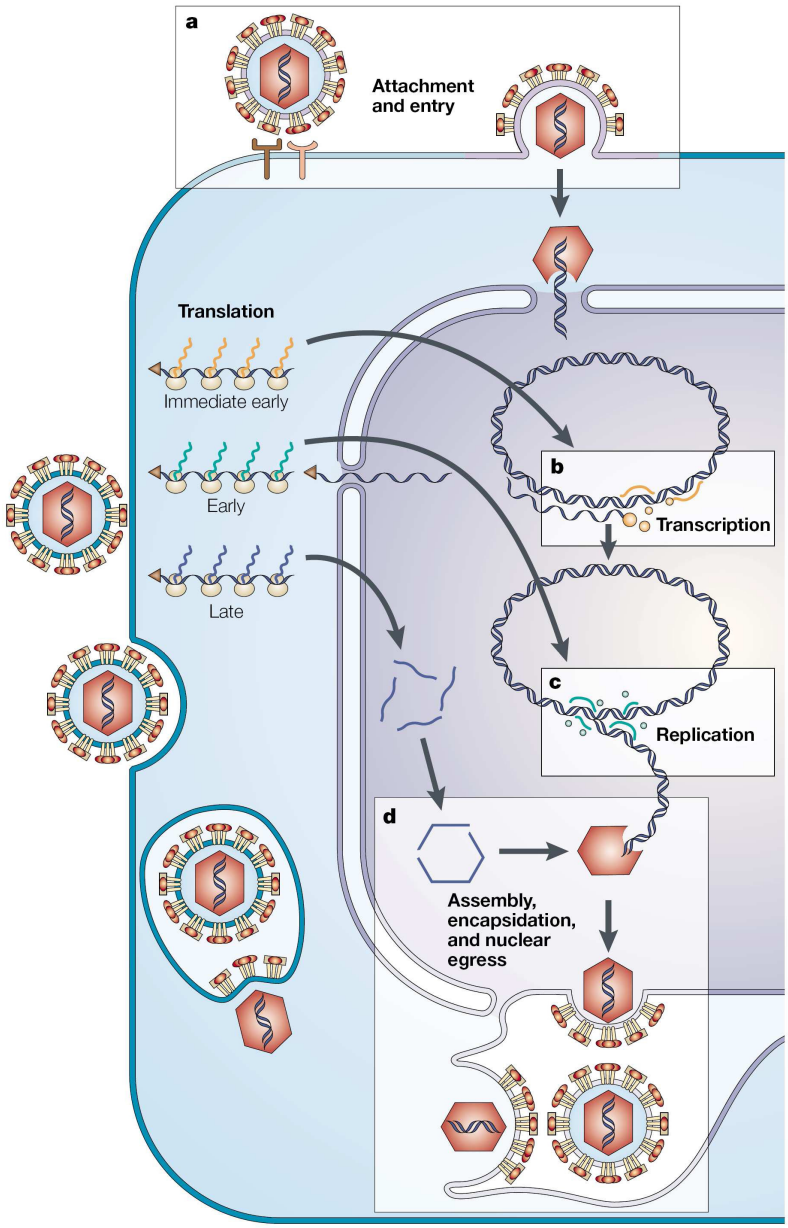


Figure 8: Herpesviral replication cycle.
a) Attachment and virus entry. Viral membrane proteins on virus particles bind to cellular receptors on the plasma membrane of the host cell, initiating fusion of the two membranes. Nucleocapsids containing the viral genome (red hexagons) are released into the cytoplasm and transported to nuclear pores. The linear viral genome is translocated into the nucleus and circularizes. **b**) Transcription. Three classes of viral genes are transcribed and translated into proteins. Immediate-early proteins (yellow) participate in further transcription. **c**) Replication. Early proteins (green) synthesize new viral DNA molecules using circularized viral DNA as a template. **d**) Assembly, encapsidation and nuclear egress. Late proteins (blue) assemble into capsids, which are filled with newly replicated viral DNA. Nucleocapsids leave the nucleus by budding through the inner nuclear membrane (a process termed 'envelopment') into the perinuclear space. Through a complex process of de- and re-envelopment, mature virus particles reach exocytic vesicles, which fuse with the plasma membrane and release new virus particles into the extracellular space. Figure from Coen and Schaffer, 2003.

The concatemeric DNA molecule is subsequently cleaved into single genomes during the packaging process. In VZV, there are seven virus-encoded proteins which are essential for the packaging of the viral genome. They are encoded by the ORFs number 25, 26, 30, 34, 43, 45/42 and 54 (Visalli et al., 2007). These proteins have been designated as DNA packaging proteins since deletion or mutation of any of these genes results in partial or no packaging of the viral genome. The two proteins pORF30 and pORF45/42, which is the splicing product of the ORFs 45 and 42, make up the Terminase complex and have been reported to be involved in cleavage of the concatemeric DNA as it is packaged into the capsid (Abbotts et al., 2000; Visalli et al., 2007). UL33, the homologue of ORF25 in VZV, was suggested to be a part of the Terminase complex, and that it interacts with UL28 (ORF30 in VZV) and stabilizes the UL15/UL28 (ORF45/42-ORF30) complex (Beard et al., 2002; Yang and Baines, 2006). After packaging of the viral DNA, the capsid is transported through the inner nuclear membrane into the perinuclear space, thereby obtaining an initial viral envelope. Two viral proteins, UL31 and UL34 in HSV-1 (ORF27 and ORF24 in VZV), called the nuclear egress complex (NEC), have been reported to play an important role in the nuclear egress of herpesviruses (Reynolds et al., 2001). Orthologs in Pseudorabies virus (PrV) (Fuchs et al., 2002), MCMV (Muranyi et al., 2002) and EBV (Lake and Hutt-Fletcher, 2004) have been reported to share the same function. The initial viral envelope is believed to get lost when the viral particle fuses with the outer nuclear membrane thus releasing the uncoated virus into the cytoplasm. In HSV-1, the viral kinase US3 (VZV ORF66) is reported to play an important role in this process, as US3 deletion mutants accumulate in the perinuclear space (Reynolds et al., 2002). In the cytosol the viral particles acquire the final tegument and a second envelope when being translocated into the trans-Golgi network (TGN) (Granzow et al., 2001; Skepper et al., 2001). The details how the viral particle is transported into the TGN is not fully understood, but for PrV and EHV-1 (Equine herpesvirus 1) the conserved proteins gM (VZV ORF50) and UL11 (VZV ORF49) have been reported to be involved (Kopp et al., 2004; Kopp et al., 2003; Seyboldt et al., 2000). Glycoprotein M may also mediate the final viral egress from the infected cell via secretorial vesicles budding off from the trans-Golgi network, leaving the cell through exocytosis.

1.3.6 Latency

All herpesviruses are able to enter a latent infection state where only a few genes are expressed. By limiting the number of proteins expressed, the virus can minimize the viral epitopes presented by class I MHC (major histocompatibility complex I), and thus prevent detection by cytotoxic T-lymphocytes (CTL). In HSV, the LAT (Latency associated transcript) has been reported to be important for heterochromatin formation on lytic genes during latency (Wang et al., 2005). For KSHV, demethylation of the ORF50 promoter, encoding the lytic transactivator RTA (replication and transcription activator) is enough to change the infection from a latent to lytic state (Chen et al., 2001). More recently it has been reported that some herpesvirus species also express microRNAs (miRNAs) during latency (Cai et al., 2005). MiRNAs are short RNA transcripts (about 22 nucleotides), involved in gene regulation through binding specific mRNAs, thereby inhibiting translation.

VZV remains latent in neurons in cranial nerve-, dorsal root-, and autonomic ganglia along the entire human neuraxis (Gilden et al., 1987; Hyman et al., 1983; LaGuardia et al., 1999). VZV DNA is present in a circular or concatemeric (end-to-end) state during the latent state (Clarke et al., 1995) and is present as two to nine copies in one to seven percent of individual neurons (Pevenstein et al., 1999). The wide range of VZV copy numbers during latent infection may reflect the strength of the primary infection. Transcripts corresponding to the VZV ORFs 21, 29, 62, 63, and 66 have been identified in latently infected human ganglia (Cohrs et al., 1995; Cohrs and Gilden, 2003; Cohrs et al., 1994), while the VZV ORF63 transcript is the most prevalent and abundant that can be detected (Cohrs and Gilden, 2007).

1.4 Drug discovery

The process of finding a new drug against a chosen target for a particular disease usually involves high-throughput screening (HTS), where large libraries of chemicals are tested for their ability to modify the target. For example, if the target is a protein kinase, the chemicals are tested for their ability to inhibit that kinase.

HTS is also used to determine how selective compounds are for the chosen target. A compound should interfere selectively with the target and not with other, related targets. This cross-screening is important, because the more unspecific a compound

hits related targets, the more likely is off-target toxicity of this compound once it reaches clinical trials.

Another important method for drug discovery is drug design, where a prediction is made which chemical compound might fit into an active site or within a PPI interface. One example is fragment-based lead discovery (FBLD). Novel pharmacophores can emerge very rapidly from these exercises.

Fragment-based lead discovery (FBLD) is a method used for finding lead compounds as part of the drug discovery process. It is based on identifying small chemical fragments, which may bind weakly to the biological target, and then modifying them or combining them to produce a lead with higher affinity (Congreve et al., 2008). When a lead compound series with sufficient target potency -selectivity and favorable drug-like properties has emerged, one or two compounds are proposed for drug development.

A third method is virtual high throughput screening, where screening is performed *in silico* with 3D molecule structures, attempting to dock those virtual libraries to a target, e.g. the herpesviral thymidine kinase (Seifert, 2005).

Prerequisites for virtual drug development are, apart from the identification of a promising drug target, which is commonly a protein, the knowledge of its structure and the active sites within the protein which can be targeted.

1.5 Mapping of interaction epitopes using peptide arrays

1.5.1 Spot synthesis

SPOT synthesis was originally introduced by Frank (Frank, 1992). It is a convenient and flexible technique for simultaneous and parallel chemical synthesis of peptides at distinct positions on a solid support like a cellulose membrane. SPOT method has opened up great opportunities to synthesize and subsequently screen large arrays of synthetic peptides (Landgraf et al., 2004; Otte et al., 2003).

Peptide arrays prepared by SPOT synthesis can be used to characterize molecular interactions, for example by epitope mapping, which is the analysis of protein-protein- and protein-nucleic acid interactions. This allows the identification of biologically active peptides. Moreover, peptide arrays can be used to describe molecular recognition events on the single amino acid level.

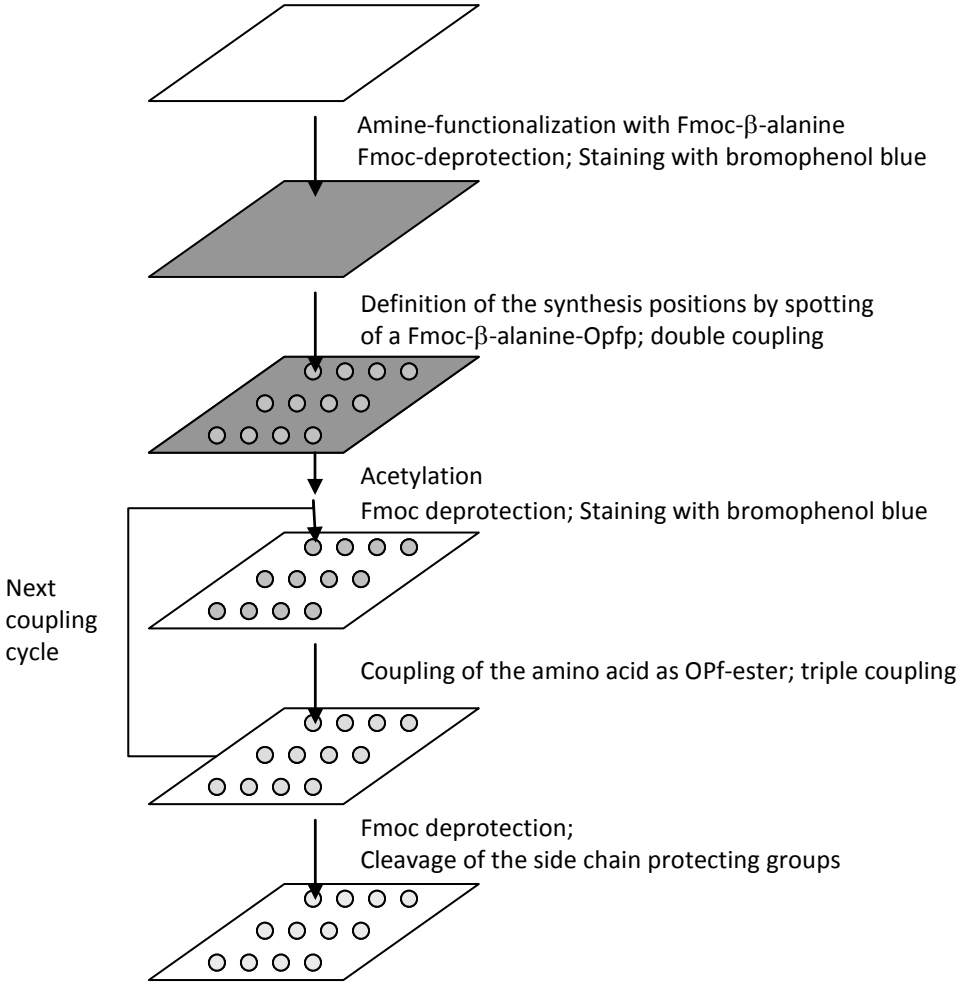


Figure 9: SPOT synthesis following solid phase peptide synthesis.
After membrane functionalization, the amino acids are coupled as an activated ester solution. To increase peptide density and to limit spot diameter, a triple coupling of each amino acid solution per cycle is performed. The coupling can be monitored by staining the free amino groups with bromophenol blue (optional). After the amino acid coupling, the remaining free amino groups are blocked by acetylation (capping). Subsequently, the Fmoc group is removed by treatment with 20 % piperidine/DMF to prepare the membrane for the next coupling step. After the last coupling step, the side-chain protection groups are removed by treatment with TFA. Figure adapted from Hilpert et al., 2007.

Figure 9 shows the scheme of the SPOT synthesis method. A commercially available filter paper (e.g. Whatman 50) with free functional hydroxyl groups can be modified to amino acid coupling by coupling with activated Fmoc-ß-Ala-OH and the subsequent removal of the Fmoc-group. The peptide chain elongation steps follow solid phase peptide synthesis (Merrifield, 1963).

1.5.1.1 Analysis of protein-peptide contact sites based on SPOT synthesis

Protein interactions can be studied at the amino acid level by peptide scans of overlapping peptides (Geysen et al., 1984). Proteins interact via surface accessible interaction sites which involve amino acid residues and backbone contacts either along a continuous segment of the protein chain (linear epitopes) or they involve amino acid residues from at least two segments close in space by the folded conformation (conformational or discontinuous epitopes) but separated in the primary sequence (Figure 10). An epitope is defined as the contact site of a protein that interacts with a binding partner. In linear epitopes, the key amino acids mediating the contacts with the binding partner are located within one part of the primary structure and comprise no more than 15 amino acids in length, usually 8-12 amino acids. Peptides covering such epitopes are usually able to compete with the protein-protein interaction and have similar affinities as the intact protein from which they are derived (Reineke et al., 1999). The mapping of linear epitopes by standard SPOT technology using overlapping peptides derived from the primary AA sequence of a protein is an easy and efficient approach. The whole protein sequence is fragmented and synthesized on cellulose with short overlapping peptides, in my standard conditions 15 amino acids in length, shifted by 3 amino acids to at the next spot (Figure 10A). The resulting peptide scan is subsequently probed for binding to the respective interacting protein, which was identified e.g. by Y2H analysis. Binding assays can be performed directly on the peptide array by immunodetection of bound protein or by autoradiography of a radiolabeled probe. Transfer to a nitrocellulose membrane is also possible, but leading to a loss of protein.

Since protein-protein interactions are often mediated by large contact patches (Jones and Thornton, 1996), most proteins have more than one binding site which interacts with a binding partner.

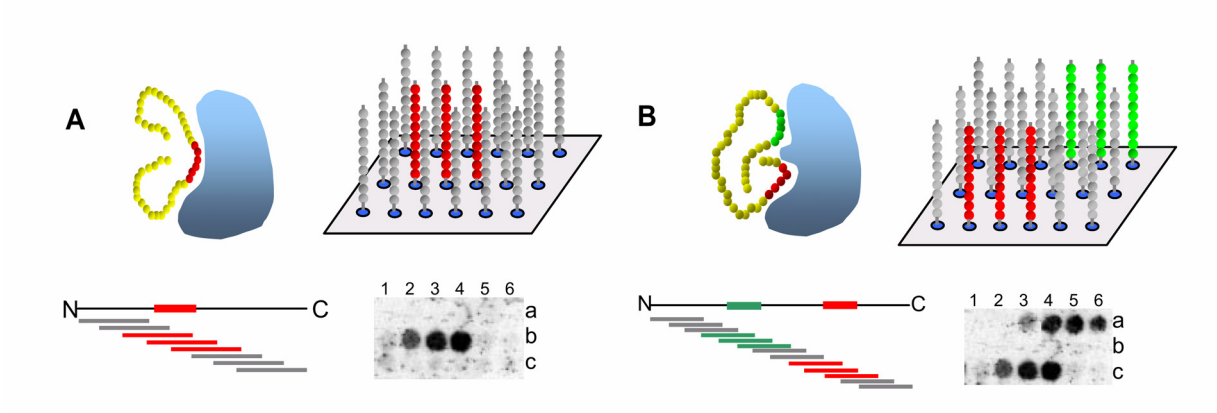


Figure 10: Scheme of mapping PPI epitopes.
A) Linear and **B**) discontinuous protein-protein interaction epitopes can be detected in binding assays using peptide arrays. The sequences colored red in **A**) are the linear contact site with detectable signals in binding assay. The sequences colored red and green in **B**) are separated in the protein sequence and are detected individually in the binding assay.

In discontinuous binding sites the key residues contributing to the binding affinity are separated in the protein sequence, but they are close to each other in the folded protein, forming a combined binding epitope (Figure 10B). The mapping of discontinuous epitopes using fragments of protein either generated chemically or biologically suffers from the drawback that peptides comprising only single binding region of a discontinuous binding site generally have very low affinities for the binding partner in solution, even if the complete binding site mediates a high affinity interaction. Thus, the mapping of discontinuous epitopes with overlapping peptide scan on cellulose might be accordingly very difficult. However, peptide arrays have the advantage of a high density of immobilized peptides on membranes, which can be estimated between 0.2 and 0.4 µmol/cm² using β-Alanine coated Whatman 50 cellulose membranes (Frank and Overwin, 1996), which achieves a high local peptide concentration (mM level or higher, estimated) and thus increases the screening sensitivity. In this way, the inherent defect that peptides comprising only the individual binding regions normally having low affinities for the binding partner can be overcome to a great extent by SPOT method.

1.6 Aims

The present study is composed of two major parts. First, an improved strategy for Yeast two-hybrid screening was developed, applied and validated. The system was applied to a VZV ORFeome, and validated using intrinsic quality features as well as external available interaction data. A second screen was performed with the novel system, mapping intraviral protein-protein interactions of Hepatitis E virus, which is going to be published together with an intracellular localization study performed at the Max-von-Pettenkofer Institut (Ludwig-Maximilians-Universität, München) (Osterman et al., 2010).
In the second part, a promising drug-target for development of anti-VZV compounds, pORF25, was further characterized by mapping its self-interaction contact sites. This was part of a collaborate study which investigated this highly conserved protein among the herpesvirus-family, proposing it to be a chaperone involved also in Terminase complex formation (Vizoso Pinto et al., 2010).

1.6.1 Combinatorial Y2H-screening with permutated fusion tags

Irrespective of sterical conditions, small variations of the Y2H system were shown to result in differences in the resulting interactions (Rajagopala et al., 2009). As no other large-scale screen has used multiple Y2H systems within a single study, not much attention has been paid to this phenomenon so far. For example, two systematic screens of protein-protein interactions among proteins of Kaposi's sarcoma associated herpesvirus (KSHV) yielded very little overlap (Rozen et al., 2008; Uetz et al., 2006). Clearly, slight variations of vectors, strains, or assay conditions can strongly affect the resulting interactions, even when identical proteins are used. Together with the knowledge about high false-negative rates, my aim was to create a Y2H system, made up of two compatible bait- and prey vectors to amplify the interactome of Varicella zoster virus. Traditional Y2H vector systems carry the cloning sites downstream of the DBD and AD, so that the bait and prey proteins carry N-terminal fusion domains. This assay layout allows screening of eight possible steric combinations of every possible protein pair, increasing the chance to overcome possible steric hindrance effects (see Figure 11 for a schematic overview).

Several vectors have been described that use C-terminal fusions of AD and DBDs (Beranger et al., 1997; Brown and MacGillivray, 1997; James, 2001; Millson et al., 2003). These studies have shown that the topology of the fusion site is critical when selected protein pairs were tested (Brown and MacGillivray, 1997). However, those previously studied vectors are not applicable for high-throughput screening because they require conventional cloning by restriction digest and ligation. As a consequence, it remained unclear how N- or C-terminal Yeast two-hybrid fusions behave when larger sets of proteins or whole genomes are analyzed in standardized Y2H assays.

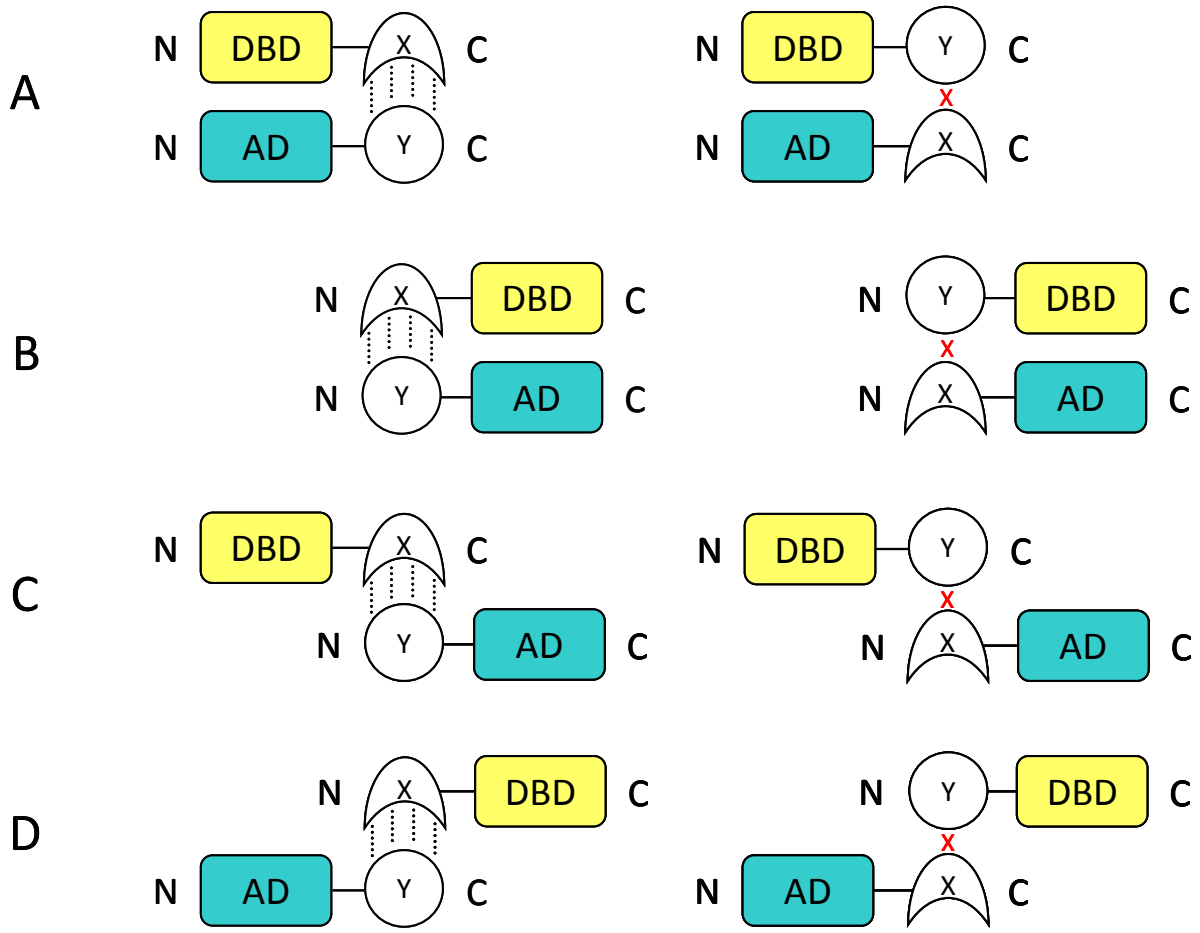


Figure 11: Principle of combinatorial Y2H screening.
Each protein pair X-Y is tested eight times (**A**-**D**) instead of two times in a classical screen (**A**). **A**) Classical NN-topology, both fusion tags are fused to N-termini of the hybrid proteins X and Y. In a pairwise whole proteome screen each protein is tested once as bait and every time as prey, so every protein pair X-Y is tested twice, as bait-X/prey-Y and bait-Y/prey-X. **B**-**D**) CC, NC and CN-topologies are screened in the same way as the classical system, thereby quadrupling the screening combinations. Different scenarios are supposable allowing-, or preventing detectability of a PPI in Y2H assays. **A**)-**D**), left side. It is assumed that an interacting protein pair X-Y does interact and the fusion tags allow physical interaction (dotted lines). One or more topological DBD-AD tag combinations subsequently may allow driving reporter gene expression. It seems plausible that a certain distance range between the fusion tags promotes the reconstitution of the native-like transcription factor. **A**)-**D**) Right side. It is postulated that the fusion tags sterically hinder the interaction, so that no TF-reconstitution occurs (indicated by red crosses).

1.6.2 Mapping of the homomerization domain of VZV ORF25

The VZV DNA-encapsidation proteins are involved into the packaging of the linear virus genomes into the capsid (Visalli et al., 2007). The concatemeric dsDNA is packaged into the procapsid through a channel formed by a portal protein complex embedded in a unique vertex of the procapsid and cut into single genomes. This is a mechanism that is unique to herpesviruses and many tailed dsDNA bacteriophages (Rao and Feiss, 2008). The protein encoded by VZV ORF25 comprises 156 amino acids and a molecular mass of 17,461 kDa. It is conserved throughout the family of the *Herpesviridae* and plays a central and essential role in the DNA-packaging process (Beard et al., 2002; Visalli et al., 2007; Vizoso Pinto et al., 2010). It may be a promising new drug-target, as we are facing first resistances against established nucleoside analogs which are used for antiviral therapy to date. A prerequisite for virtual drug design was achieved in this study by revealing the interaction sites and associated essential residues.

2 Materials and Methods

2.1 Materials

2.1.1 Instruments

12-channel pipette	Biohit, Frankfurt
8-channel pipette	Eppendorf, Hamburg
Agarose gel electrophoresis chambers	Peqlab, Erlangen
Benchtop centrifuge 5810R	Eppendorf, Hamburg
Benchtop centrifuge Labofuge 400R	Heraeus, Stuttgart
Biomek 2000 automated workstation	Beckman Coulter, Krefeld
Bioruptor	Diagenode, Liège
Developer	Kodak, New Haven, USA
Glassware	Schott, Mainz
High Density Replication Tool	Beckman Coulter, Krefeld
Incubator	Heraeus, Stuttgart
Incubator, shaker	Infors, Bottmingen, Switzerland
Microcentrifuge, Biofuge pico	Heraeus, Stuttgart
MultiPep peptide spotter	INTAVIS AG, Köln
NanoDrop ND-1000, spectrophotometer	Peqlab, Erlangen
PAGE-chambers	GE Healthcare, München
PH-meter	Eppendorf, Hamburg
Photometer	BioRad, München
Piston driven pipettes	Gilson, Middleton, USA
Semidry-Blotter	Amersham, Freiburg
COREX 8445 centrifugation tubes	Corning Glass Works, New York
Stand alone centrifuge J2-HS	Beckman Coulter, Krefeld
Stand alone centrifuge Avanti-J20	Beckman Coulter, Krefeld
Thermocycler	BioRad, München
UV-lightbox	Peqlab, Erlangen

2.1.2 Consumable Materials

96-deep-well plates	VWR International, Darmstadt
96-well plates, flat bottom	Sarstedt, Nümbrecht
96-well plates, round bottom	Sarstedt, Nümbrecht
Bottle top sterile filters, Ø 0.22 µm	Nalge Nunc, Rochester, USA
Conical tubes, 15 ml & 50 ml	Greiner Bio-One, Frickenhausen
Dialysis membrane, Type 8/32	Roth, Karlsruhe
Microcentrifuge tubes, 1.5 ml & 2 ml	Eppendorf, Hamburg
Omnitray plates	Nalge Nunc, Wiesbaden
PCR tubes	Corning, Amsterdam, The Netherlands
Petri dishes	Greiner, Nürtingen
PVDF membrane	Millipore, Schwalbach
Sterile syringe filters, Ø 0.22 µm	Schleicher & Schuell, Dassel
Syringes	Braun, Melsung
Whatman paper	Bender und Hobein, Karlsruhe
Whatman 50 paper, (hardened low ash filter paper, grade 50)	Bender und Hobein, Karlsruhe

2.1.3 General Chemicals

1,4-dithiothreitol	Sigma-Aldrich, Taufkirchen
3-amino-1,2,4-triazole	Sigma-Aldrich, Taufkirchen
5-fluoroorotic acid	Fermentas, St Leon-Rot
Acetic acid, glacial	Merck, Darmstadt
Acrylamide/N,N'-methylenbisacrylamide	Roth, Karlsruhe
Activated charcoal	Roth, Karlsruhe
Agarose	Peqlab, Erlangen
Anhydrotetracycline	Sigma-Aldrich, Taufkirchen
Ammonium persulfate	Sigma-Aldrich, Taufkirchen
Amylose Resin High Flow	New England Biolabs, Frankfurt
β-mercaptoethanol	Roth, Karlsruhe
Bovine serum albumin	PAA Laboratories, Marburg
Bradford reagent	Roth, Karlsruhe
Bromophenol blue	Sigma-Aldrich, Taufkirchen

$CaCl_2$	Roth, Karlsruhe
Complete protease inhibitor mix tablets	Roche, Mannheim
Coomassie brilliant blue G-250	Roth, Karlsruhe
DMSO	Roth, Karlsruhe
dNTPs	Metabion, Martinsried
EDTA	Roth, Karlsruhe
Ethanol	Roth, Karlsruhe
Ethidium bromide	Roth, Karlsruhe
Glass beads	Sigma-Aldrich, Taufkirchen
Glycerol	Roth, Karlsruhe
HCl	Roth, Karlsruhe
Imidazole	Sigma-Aldrich, Taufkirchen
IPTG	Roth, Karlsruhe
Isopropanol	Merck, Darmstadt
Igepal CA-630	Sigma-Aldrich, Taufkirchen
Lithium acetate	Sigma-Aldrich, Taufkirchen
Lysozyme, from chicken egg white	Sigma-Aldrich, Taufkirchen
Maltose	Roth, Karlsruhe
Skimmed milk powder	Saliter, Obergünzburg
NaOH	Roth, Karlsruhe
Ni-NTA-Agarose	Qiagen, Hilden
PEG P3640	Peqlab, Erlangen
PMSF	Serva, Heidelberg
Ponceau-S	Roth, Karlsruhe
Potassium acetate	Roth, Karlsruhe
Salmon sperm DNA	Sigma-Aldrich, Taufkirchen
SDS	Roth, Karlsruhe
Sodium hypochlorite, 12 %	Sigma-Aldrich, Taufkirchen
TEMED	Roth, Karlsruhe
Tris-base / -HCl	Roth, Karlsruhe
Triton-X-100	Sigma-Aldrich, Taufkirchen
Tween20	Roth, Karlsruhe

2.1.4 Kits

ECL western blotting analysis system	Amersham, Freiburg
PCR-product clean-up kit, “SV Wizard”	Promega, Mannheim
Plasmid maxiprep kit	Qiagen, Hilden
Plasmid miniprep 96 kit, “Montage”	Millipore, Schwalbach
Plasmid miniprep kit, “Quick Lyse”	Qiagen, Hilden

2.1.5 Compounds of Bacteria- and Yeast Media

Adenine	Sigma-Aldrich, Taufkirchen
Agar-agar	Otto-Nordwald KG, Hamburg
Ammonium sulfate	Roth, Karlsruhe
Ampicillin	Roth, Karlsruhe
Arginine	Merck, Darmstadt
Asparagine	Merck, Darmstadt
Aspartic acid	Merck, Darmstadt
Bacto™ peptone	Roth, Karlsruhe
Bacto™ Yeast extract	Roth, Karlsruhe
Chloramphenicol	Sigma-Aldrich, Taufkirchen
D(+)-glucose	Roth, Karlsruhe
Gentamicin	Roth, Karlsruhe
Glutamine	Merck, Darmstadt
Histidine	Merck, Darmstadt
Isoleucine	Merck, Darmstadt
Kanamycin	Roth, Karlsruhe
Leucine	Merck, Darmstadt
Lysine	Merck, Darmstadt
NaCl	Roth, Karlsruhe
Phenylalanine	Merck, Darmstadt
Serine	Merck, Darmstadt
Spectinomycin	Duchefa, Haarlem, The Netherlands
Threonine	Merck, Darmstadt
Tryptophan	Merck, Darmstadt
Tyrosine	Merck, Darmstadt

Uracil	Sigma-Aldrich, Taufkirchen
Valine	Merck, Darmstadt
Yeast nitrogen base	Gibco, Karlsruhe

2.1.6 Chemicals for Peptide Synthesis

DMF (N,N- Dimethylformamide)	Roth, Karlsruhe
DIC (N,N'-Diisopropylcarbodiimide)	Merck, Darmstadt
NMI (1-Methylimidazole)	Merck, Darmstadt
Methanol	Roth, Karlsruhe
HOBt (1-Hydroxybenzotriazole)	Sigma-Aldrich, Taufkirchen
NMP (1-Methyl-2-pyrrolidone)	Fisher Scientific, Schwerte
Piperidine	Roth, Karlsruhe
Acetic anhydride	Sigma-Aldrich, Taufkirchen
TFA (2,2,2-Trifluoroacetic acid)	Sigma-Aldrich, Taufkirchen
TIPS (Triisopropylsilane)	Merck, Darmstadt
Phenol (Hydroxybenzene)	Merck, Darmstadt

AA-Derivative	Novabiochem No.	M [g/mol]	m [0,5 mmol]
Fmoc-Ala-OH	04-12-1006	311,3	155,7 mg
Fmoc-Arg(Pbf)-OH	04-12-1145	648,8	324,4 mg
Fmoc-Asn(Trt)-OH	04-12-1089	596,7	298,4 mg
Fmoc-Asp(OtBu)-OH	04-12-1013	411,5	205,8 mg
Fmoc-Cys(Trt)-OH	04-12-1018	585,7	292,9 mg
Fmoc-Gln(Trt)-OH	04-12-1090	610,7	305,4 mg
Fmoc-Glu(OtBu)-OH	04-12-1020	425,5	212,8 mg
Fmoc-Gly-OH	04-12-1001	297,3	148,7 mg
Fmoc-His(Trt)-OH	04-12-1065	619,7	309,9 mg
Fmoc-Ile-OH	04-12-1024	353,4	176,7 mg
Fmoc-Leu-OH	04-12-1025	353,4	176,7 mg
Fmoc-Lys(boc)-OH	04-12-1026	468,5	234,3 mg
Fmoc-Met-OH	04-12-1003	371,5	185,8 mg
Fmoc-Phe-OH	04-12-1030	387,4	193,7 mg
Fmoc-Pro-OH	04-12-1031	337,4	168,7 mg
Fmoc-Ser(tBu)-OH	04-12-1033	383,4	191,7 mg
Fmoc-Thr(tBu)-OH	04-12-1000	397,5	198,8 mg
Fmoc-Trp(Boc)-OH	04-12-1103	526,6	263,3 mg
Fmoc-Tyr(tBu)-OH	04-12-1037	459,6	229,8 mg

AA-Derivative	Novabiochem No.	M [g/mol]	m [0,5 mmol]
Fmoc-Val-OH	04-12-1039	339,4	169,7 mg
Fmoc-ß-Ala-OH	04-12-1044	319,3	128,0 mg

Table 1 Fmoc protected amino acids for SPPS

2.1.7 DNA and Protein Ladders

1 kb DNA-ladder	Invitrogen, Karlsruhe
PeqGOLD prestained protein ladder	Peqlab, Erlangen

2.1.8 Enzymes

BP Clonase Mix II	Invitrogen, Karlsruhe
LR Clonase Mix II	Invitrogen, Karlsruhe
PrimeStar HS DNA Polymerase	TaKaRa Bio Europe, Potsdam
Proteinase K	Invitrogen, Karlsruhe
Restriction endonucleases	Promega, Mannheim and New England Biolabs, Frankfurt
RNAse A	Qiagen, Hilden
Shrimp Alkaline Phosphatase	Promega, Mannheim
T4 DNA Ligase	Promega, Mannheim
Taq DNA Polymerase	Promega, Mannheim

2.1.9 Media for Bacterial Culture

2.1.9.1 LB liquid medium

0.5 % (w/v) Bacto™ Yeast extract

1 % (w/v) Bacto™ Peptone

1 % (w/v) NaCl

Autoclave 20 min at 121 °C. If required, add antibiotics for selection after the medium has cooled down to approximately 50 °C.

2.1.9.2 LB agar plates

Additionally to the liquid medium add 1.6 % Agar-Agar prior to autoclaving. Pour the medium 5 cm or 10 cm Petri dishes after cooling down to approximately 50 °C.

2.1.9.3 Antibiotics

Name	Working concentration
Ampicillin	100 µg/ml
Chloramphenicol	34 µg/ml
Gentamicin	50 µg/ml
Kanamycin	50 µg/ml
Spectinomycin	100 µg/ml

Antibiotic stocks were prepared in 1000 x concentration and stored at -20 °C.

2.1.10 Media for Yeast Culture

2.1.10.1 YPD liquid medium

1 % (w/v) Bacto™ Yeast extract

2 % (w/v) Bacto™ Peptone

2 % (w/v) Glucose

Autoclave 20 min at 121 °C.

2.1.10.2 YPD agar medium

Additionally to the liquid YPD medium add 1.6 % Agar-Agar prior to autoclaving. After autoclaving add 4 ml of a 1 % adenine solution (1 % adenine in 0.1 M NaOH). Pour the medium in OmniTrays after cooling down to approximately 50 °C.

2.1.10.3 SD media for yeast selective plates

- **5 x Medium concentrate, 1 l**

8.5 g Yeast nitrogen base

25 g Ammonium sulfate

100 g Glucose

7 g Dropout mix

Add deionized water up to 1 l. Sterile filter through a bottle top filter.

- **Dropout mix (-Leu,-Trp,-His)**

Amino acid / Nucleobase	m [g]
Methionine	1
Arginine	1
Phenylalanine	2.5
Lysine	3
Tyrosine	3
Isoleucine	4
Glutamic acid	5
Aspartic acid	5
Valine	7.5
Threonine	10
Serine	20
Adenine	1
Uracil	1

Mix all components and store under dry conditions

- **Amino acid stock solutions**

Leucine (Leu, L): dissolve 7.2 g/l in DIW, sterile filter.
Tryptophan (Trp, (W), T): dissolve 4.8 g/l in DIW, sterile filter.
Histidine (His, H): dissolve 4 g/l in DIW, sterile filter.
Note: Even though the official one-letter code for Tryptophan is W in this study it is abbreviated T, which is commonly used for Threonine.

- **3-AT Stock Solution**

Dissolve 1 mol/l 3-AT in deionized water.
Sterile filter through a bottletop filter.

- **Preparation of selective plates**

Per liter medium autoclave 16 g Agar in 800 ml DIW, then let it cool down to 60 °C to 70 °C. Add 200 ml of 5 x medium concentrate. Depending on the required selective plates, add the missing amino acids or the required amount of 3-AT from the stock solution as follows.

- -Trp (-T): 8.3 ml Leucine and 8.3 ml Histidine stock solution
- -Leu (-L): 8.3 ml Tryptophan and 8.3 ml Histidine stock solution
- -Leu-Trp (-LT): 8.3 ml Histidine stock solution

- -Leu-Trp-His (-LTH): Nothing to be added
- -Leu-Trp-His, x mM 3-AT: x ml of 3-AT stock solution

2.1.10.4 Solutions for Yeast transformation

- Salmon Sperm DNA: Dissolve 7.75 mg/ml salmon sperm DNA in water, autoclave 15 min at 121 °C and store at -20 °C
- 96 PEG solution (100 ml):
  - 45.6 g PEG
  - 6.1 ml of 2 M LiOAc (Lithium acetate)
  - 1.14 ml of 1 M Tris pH 7.5
  - 232 µl 0.5 M EDTA
  - Fill up to 100 ml with sterile water and autoclave

2.1.11 General Buffers and Solutions

- **6 x DNA loading buffer**

30 % (v/v) Glycerol
0.01 % (w/v) Bromophenol blue
0.001 % (w/v) RNAse A

- **50 x TAE buffer**

242 g Tris base
100 ml 0.5 M EDTA
57.1 ml Glacial acetic acid
Add enough DIW to dissolve solids, adjust pH 7.6-7.8 with HCl then fill up to 1000 ml.

- **Plasmid miniprep solution 1**

50 mM glucose
25 mM Tris-HCl (pH 8.0)
10 mM EDTA (pH 8.0)
Autoclave, add 50 µg/ml RNAse A and store at 4 °C.

- **Plasmid miniprep solution 2**

200 mM NaOH
1 % (w/v) SDS

- **Plasmid miniprep solution 3**

3 M Potassium acetate

Adjust pH 5.5 with glacial acetic acid.

2.1.12 Plasmids

Name	Promoter	Marker	Tag	Host	Origin	Source	Description
pDONR 201	none	kan^r	none	TOP10	pUC	Invitrogen	Gateway® donor vector
pDONR 207	none	gen^r	none	TOP10	pUC	Invitrogen	Gateway® donor vector
VZV ORFeome	none	gen^r	none	TOP10	pUC	(Uetz et al., 2006)	All VZV ORFs and ORF-fragments cloned into pDONR207
pDONR 223	none	$spec^r$	none	TOP10	pUC	(Rual et al., 2004)	Gateway® donor vector
pETG-10A	T7/lac	amp^r	N-His	BL21 (DE3)	pBR322	A. Geerlof, EMBL	expression vector based on pET-22b
pETG-30A	T7/lac	amp^r	N-His N-GST	BL21 (DE3)	pBR322	A. Geerlof, EMBL	expression vector based on pET-22b
pETG-40K	T7/lac	kan^r	N-MBP	BL21 (DE3)	pBR322	A. Geerlof, EMBL	expression vector based on pET-24d
pGADT7g	pADH1	amp^r	N-Gal4-AD	TOP10 AH109	pUC, 2µ	(Uetz et al., 2006)	Gateway® Y2H prey expression vector
pGBKT7g	pADH1	kan^r	N-Gal4-DBD	TOP10 Y187	pUC, 2µ	(Uetz et al., 2006)	Gateway® Y2H bait expression vector
pGBGT7g	pADH1	gen^r	N-Gal4-DBD	TOP10 Y187	pUC, 2µ	(Uetz et al., 2006)	Gateway® Y2H bait expression vector
pGADCg	pADH1	amp^r	C-Gal4-AD	TOP10 AH109	pUC, 2µ	(Stellberger et al., 2010)	Gateway® Y2H prey expression vector
pGBKCg	pADH1	kan^r	C-Gal4-DBD	TOP10 Y187	pUC, 2µ	(Stellberger et al., 2010)	Gateway® Y2H bait expression vector
pNusA	tet	amp^r	N-NusA N-His	BL21 (DE3)	unknown	Santhera Pharmaceuticals	expression vector based on pASK75

2.1.13 Bacterial Strains

Name	Description	Source
E. coli TOP10	Chemically competent bacterial cells for amplification of plasmids	Invitrogen
E. coli DB3.1	Electrocompetent and chemically competent cells suitable for propagation of plasmids containing the ccdB gene	Invitrogen
E. coli BL21-CodonPlus® (DE3)-RIL	Encodes T7 RNA polymerase under control of the lacUV5 promoter for easy protein expression	Stratagene

2.1.14 Yeast Strains

Name	Genotype	Source
AH109	MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::$GAL1_{UAS}$-$GAL1_{TATA}$-HIS3, $GAL2_{UAS}$-$GAL2_{TATA}$ADE2, URA3::$MEL1_{UAS}$-$MEL1_{TATA}$-lacZ	BD Biosciences Clontech, Palo Alto, USA
Y187	MATα, ura3- 52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met-, gal80Δ, URA3::$GAL1_{UAS}$-$GAL1_{TATA}$-lacZ	BD Biosciences Clontech, Palo Alto, USA
MaV103	MATa, leu2-3,112, trp1-901, his3Δ200, ura3-52, ade2-101, gal4Δ, gal80Δ, $cyh2^R$, $can1^R$, GAL1::HIS3@LYS2, GAL1::lacZ, SPAL10::URA3@ura3	(Vidal et al., 1996)
MaV203	MATα, leu2-3,112, trp1-901, his3Δ200, ura3-52, ade2-101, gal4Δ, gal80Δ, $cyh2^R$, $can1^R$, GAL1::HIS3@LYS2, GAL1::lacZ, SPAL10::URA3@ura3	(Vidal et al., 1996)

2.1.15 PCR-Primers

Name	Sequence [5´→3´]
VZV_ORF21_FW	AAAAAGCAGGCTCCGCCATGGAAGAACCAATTTGTTATGAT
VZV_ORF21_REV	AGAAAGCTGGGTCAGGGTCACTCCCACTTGTAT
VZV_ORF22_FW	AAAAAGCAGGCTCCGCCATGGATATAATTCCGCCTATAG
VZV_ORF22N_REV	AGAAAGCTGGGTATCTCGGTAGTTAGGTATTCCATTAATAG
attB1_FW	GGGGACAAGTTTGTACAAAAAAGCAGGCT
attB2_REV	GGGGACAAGTTTGTACAAGAAAGCTGGGT
GAL4_AD_REV	GGTTCGGACCGTTGCTACT
GAL4_DBD_REV	CGGCAAATATCGCATGCTTG
HEV_ORF2_P1_FW	AAAAAGCAGGCTCCGCCATGGGTAATACCAACACGCGGGTC
HEV_ORF2_P1_REV	AGAAAGCTGGGTCTGTCGGTCGGTCCTGCTCATG
HEV_ORF2_P2_FW	AAAAAGCAGGCTCCGCCATGCCTTCCCCAGCCCCATCG
HEV_ORF2_P2_REV	AGAAAGCTGGGTCTGCTAGCGCAGAGTGGGGG
HEV_ORF2_S_FW	AAAAAGCAGGCTCCGCCATGGCGGTCGCTCCGGCCCAT
HEV_ORF2_S_REV	AGAAAGCTGGGTCGGGGGTGAGGTTGCGGAAC
KpnI-GWY RFB-KpnI_FW	AAAAGGTACCGCATCAACAAGTTTGTACAAAAAAGCTGAACGA

KpnI-GWY RFB-KpnI_REV	AAAAGGTACCATCAACCACTTTGTACAAGAAAGCTGAAC
M13_FW	CGTTGTAAAACGACGGCCAG
M13_REV	GCCAGGAAACAGCTATGACC
pDONR207_FW	TTCCCGAGGTAATCGGAGTCCGGCT
pDONR207_REV	TGTGACAAAATAAAAACATCTACC
pETG-40K_FW	CAGCGGTGGCAGCAGCC
pETG-40K_REV	GCCGCCAGCGGTCGTA
pGADC_XbaI/BclI_FW	GATCCTCGAGTACGACGTACCAGATTACGCTTAGTGATCA
pGADC_XbaI/BclI_REV	CGTGATCACTAAGCGTAATCTGGTACGTCGTACTCGAG
pGADT7g_FW	TTTAATACGACTCACTATAGGGCG
pGADT7g_REV	CTGTGCATCGTGCACCATCT
pGBKC_8mer	AATTTGCA
pGBKC_A	CGGCATTGATATCGTCCAACTG
pGBKC_B	TAGCTTCATCTTGGTACCTCAGGAGGCTTGCTTCAAGCTT
pGBKC_C	AAGCTTGAAGCAAGCCTCCTGAGGTACCAAGATGAAGCTA
pGBKC_D	GGTCTTCTCGAGGAAAAATCAGTAG
pGBKT7g_FW	GTAATACGACTCACTATAGGGCG
pGBKT7g_REV	GACTCTTAGGTTTTAAAACGAAAA
pNusA_FW	AGGGCATTGATGATCTGGCT
pNusA_REV	AAATGTCGCACAATGTGCGC
T7_FW	TAATACGACTCACTATAGGG
T7_REV	TAGTTATTGCTCAGCGGTG
tADH1_REV	CGACCTCATGCTATACCTGAG

2.1.16 Antibodies

Name	Source
Anti-His-Tag, mouse monoclonal antibody	New England Biolabs, Frankfurt
Anti-HA, mouse monoclonal antibody	Covance, Freiburg
Anti-MBP, mouse monoclonal antibody	New England Biolabs, Frankfurt
Anti mouse IgG HRP conjugate, Goat polyclonal antibody	Sigma-Aldrich, Taufkirchen

2.2 Methods

2.2.1 General DNA-related Methods

2.2.1.1 Competent bacteria for DNA transformation ($CaCl_2$ method)

- Inoculate 150 ml LB medium with 1 ml bacterial pre-culture.
- Incubate at 37 °C to OD_{600} = 0.35 and cool the bacterial culture on ice for 10 min.
- Centrifuge for 10 min at 6000 rpm (Beckmann J2-HS) and 4 °C; discard the supernatant.
- Wash cells with 30 ml 100 mM $CaCl_2$, centrifuge for 10 min at 4000 rpm (Eppendorf 5810R) and 4 °C.
- Resuspend pellet in 30 ml 100 mM $CaCl_2$, and incubate for 1 h on ice.
- Centrifuge again and resuspend in 3 ml 100 mM $CaCl_2$ with 10 % Glycerol.
- Freeze 100 µl aliquots in liquid nitrogen and store at -80 °C.

2.2.1.2 Transformation of chemical competent bacteria

- Take the required number of tubes or 96-deep well plates containing competent *E.coli* from the -80 °C freezer.
- Thaw the cells on ice.
- Add the transforming DNA, vortex briefly to mix and incubate on ice for 20-30 minutes.
- Transfer the cells into a 42 °C water bath or a heat block for 90 s.
- Add 600 µl of RT LB medium (without antibiotic) and shake the cultures at 37 °C for 1 h.
- Centrifuge the cells at 13.000 rpm for 1 min and resuspend in 200 µl sterile Water.
- Transfer each suspension onto a LB agar plate with the appropriate antibiotic.
- Plate out each suspension by the glass bead shaking method (see next paragraph).
- Remove the glass beads
- Incubate at 37 °C

2.2.1.3 Glass bead streaking method for high-throughput plating

Add 3 to 4 glass beads to each plate. Swirl single or a whole stack of plates to spread bacteria. Recycle beads by washing in EtOH, water and autoclaving. Glass beads can be conveniently distributed from a 50 ml conical tube with a small hole in the lid.

2.2.1.4 Plasmid preparation (small scale, individual tubes)

- Pick an individual bacterial colony from culture plate and inoculate in 2-5 ml LB liquid medium (with antibiotic for plasmid selection).
- Incubate o/n with shaking at 37 °C.
- Pellet 1.5 ml of the culture in a Microcentrifuge (1 min at top speed).
- Resuspend pellet in 150 µl buffer P1.
- Add 150 µl buffer P2, mix gently by inverting the Microcentrifuge tube 4-5 times and incubate on ice for 5 min.
- Add 150 µl buffer P3, vigorously mix by shaking and centrifuge at top speed for 6 min. Transfer the supernatant to a fresh Microcentrifuge tube.
- Add 1/10 Vol. of 3 M NaOAc solution (pH 5.5).
- Add 0.7 Vol. isopropanol, mix thoroughly.
- Put the tube into a -20 °C freezer for 10 min.
- Centrifuge at top speed for 5 min.
- Wash the pellet with 70 % ethanol, air dry the pellet and dissolve it in 100 µl sterile H_2O.
- Alternatively commercial available Kits with silica columns for DNA purification were used (e.g. QuickLyse from Qiagen, Hilden).

2.2.1.5 Plasmid preparation (96-well plasmid preparation kit)

Plasmid DNA isolation from 96-well cultures was performed by the “Montage Plasmid Miniprep 96-well kit” (Millipore, Schwalbach).

2.2.1.6 Determination of nucleic acid concentration

The concentration of a DNA-solution was automatically determined with a NanoDrop UV spectrophotometer. The measurement is based on determination of the optical density at 260 nm and 280 nm. The OD_{260} value of one is equivalent to 50 µg/ml of

double stranded DNA. Pure DNA in aqueous solution should have an OD_{260}/OD_{280} ratio of 1.6-1.8.

2.2.1.7 Restriction digestion of DNA

Commonly 2 to 3 units of a restriction enzyme per µg of DNA were used. 1 to 10 µg of DNA was digested in the buffer recommended by the supplier. The reaction was carried out between 2 to 4 hours at enzyme supplier recommended temperature. The quality of the digestion was controlled by DNA agarose gel electrophoresis.

2.2.1.8 Nucleic acid analysis by agarose gel electrophoresis

Agarose was dissolved in 1x TAE-buffer by boiling in a microwave oven (final concentration between 0.8 % and 1.5 %). Ethidium bromide was added to a final concentration of 0.3 µg/ml. The molten gel was poured into a horizontal agarose gel electrophoresis chamber. Immediately insert the combs and let solidify. The DNA sample was mixed with the required amount of 6 x DNA loading buffer and loaded on the gel. For determination of the size an additional sample of DNA-ladder was loaded in parallel. The separation in the gel was commonly performed by electrophoresis for 1 h at 100 V. After separation the DNA was visualized by transillumination with 302 nm ultraviolet radiation.

2.2.1.9 Gel purification of DNA

The DNA band of interest was isolated electrophoretical by running the gel until the DNA band of interest was separated from adjacent contaminating fragments. The DNA was cut out of the gel with a scalpel and recovered into aqueous solution with the PCR-product clean-up kit, “SV Wizard”. The purification was carried out according to the protocol of the supplier.

2.2.1.10 Ligation of DNA

In all cases the restriction digested insert and vector were loaded on an agarose gel to check the DNA content before ligation. Ligation was performed in a total volume of 20 µl with a molar insert-vector ratio of approximately 4:1 in 1 x ligation buffer and one unit of T4 DNA ligase and incubated for 2 h at RT or at 16 °C overnight (blunt end ligation).

2.2.1.11 Polymerase chain reaction

- **PCR reaction setup**
  - 1 µl template DNA (20-200 ng)
  - 1 µl dNTP mix (dATP, dTTP, dCTP, dGTP. c = 10 mM each)
  - 5 x Polymerase buffer
  - 2 µl forward primer (c = 10 pmol/µl)
  - 2 µl reverse primer (c = 10 pmol/µl)
  - 0.5-1.0 µl polymerase (Taq polymerase for analytical, PrimeStar polymerase for preparative PCR)
  - In a total volume of 50 µl
- **PCR program**

A commonly used PCR program is given below. This program was modified according to specific PCR requirements.

- Step 1: 1 min at 95 °C
- Step 2: (30 s at 95 °C, 30 s at annealing temperature, 1 min per 1 kb DNA to be amplified at 72 °C) Repetition of Step 2 for 30 times
- Step 3: Store at 8 °C

The annealing temperature was commonly set to Tm - 5 °C.

2.2.1.12 Colony PCR (from yeast or bacterial cells)

Resuspend a single colony in 10 µl H_2O. Boil sample for 10 min at 95 °C. Spin down the debris and use 2 µl of the supernatant as template-DNA for PCR.

2.2.2 Yeast two-hybrid screening

2.2.2.1 Gateway®-cloning

The Gateway® system (Invitrogen) provides a universal technology to clone DNA sequences for Y2H assays, functional analysis and protein expression in multiple systems. It is based on site-specific recombination of bacteriophage lambda (Landy, 1989). The lambda recombinase is employed to recombine attachment sites (att) from different vectors or from a linear DNA strand to a vector. This allows cloning any

desired DNA sequence which can be PCR-amplified or synthetically derived into a donor vector from which it can be recombined in any destination vector.

2.2.2.2 Cloning of ORFs with the Gateway® system - the BP reaction

The first step of ORF cloning is the primer design. The PCR-amplification of the ORFs will be done in two steps: in the 1st PCR the ORF is amplified with ORF-specific primers, in the 2nd PCR attB-sites for cloning are completed to the resulting PCR product. For primer design, the "Oligo Calc: Oligonucleotide Properties Calculator" was used (Kibbe, 2007). The following parameters can serve as guidelines: optimal Tm is 55 °C, the Tm range is 10 °C above and below the optimal Tm, and the maximum Tm difference between forward and reverse primers is 5 °C.

- **The primers for the 1st PCR are designed as follows**

  - Forward Primer:
    5'-AA AAA GCA GGC TCC GCC ***ATG*** -18-20 nt (ORF specific sequence)-3'
  - Reverse Primer:
    5'-A GAA AGC TGG GTA ***CTA*** - 18-20 nt (specific ORF sequence)-3'

    ATG: start codon

    CTA: stop codon (reverse-complement of TAG); must be omitted for C-terminal fusions.

The 1st PCR is done with a standard PCR protocol employing a proof-reading polymerase (e.g., PrimeStar polymerase). However, the number of cycles in the 1st and 2nd PCR can be reduced if errors are expected.

The PCR-products are checked by agarose gel electrophoresis and 1-5 µl of the products is used for the 2nd PCR reaction with the primer pair:

- Forward Primer (attB1):
  5'-G GGG ACA AGT TTG TAC A**AA AAA GCA GGC T**-3'
- Reverse Primer (attB2):
  5'-GGG GAC CAC TTT GTA CA**A GAA AGC TGG GT**-3'

If unspecific bands are detected, the PCR product is gel purified. Otherwise, a column purification of the product is sufficient. The purified PCR product is used in a BP recombination to create the entry clone (Figure 12).

Figure 12: Gateway® BP reaction.
For generation of entry clones an attB substrate (PCR product) is recombined with an attP substrate (donor vector) mediated by BP Clonase® Mix II. Source: Gateway® manual (Invitrogen).

- **BP-reaction setup**
  - 0.5 µl BP Clonase Mix II
  - 1 µl destination vector (pDONR207 / 150 ng/µl)
  - 1 µl purified PCR fragment
  - Cover with 10 µl mineral oil
- **Performing BP-reaction**
  - Incubate 2 h at 25 °C
  - Add 1 µl Proteinase K
  - Incubate for 10 min at 37 °C to stop the enzymatic reaction
  - Add 10 ml sterile water
  - Transform into TOP10 cells
  - Select on LB-agar plate containing gentamicin

Picking of a single colony is usually sufficient. Constructs were verified by PCR with the primer pair pDONR207_FW and pDONR207_REV.

2.2.2.3 Destination vector creation with the Gateway® system - the LR reaction

The entry clones in pENTR207 can be transferred into different destination vectors with an LR reaction.

Figure 13: Gateway® LR reaction.
For generation of destination clones an attL substrate (entry clone) is recombined with an attP substrate (destination vector) mediated by LR Clonase® Mix II. Source: Gateway® manual (Invitrogen).

For Y2H assays, the pGBKT7g (bait) and pGADT7g (prey) destination-vectors were used (Uetz et al., 2006) plus novel Y2H-vectors with C-terminal DBD or AD fusion-tag (pGBKCg and pGADCg) designed in this study (Stellberger et al., 2010).

- **LR-reaction setup**
  - 0.5 µl LR Clonase Mix II
  - 1 µl Entry vector (150 ng/µl)
  - 1 µl Destination vector (150 ng/µl)
- **Performing LR-reaction**
  - Incubate 2 h at 25 °C
  - Add 1 µl Proteinase K
  - Incubate for 10 min at 37 °C to stop the enzymatic reaction
  - Transform into TOP10 cells and select on appropriate antibiotic

2.2.2.4 Yeast transformation

This protocol is suitable for 100 yeast transformations and was scaled up or down to experimental requirements. Bait constructs were transformed into Y187 and prey constructs into AH109 laboratory strains, each harboring the HIS3-Marker gene used for detection of PPIs in the Y2H-assay.

- **Required material**
  - Salmon sperm DNA
  - DMSO
  - Competent host yeast strains, in this study Y187 for baits and AH109 for preys
  - Lithium Acetate (0.1 M)
  - Selective SD agar plates
  - 96PEG solution
  - Carrier DNA (salmon sperm DNA)
- **Preparation of competent yeast cells**
  - Inoculate 50 ml YPD liquid medium with yeast in a 250 ml flask and shake o/n at 30 ºC (minimum 15 h, max. 24 h)
  - Spin down cells in 50 ml conical tube (3500 rpm, 5 min at RT)
  - Decant supernatant and dissolve pellet in 2 ml 0.1 M LiOAc

- Transfer yeast into 15 ml conical tube. Spin down yeast and resuspend in a total volume of 1.8 ml 0.1 M LiOAc

- **Preparation of "CT110" for yeast transformation**
  - Mix the following solutions in a 50 ml conical tube:
    - 20.73 ml 96PEG
    - 0.58 ml boiled salmon sperm DNA (boil frozen salmon sperm DNA at 95 ºC for 5 min and cool on ice before use)
    - 2.62 ml DMSO, shake immediately
    - 20 ng prey or bait plasmid construct
    - Vortex 1 min
  - Add the competent yeast cells and mix vigorously by hand or by vortexing for 1 minute. Immediately pipette 245 µl into 96-well plate or individual microcentrifuge tubes
  - Seal and vortex for 4 minutes
  - Incubate at 42 °C for 30 minutes
  - Centrifuge the 96 well plate or tubes; discard the supernatant and tap on cotton napkin for a couple of times
  - Add 150 µl of sterile water to all 96 wells, resuspend and plate them on SD plates lacking Leucine (prey) or Tryptophan (bait)
  - Incubate the plates at 30 ºC for 2-3 days. After one day the colonies start to appear; pick colonies after 2-3 days and make glycerol stocks (20 % (v/v), store at -80 °C)

2.2.2.5 Setup of the prey array

Preys rather than baits are arranged on an array, because the former do not generally show self-activation of transcription. For creation of the prey array, the layout of the array is defined first. Each prey construct of the proteome is given a specific position of a particular 96-well-plate, e.g. position A06 of prey plate #1. The wells of these 96-well-plates are filled with 100 µl YPD medium. Several colonies from a specific prey transformation are combined and manually transferred into the well at the previously defined position. These 96-well plates carrying the prey strains are incubated o/n at 30 °C and replicated to Omnitray plates with solid media - at least one selective Omnitray plate (-L) and one YPD plate. 50 µl of 50 % (v/v)

glycerol is added to the liquid culture plate. The plate is sealed and transferred to -80 °C for long term storage. The solid prey plates with prey strains in the 96-format were quadruplicated to the 384-format using an automatic workstation (Biomek 2000 laboratory robot). For increased throughput, duplicates rather than quadruplicates from two 96-formatted plates can be combined on one 384-formatted plate. This prey array was generally stored on selective plates (-L). Copies for Y2H screens were pinned on solid YPD medium because of faster growth and higher yield of cell mass.

2.2.2.6 Bait construction

If not provided by collaborators, baits were constructed by Gateway®-cloning procedure as described. Baits were transformed into the Y187 yeast strain and stored on -Trp plates and as glycerol stocks.

2.2.2.7 Autoactivation test

Prior to the two-hybrid analysis, the bait yeast strains (DB-X) must be examined for potential autoactivation properties. Autoactivation is defined as a detectable DB-X-dependent two-hybrid reporter gene activation in the absence of a prey construct (AD-Y). Weak to intermediate-strength self-activating baits can be used in two-hybrid array screens because the corresponding DB-X/AD-Y interactions produce significantly stronger signals that can be detected despite of some autoactivation background. In case of the HIS3 reporter gene, the unspecific background can be titrated by different concentrations of 3-AT, a competitive inhibitor of His3p (Hilton et al., 1965) (Figure 14). The lowest concentration of 3-AT that suppresses growth in this test is later on used for the Y2H interaction screen.

- **Required material**

  - Full medium and selective medium agar in single-well microtiter plates (OmniTray plates):
  - YPD plates
  - -LT plates
  - Selective plates without Leu, Trp, His and different concentrations of 3-AT, e.g., 0 mM, 1 mM, 3 mM, 10 mM, 25 mM and 50 mM (-LTH/3-AT plates)
  - Prey strain carrying the empty prey plasmid (AH109 with pGADT7g)

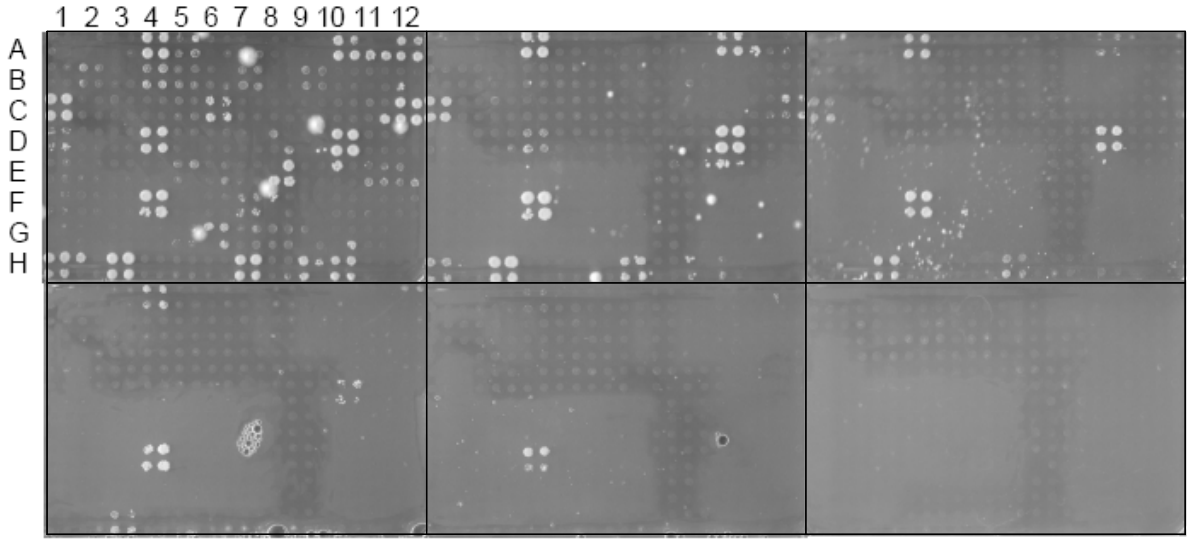


Figure 14: Self activation test of VZV-ORFs in pGBKCg.
VZV bait array in pGBKCg (Table 2) transformed into Y187 and mated against pGADT7g. After mating the diploid cells were selected on -LT agar and replicated onto -LTH agar with rising concentration of 3-AT. 3-AT concentrations (from upper left to lower right): 0, 1, 3, 10, 25, 50 mM. The lowest concentration is listed and used for the following Y2H-screen against the VZV prey-array.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	1	1 N	2	3	4	5	5 F	6	7	8	9	9a
B	9aN	10	11	12	12N	12C	13	14	14N	15	15N	15F
C	16	17	18	18N	18C	19	20	21	22N	23	24	24N
D	25	26	27	28	29	30	31	31N	31C	32	33	33.5
E	34	35	36	37	37N	38	39	39N	40	41	42	43
F	43C	44	45	46	47	48	49	50	50C	51	52	53
G	54	55	56	56C	57	58	59	60	60 C	61	62	63
H	64	65	65N	66	67	67N	67C	68	68F	68C	0	0 C

Table 2: VZV array-layout.
96 position array layout of the VZV library. Positions are defined by the column number in the first row (1 to 12) and the row number (A to H).

- **Testing procedure**

  - Bait strains are arrayed onto a single-well Omnitray agar plate; either the standard 96-well format or the 384-well format is used. Baits are first inoculated at the different positions of a 96-well plate as liquid culture, and then cells are transferred (manually or with robot) to Omnitray plates. In this step the 96-well

format can also be converted into the 384-well format, this will position all baits in quadruplicates on the 384-well formatted plate. Full medium agar (YPD agar) can be used, however, for long term storage of the array selective agar (-T) is suggested to prevent loss of plasmids.

- The arrayed bait strains are mated with the prey strain carrying the empty prey plasmid, AH109 with pGADT7g. Mating is conducted according to the standard screening protocol. Note: Compared to the actual screening protocol bait and prey strains are exchanged.
- After selecting for diploid yeast cells (on -LT agar) the cells are transferred to –LTH-medium selecting for the HIS3 reporter gene activity (His3p-expression). The transfer is repeated on several selective plates with increasing concentrations of the competitive inhibitor of His3p, 3-Aminotriazole (3-AT). Suggested are 3-AT concentrations of 0 mM, 1 mM, 3 mM, 10 mM, 25 mM, and 50 mM.
- These -LTH/3-AT plates are incubated for 1 week at 30 °C. The self-activation level of individual baits is assessed: the lowest 3-AT concentration that completely prevents colony growth is noted. As this concentration of 3-AT suppresses reporter activation in absence of an interacting prey this 3-AT concentration is added to -LTH plates in the subsequent interaction screening.

2.2.2.8 Screening for protein interactions using a Y2H protein array

The Y2H prey array is screened for protein interactions by a mating procedure that is carried out using robotic support. A strain expressing a single candidate protein as a DBD fusion (bait strain) is mated to all the colonies in the prey array. After mating, the colonies are transferred to double-selective medium (-LT to select for diploid yeast cells) and then to triple-selective medium (-LTH). A robotic workstation (Biomek 2000, Beckman Coulter) was used for the screening procedure. A 384 pin replicating tool (High Density Replication Tool, Beckman Coulter) was used to transfer the colonies from one plate to another. Between the transfer steps, the pinning tool must be sterilized by sequential immersion into a 20 % (v/v) bleach solution (20-40 s), sterile water (1 s), 95 % (v/v) ethanol (20-40 s), and sterile water (1 s). The level of these liquids should be 2 to 4 mm from the base of the pin. Attention: Take care that the ethanol does not evaporate.

It is important to ensure that plasticware is compatible with the movement of the robot. Here, the prey array was dissected on 86 x 128-mm OmniTrays in 384-colony format.

- **Required material**
  - 384 pin replicating tool
  - 20 % (v/v) bleach (1 % sodium hypochlorite)
  - 95 % (v/v) ethanol
  - Single-well microtiter plates (OmniTray; Nalge Nunc international) containing solid YPD medium, -LT, -LTH w/o and with different concentrations of 3-AT
  - Bait liquid culture (DBD-fusion expressing yeast strain)
  - Yeast protein array on solid YPD plates (prey array)

- **Sterilization**

Sterilize the 384-pin replication-tool by dipping the pins into 20 % bleach for 20 s, sterile water for 1 s, 95 % ethanol for 20 s, and sterile water again for 1 s. Repeat this sterilization prior to each transfer.

- **Preparing bait liquid culture**

Inoculate 20-30 ml of liquid YPD medium in a 250 ml conical flask with a bait strain (DBD fusion-expressing yeast strain) and grow overnight in 30 ºC shaker. If the Bait strains are frozen, they are streaked out or pinned on selective solid medium (-T) and grown 1-2 days at 30 ºC. Bait colonies from this plate are then used to inoculate the liquid YPD medium.

- **Mating procedure**
  - Pour the dense bait strain culture into a sterile Omnitray. Dip the sterilized pins of the replication-tool (thick pins should be used to pin baits) into the bait liquid culture and place directly onto YPD agar and allow the yeast to dry on the plates for 10-20 minutes.
  - Pin the prey array with sterilized pins (use thin pins for the preys) and transfer them directly onto the baits pinned onto the YPD plate, so that each of the 384 bait spots per plate receives a quadruplicate of different prey yeast cells. Incubate 1-2 days at 30 ºC to allow mating. Mating will take place within 15 h

but longer period is recommended, as some baits strains show poor mating efficiency.

- Selection of diploids: For the selection of diploids, transfer the colonies from YPD mating plates to single-well microtiter plates containing -LT medium using the sterilized pinning tool (thin pins should be used in this step). Grow for 2-3 days at 30 ºC until the colonies are > 1 mm in diameter. This step is essential because only diploid cell containing Leucine and Tryptophan markers on prey and bait vector, respectively, will grow on this medium. This step also leads to an enrichment of diploid cells which increases the efficiency of the next selection step.

- **Interaction selection**

  - Transfer the colonies from -LT plates to a single-well microtiter plate containing solid -LTH agar using the sterilized pinning tool. If the baits are self-activating, they have to be transferred on -LTH with a predetermined concentration of 3-AT. Incubate for 6-10 days at 30 ºC.
  - Score the interactions by looking for growing colonies that are significantly above the background by size and that are present for > 1 of the quadruplicate colonies. The plates should be examined every day. Most two-hybrid positive colonies appear within 3 to 5 days, but occasionally positive interactions can be observed later. Very small colonies are usually designated as background. However, there is no absolute measure to distinguish between the background and real positives. When there are many (i.e., > 30) large colonies per array of 6000 positions, these baits are considered as random activators. In this case the screen should be repeated.

2.2.2.9 Evaluation of raw results

Filtering of the obtained raw results significantly improves the data quality of the protein interaction set. For filtering at least three parameters should be considered. Detected protein interactions that are not reproducible should be discarded. The reproducibility is assessed from the screen by pinning quadruplicates or by conducting an independent retest of initially detected protein interactions. For each prey the number of different interacting baits is calculated. Preys interacting with a significant high number of baits - judged by evaluating the distribution of these

numbers - are assumed to interact unspecific and are neglected (“sticky preys”). The last parameter is the background activation activity of the tested bait. The activation strength of interaction pairs must be significantly higher than with remaining pairs. In principle, at least with the HIS3-reporter, no activation (no colony growth) should be observed in non-interacting pairs.

2.2.3 Bioinformatical Analysis

2.2.3.1 Y2H Interaction-data: Storage and analysis

The Yeast two-hybrid interaction data of the combined VZV screen was stored and analyzed with the help of FileMaker Pro v.8.5 databases (http://www.filemaker.de). At this point I want to thank my colleague Roman Häuser, who trained me in the use of the FileMaker software and helped me to develop the databases, which allowed the detailed data analysis presented in this study.

2.2.3.2 Protein network visualization

Protein interaction networks were visualized with the Cytoscape software package (Shannon et al., 2003). Data analysis was done with perl (www.perl.org) and R (R-Development-Core-Team, 2004). Statistical analysis was done with R and MS-Excel. Tightly connected clusters in networks were identified with the MCODE algorithm as implemented into the Cytoscape software package (Bader and Hogue, 2003). Topological parameters of networks were computed with the NetworkAnalyzer plug-in for Cytoscape (MPI for Informatics, Germany, med.bioinf.mpi-inf.mpg.de/netanalyzer). Network centralities (node degree, centroid value) were calculated with CentiBiN (Junker et al., 2006).

2.2.3.3 Network analysis

Network parameters were calculated using the "NetworkAnalyzer" plugin (Assenov et al., 2008) for Cytoscape network visualization software (Shannon et al., 2003).

2.2.3.3.1 Node degree and degree distribution

The degree of a node, abbreviated *k*, tells how many links a node has to other nodes. In undirected networks, the node degree of a node *n* is the number of edges linked to *n*. A self-loop of a node is counted like two edges for the node degree (Diestel,

2005). An undirected network with N nodes (proteins) and L edges (links, interactions) is characterized by an average degree $<k> = 2*L/N$ while self-interactions are subtracted. The degree distribution $P(k)$ of a network is defined to be the fraction of nodes in the network with degree k. Thus if there are n nodes in total in a network and n_k of them have degree k, we have $P(k) = n_k/n$.
The node degree distribution gives the number of nodes with degree k for $k = 0, 1…$

2.2.3.3.2 Clustering coefficients

In undirected networks, the clustering coefficient Cn of a node n is defined as $Cn = 2e_n/(k_n(k_n-1))$, where k_n is the number of neighbors of n and e_n is the number of connected pairs between all neighbors of n *(Barabasi and Oltvai, 2004; Watts and Strogatz, 1998)*. The clustering coefficient is a ratio N / M, where N is the number of edges between the neighbors of n, and M is the maximum number of edges that could possibly exist between the neighbors of n. The clustering coefficient of a node is always a number between 0 and 1.
The average clustering coefficient distribution gives the average of the clustering coefficients for all nodes n with k neighbors for $k = 2,….$

2.2.3.3.3 Characteristic path length and attack tolerance

The characteristic path length (CPL) is for example the number of clicks which will lead you from one website to another, in this case the number of interactions leading from one protein to another on an average within the whole network (average distance between any two proteins).
If d(X; Y) is the length of the shortest path between the nodes X and Y, then the characteristic path length, CPL, is d(X; Y) averaged over all pairs of nodes.
The robustness of the networks is investigated by a step-by-step assault strategy and subsequent recalculation of the characteristic path length. The most highly connected nodes are removed in decreasing order. After each node is removed, the new network CPL of the remaining network is plotted as a multiple or fraction of the original parameters. A higher attack tolerance is ascertained when the increase in path length is considerably smaller.

2.2.3.4 Multiple sequence alignments

Amino acid sequences were retrieved from the Uniprot Knowledgebase (UniProt-Consortium, 2010). Sequences were aligned directly from the Uniprot website (www.uniprot.org) using ClustalW version 2 online via the EBI ClustalW server (Larkin et al., 2007). Figures of Alignments were derived from the Jalview alignment editor (Waterhouse et al., 2009) using the ClustalX colour scheme. ClustalX is the graphical user interface of ClustalW derived sequence alignments. Each amino acid in the alignment is assigned a colour if the amino acid profile of the alignment at that position meets some minimum criteria specific for the residue type according to the ClustalX colour scheme (see Table 3).

Residue at position	Applied Colour	{ Threshhold, Residue group }
A,I,L,M,F,W,V	BLUE	{+60%, WLVIMAFCHP}
R,K	RED	{+60%,KR},{+80%, K,R,Q}
N	GREEN	{+50%, N}, {+85%, N,Y}
C	BLUE	{+60%, WLVIMAFCHP}
C	PINK	{100%, C}
Q	GREEN	{+60%,KR},{+50%,QE},{+85%,Q,E,K,R}
E	MAGENTA	{+60%,KR},{+50%,QE},{+85%,E,Q,D}
D	MAGENTA	{+60%,KR}, {+85%, K,R,Q}, {+50%,ED}
G	ORANGE	{+0%, G}
H,Y	CYAN	{+60%, WLVIMAFCHP}, {+85%, W,Y,A,C,P,Q,F,H,I,L,M,V}
P	YELLOW	{+0%, P}
S,T	GREEN	{+60%, WLVIMAFCHP}, {+50%, TS}, {+85%,S,T}

Table 3: ClustalX colour scheme applied to multiple sequence alignments.
The table gives these criteria as clauses: (+X%, xx, y), where X is the minimum percentage presence for any of the xx (or y) residue types. Source: www.clustal.org.

2.2.4 General protein related procedures

2.2.4.1 SDS-PAGE Gels (for 2 gels 10 cm x 8 cm with 1mm spacers)

- **Required material**

  - Gel chambers (2 gels, 10 cm x 8 cm with 1 mm spacers)
  - 0.5 M EDTA
  - 30 % Acrylamide/N,N'-methylenbisacrylamide
  - TEMED

- Protein molecular weight marker
- Buffers and solutions

Solution	Composition
Tris-HCl pH 8.8	1 M Tris-HCl pH 8.8
Tris-HCl pH 6.8	1 M Tris-HCl pH 6.8
Separating gel buffer (4x)	18.17g Tris base 4 ml 10 % (w/v) SDS adjust pH to 8.8 with HCl add DIW to 100 ml autoclave
Stacking gel buffer (4x)	6.06 g Tris base 4 ml 10 % (w/v) SDS adjust pH to 6.8 with HCl add DIW to 100 ml autoclave
10 % APS	10 % (w/v) APS in DIW store at - 20 °C
20 % SDS	20 % (w/v) SDS in DIW
1 x Laemmli running buffer	30 g tris base 144 g Glycine 100 ml 10 % (w/v) SDS add DDW to 1 l
2. Laemmli sample buffer	125 mM Tris-HCl, pH 6.8 4 % (w/v) SDS 20 % (v/v) Glycerol 10 % β-mercaptoethanol (or 100 mM DTT) 0.004 % (w/v) Bromophenol blue

- **Gel preparation**

Component	Stacking Gel (5%)	Separating Gel (10%)
30 % Acrylamide/N,N'-methylenbisacrylamide	1.5 ml	4 ml
Separation gel buffer (4x)	-	3 ml
Stacking gel buffer (4x)	2 ml	-
DIW	4.6 ml	5 ml
10 % APS	15 µl	20 µl
TEMED	40 µl	60 µl

Pour separating gel (add reagents as indicated above) into gel chamber (so that it fills up 2/3 of the chamber) and cover with isopropanol and let solidify.
Remove isopropanol and pour stacking gel. Immediately insert the combs and let solidify.

2.2.4.2 Coomassie Blue staining of PAGE Gels

- **Required buffers and solutions**

Solution	Composition
Coomassie staining solution	0.2 % (w/v) Coomassie brilliant blue R250 50 % (v/v) Methanol 10 % (v/v) Glacial acetic acid
Coomassie destaining solution	30 % (v/v) Methanol 10 % (v/v) Glacial acetic acid

- Incubate protein gel in Coomassie staining solution for 30 min
- Recover Coomassie stain (can be used several times)
- Incubate the Coomassie stained gel in destaining solution (change destaining solution several times) until the background is sufficiently reduced
- Recover destaining solution by filtering through activated charcoal
- Rinse the gel in water and dry the gel for longer storage

2.2.4.3 Western Blot (semi-dry)

- **Required material**
  - Blot chamber
  - Whatman paper
  - PVDF membrane
  - Buffers and solutions

Solution	Composition
Semi-dry western blot transfer buffer	25 mM Tris base 190 mM Glycine 20 % (v/v) Glycerol 0.05 % (w/v) SDS
Ponceau-S staining solution	0.25 % (w/v) Ponceau-S 40 % (v/v) Methanol 15 % (v/v) Glacial acetic acid
Membrane blocking solution	5 % (w/v) Skimmed milk powder in TBS(-T)
TBS-T	50 mM Tris-HCl 150 mM NaCl 0.2 % (v/v) Tween-20 Adjust pH 8.0 with 10 M NaOH

- **Western blotting procedure**
  - Cut blot paper (6 sheets of Whatman paper per gel) according the size of the gel and equilibrate in transfer buffer
  - Cut blotting membrane with the same size and activate in methanol for 2 minutes, then equilibrate in transfer buffer
  - Disassemble electrophoresis chamber and carefully transfer gel (cut off stacking gel) into a tray with transfer buffer, briefly equilibrate
  - Build up blot: To bottom platinum anode of Blotting-chamber place:
    - Pre-wet filter paper (3 sheets of Whatman paper)
    - Pre-wet membrane
    - Gel
    - Pre-wet filter paper (3 sheets of Whatman paper)
    - Roll out air bubbles
  - Secure safety cover and connect to power supply
  - Run for 60 min at 110 mA per gel
  - Stop transfer, discard filter paper and briefly wash blot in H_2O
  - Check transfer by staining blot in Ponceau S solution and destain in H_2O
  - Blocking: Incubate membrane for 30 min with blocking solution
  - Incubate with 1^{st} antibody (1:10.000 dilution) in blocking solution for 1 h at RT (alternatively o/n at 4 °C)
  - Wash 3 x with TBS-T for 10 min each step

- Incubate with 2nd antibody (1:5.000 dilution; HRP labeled) in TBS-T for 1 h at RT
- Wash 3 x with TBS-T and perform chemoluminescence detection for bound antibodies

2.2.4.4 Protein expression and purification

- **Required material**
  - *E. coli* BL21-CodonPlus® (DE3)-RIL cells containing the respective expression clone
  - LB liquid medium with the appropriate antibiotic
  - Maltose and Amylose resin High Flow for MBP-Tag purification
  - Imidazole and Ni-NTA-Agarose for His-Tag purification
  - Buffers and solutions:

Compound	Composition
Washing buffer	50 mM Tris-HCl, pH 8.0 150 mM NaCl 1 mM PMSF 1 mM EDTA 1 mM DTT (20 mM Imidazole for His-Tag purification)
Lysis buffer	Washing buffer 0.5 % Igepal CA-630 50 µg/ml Lysozyme 1 mM PMSF Complete protease inhibitor mix stock solution, 1:50 dilution (10 mM Imidazole for His-Tag purification)
Proteinase inhibitor stock solutions	200 mM PMSF dissolved in isopropanol Complete protease inhibitor mix tablets (1 Tablet/2 ml DIW)
IPTG stock solution	1 M IPTG in sterile water
AHT stock solution	2 mg/ml
Elution buffer (for MBP-Tag purification)	Washing buffer 10 mM Maltose (from 1M stock solution)
Elution buffer (for His-Tag purification)	Washing buffer 250 mM Imidazole (from 1M stock solution)

- **Procedure**

  - Start with the bacterial expression clone
  - Inoculate 5 ml LB liquid medium plus appropriate antibiotic with a single colony, shake o/n at 37 °C
  - Dilute culture in 250 ml LB liquid medium with antibiotic
  - Incubate at 37 °C until OD_{600} = 0.6
  - Induce with 1 mM IPTG or 0.2 µg/ml AHT
  - Shake 4-5 h at selected temperature (RT to 37 °C)
  - Centrifuge 10 min at 6000 rpm
  - Discard supernatant (storing at -80 °C possible) and resuspend cells in 10 ml lysis buffer
  - incubate for ~30 min on ice
  - Shear DNA by ultrasonic impulses with Bioruptor (3 x ~15 s pulses, keep on ice)
  - Remove 15 µl aliquot for SDS-PAGE (whole cellular lysate)
  - Centrifuge at 15.000 x g for 30 min at 4 °C
  - Remove 15 µl aliquot for SDS-PAGE (soluble protein extract)
  - Transfer supernatant to 15 ml conical tube

  - Affinity purification
    - Take 500 µl Amylose Resin High Flow (MBP-Tag) or Ni-NTA-Agarose (HIS-TAG), 50 % slurry
    - Wash 3 x with 10 ml TBS
    - Add supernatant
    - Rotate 30 min at RT
    - Wash 3 x with washing buffer (rotate for 15 min at 4 °C)
    - Elute 3 x with 1 ml elution buffer (5 min each, resuspend pellet gently several times). Remove 15 µl aliquot for SDS-PAGE (Eluates 1-3)
    - Add 10 % sterile Glycerol and store at -20 °C

2.2.5 SPOT Peptide Synthesis

SPOT synthesis which was originally introduced by Ronald Frank is a convenient and versatile technique for simultaneous and parallel solid phase synthesis of peptides at distinct positions on a membrane support (Frank, 2002).The SPOT method has opened up great opportunities to synthesize and subsequently screen large arrays of synthetic peptides (Landgraf et al., 2004; Otte et al., 2003). Peptide arrays prepared by SPOT synthesis can be used to study molecular recognition events such as epitope mapping, the analysis of protein-protein and the identification of biologically active peptides. Peptide arrays can be applied to precisely depict molecular recognition events on the single amino acid level.

2.2.5.1 Preparation of peptide arrays by SPOT peptide synthesis

2.2.5.1.1 Materials

- **Solvents**
  - N,N′-dimethylformamide (DMF)
  - Methanol or ethanol (MeOH or EtOH)
  - N-methylpyrrolidone (NMP)
  - Dichloromethane (DCM)
- **Preparation of membranes**
  - Cellulose membranes for the preparation of peptide arrays are prepared from filter paper Whatman 50 (alternatively Whatman 540 or Chr1). Ready-to-use modified membranes for synthesis of peptides are available from INTAVIS.
  - Diisopropylcarbodiimide (DIC)
  - N-methylimidazole (NMI)
  - Fmoc-β-alanine
- **Preparation of activated amino acid solutions**
  - Fmoc protected amino acids for SPPS (Table 1)
  - Coupling reagents: Diisopropylcarbodiimide (DIC)
  - N-hydroxybenzotriazole (HOBt)

- **SPOT synthesis on cellulose membranes**
  - Solution for Fmoc-deprotection: 20 % (v/v) piperidine in DMF
  - Capping solution: 2 % (v/v) acetic anhydride in DMF
  - Cleavage solution I: 90 % trifluoroacetic acid (v/v) (TFA) + 5 % (v/v) H_2O + 3 % (v/v) triisopropylsilane (TIPS) + 1 % (v/v) phenol + 1 % (v/v) DCM
  - Cleavage solution II: 60 % (v/v) TFA + 3 % (v/v) TIPS + 2 % (v/v) H_2O + 1 % (v/v) phenol + 44 % (v/v) DCM

2.2.5.1.2 Methods

Spot synthesis was carried out fully automated using a MultiPep spotter from INTAVIS AG (Köln) (Figure 15). Individual spots were placed with a pipetting volume of 0.2 µl.

- **Amine functionalization by esterification of filter paper**

- Preparation of a Cellulose Membrane for Spot Synthesis
  - Cut a piece of Whatman 50 paper with the size of 10 cm × 15 cm
  - Amine functionalization of the filter paper:
    - Dissolve 0.8 g Fmoc-β-alanine in 12.5 ml amine-free DMF.
    - Add 468 µl DIC and 397 µl NMI, mix well and transfer the reaction solution into a chemically resistant box with lid. Avoid air bubbles under the paper during placement of the filter paper in the box and ensure that the surface of the membrane is slightly covered by the solution.
    - Close the box and leave the membrane in the reaction mixture overnight
  - Wash the membrane 3 x with DMF for at least 30 s each. The membrane can be stored at -20 °C for several months.
  - For storage, wash the modified membrane at least twice with MeOH or EtOH and air dry in a fume hood. When needed, equilibrate the membrane to room temperature and wash the membrane 1 x with DMF for at least 20 min prior to the deprotection step.
  - Fmoc-deprotection: treat the membrane twice with 20 % (v/v) piperidine in DMF for at least 5 min each.

Figure 15: Fully automatic spot synthesizer.
The MultiPep Spotter by Intavis AG, Köln.

- **Preparation of activated amino acid solutions**

Peptide synthesis was performed using pre-activated Fmoc-protected amino acids. The advantage of this method is the use of only one reagent for each solution, making the preparation of the amino acid solutions very simple and the likelihood of mistakes low.

- Preparation of coupling solutions

- Weigh 0.5 mmol of each amino acid (see Table 1)
- Add for each amino acid 1 ml of HOBt (2.3 g in 20 ml NMP)
- Dissolve amino acids and fill up to 1.5 ml with NMP
- Aliquot 4 x 360 µl in 1.5 ml microcentrifuge tubes
- Store at -20 °C for up to 6 months

At room temperature, the solutions are stable for about 1 day. Therefore, before starting the first synthesis cycle of a day, always discard the solutions from the previous day and replace them by a freshly activated aliquot from the stock solutions.

- **Peptide synthesis**

  - Day 1

  - Thaw one set of amino acid aliquots o/n at RT
  - Activate each aliquot with 240 µl Activation Stock Solution:
    - 0.4 ml DIC
    - 2.0 ml NMP
    - 2.8 ml DMF
  - Mix and activate for 30 min at RT
  - Centrifuge 3 min at 4000 rpm
  - Pipette solution into the synthesizer vials in correct order. Avoid possibly emerged needles of urea.
  - Fill up solvents and reagents for coupling steps:
    - DMF
    - Methanol
    - 20 % Piperidine/DMF
    - Capping solution (0.3 ml Acetic Anhydride in 15 ml DMF)
  - Activate membrane in DMF
  - Start Program
  - Thaw amino acids aliquots overnight

  - Day 2 and 3

  - Activate amino acids
  - Discard old AA solutions and pipette freshly activated amino acids
  - Fill up capping and piperidine solutions
  - Thaw amino acids overnight at RT

  - After the last deprotection step, let the membrane dry o/n in a fume hood

- **TFA removal of all side chain protecting groups**

  - Incubate the membrane 30 min in 25 ml cleavage solution I (without shaking)
  - Wash 3x 1 min with DCM
  - Incubate the membrane 3 h in 25 ml cleavage solution II (without shaking)
  - Wash at least 5 x 1 min with water until pH is about 7
  - Wash 3 x 3 min with MeOH

- Let dry o/n in a fume hood

2.2.5.1.3 Protein hybridization and detection

- Activate the membrane for 10 min with MeOH
- Wash 3 x 5 min with TBS
- Block 3 h with blocking solution (5% skimmed milk powder in TBS)
- Incubate 10 µg/ml purified protein of interest in blocking solution for 1 h at RT
- Wash 3 x 5 min with 1x TBS
- Incubate 1 h at RT with primary antibody in blocking solution + 0.05 % Tween
- Wash 3 x 5 min with 1x TBS-T
- Incubate 1.5 h at RT with secondary antibody in blocking solution + 0.05 % Tween
- Wash 3 x 5 min with 1x TBS-T
- Mix ECL-detection kit solutions 1:1
- Pipette the mixture on to the membrane and let activate for 5 min
- Expose to light-sensitive film

2.2.5.1.4 Reuse of peptide arrays

- **Required solutions**

Solution	Composition
Stripping solution I	50 % (v/v) Ethanol 10 % (v/v) Glacial acetic acid
Stripping solution II	2 % (w/v) SDS 0.8 % (w/v) DTT

- **Procedure**

  - Incubate membranes alternating in stripping solution I and II for three times and for ½ h each. Incubation in stripping solution II is performed in a water bath at 55 °C. Between the washing steps rinse the membranes with deionized water.
  - After the last washing step rinse the membrane three times with DIW and finally three times with Ethanol.
  - Let the membranes dry o/n under a fume hood.

3 Results

3.1 Yeast two-hybrid screening of the VZV ORFeome

A previous Y2H screen of the VZV interactome has been published by Uetz and coworkers in 2006 (Uetz et al., 2006). In this study, a novel approach has been established to gather a more complete insight into the complex network of the VZV protein-protein interactions. Beyond that, I developed a tool to improve the method itself for future requirements. Therefore, a parallelly generated library of VZV ORFs lacking the stop codon was used, kindly provided by Prof. Dr. Jürgen Haas (Max-von-Pettenkofer Institut, München). The Library was generated by Pothineni Venkata Raveend under supervision of Dr. Armin Baiker, who was in the research group around Prof. Dr. Haas. Thereby it was possible to use the ORF collection for combinatorial screening with N- and C-terminal fusion tags.

3.1.1 Optimization of the VZV ORFeome collection

Prior to the restatement of the VZV interactome by this new screening approach, the ORFs 21 and 22N (N-terminal domain of ORF22) had to be reconstructed. The reason therefore was negative verification of the respective entry-clones by sequencing. As template for subsequent Gateway-cloning, cDNA from VZV pOKA-infected MeWo melanoma cells was used (Takahashi et al., 1974) (GenBank accession no. AB097933.1). As 1st round primers, I used VZV_ORF21_FW/ VZV_ORF21_REV and VZV_ORF22N_FW/VZV_ORF22N_REV, respectively. The 2nd round PCR was performed with those PCR-Products as template using the attB1 and attB2 primers. BP-recombination into pDONR207 was followed by LR-recombination into the Y2H destination vectors pGADT7g, pGADCg (prey plasmids) and, pGBKT7g and pGBKCg (bait plasmids). Preys were transformed into AH109 and baits into Y187 yeast reporter strains. The constructs were verified by analytical PCRs after each recombination step. Additionally, the entry clones were sequenced to assure the accuracy of the insert.

3.1.2 Generation of Y2H destination vectors with C-terminal fusion tags

The Yeast two-hybrid vectors pGBKT7g and pGADT7g generate N-terminal fusions of DNA binding (DBD) and activation domain (AD) fusions, respectively. The derivative vectors pGBKCg and pGADCg were designed to fuse the DBD and AD to the C-terminus of inserted ORFs (see Figure 16).

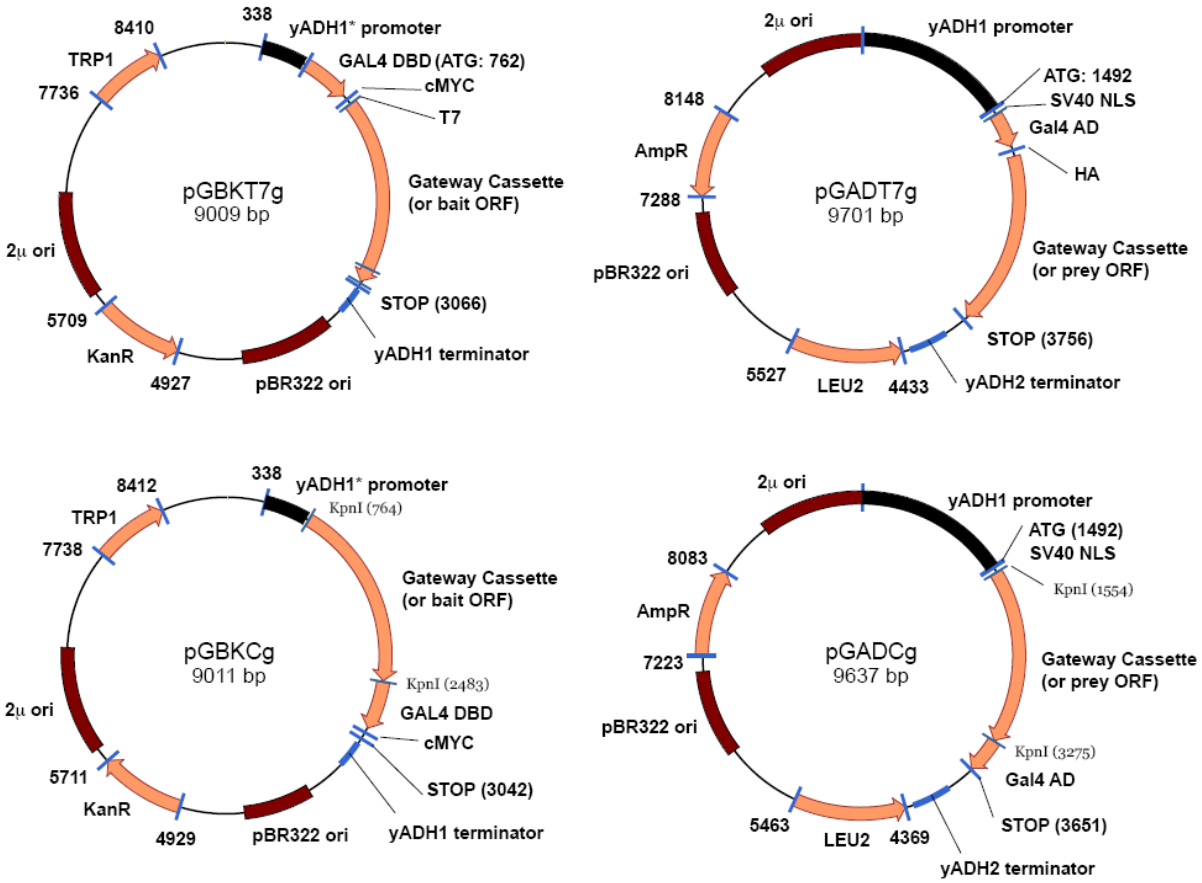


Figure 16: Gateway bait and prey vectors with N- and C-terminal fusion tags.
The parental vectors pGBKT7g and pGADT7g generate N-terminal fusions of DNA-binding (DBD) and activation domain (AD) fusions, respectively. The new vectors pGBKCg and pGADCg fuse DBD and AD at the C-terminus of inserted ORFs. Note that both pGBK- vectors contain a truncated ADH promoter (indicated by *) which may reduce expression levels and thus interaction signals (Ammerer, 1983; Legrain et al., 1994; Tornow and Santangelo, 1990). A second derivative of pGBKT7g, which was used in Chapter 3.2.5 is pGBGT7g where the Kanamycin resistance was replaced by Gentamicin resistance.(Stellberger et al., 2010)

3.1.2.1 Conversion of pGADT7g to pGADCg

The conversion of the parental vector pGADT7g into pGADCg was performed in three major steps:

1) Removal of the region downstream of the Gal4 AD.

- Double digestion of pGADT7g with BglII and ClaI to remove the coding region between Gal4 AD and the stop codon TGA (see Figure 17)
- Purification of the vector backbone by Agarose-Gelelectrophoresis
- 5´-Dephosphorylation of linearized vector by SAP treatment

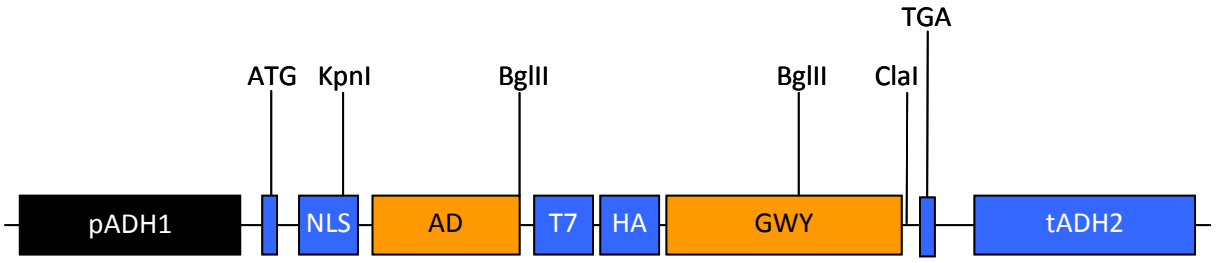


Figure 17: Transcribed region of pGADT7g flanked by promoter and terminator.
A schematic view of the vector region which is relevant for transcription and the nature of the prey-constructs expressed in the Y2H test strains. pADH1: Yeast ADH1 promoter; ATG: Start codon; NLS: Nuclear localization Signal derived from SV40 (Simian Virus 40); AD: Gal4 transactivation domain; T7: Bacteriophage T7 promoter; HA: Human influenza hemagglutinin; GWY: Gateway-cassette, reading frame B (Invitrogen); TGA: Stop codon; Restriction nuclease recognition sites as indicated.

2) Ligation of an XbaI-HA-STOP-BclI Fragment.
   - The reverse-complementary oligonucleotides pGADC_XbaI/BclI_FW and pGADC_XbaI/BclI_REV were dimerized and ligated into the linearized pGAD-backbone. The Result is an intermediate vector lacking the cloning site
   - The oligonucleotides were designed that the dimers bear 5´-overhangs that are complementary to the 3´-overhangs of the digested vector. The inserted HA-tag and stop codon could thereby be replaced by cloning with XbaI and BclI. (see Figure 18)
   - The intermediate vector was checked by analytical PCR using pADH1_FW and GAL4_AD_REV primers and subsequently sequenced.

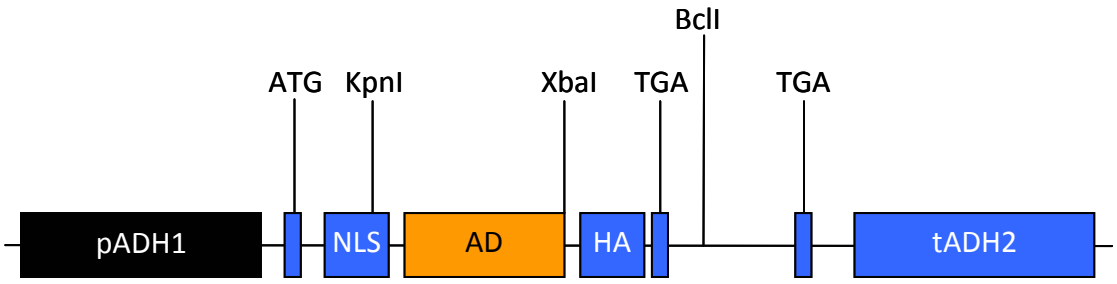


Figure 18: pGADC intermediate vector.
Compared to Figure 17 the T7-HA-GWY region is replaced by HA followed by a TGA stop codon.

3) Insertion of the Gateway-cassette between the NLS and the Gal4 AD.

In the last work step a novel Gateway-cassette has been introduced into the intermediate vector to generate pGADCg (Figure 19).

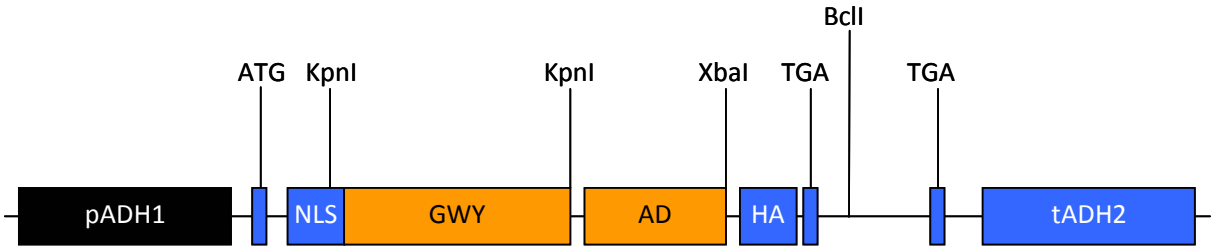


Figure 19: Resulting vector pGADCg
The Gateway cassette has been ligated into the KpnI restriction site (cf. Figure 18)

- To generate copies of a Gateway-cassette that can be cloned between the NLS and AD-region a KpnI restriction site was fused by a preparative PCR using pGADT7g as a template. The cassette has the reading frame B (RFB, see Gateway manual, Invitrogen) with the primers KpnI-GWY RFB-KpnI_FW and KpnI-GWY RFB-KpnI_REV.
- Cleanup of the PCR product was followed by a KpnI digest.
- Parallel the intermediate vector was digested with KpnI and dephosphorylated and cleaned up.
- The cassette was ligated into the vector and transformed into DB3.1-cells. Colony-PCR was performed to identify positive clones with inserted Gateway-cassette (pADH1_FW and Gal4_AD_REV primers).
- The correct orientation of the cassette was controlled by sequencing.

The resulting Vector was amplified within DB3.1 cells and purified by plasmid-DNA maxi preparation. Now any desired protein coding sequence or fragment thereof can be cloned into this vector by LR-recombination reaction as long as it is available as Entry-clone with other than an ampicillin resistance marker. A detailed view of the recombination event is given in Figure 20.

3.1.2.2 Conversion of pGBKT7g to pGBKCg

The Yeast two-hybrid vector for expression of C-terminally fused DBD domain was also performed in three main steps:

1) Removal of the Gateway-cassette.
   - First a double digestion of pGBKT7g (Figure 21) with EcoRI/PstI was performed. There are for both enzymes two restriction sites, each with one flanking the recombination cassette and a second one within it. Therefore special attention was paid to achieve a complete as possible turnover of plasmids to be digested.

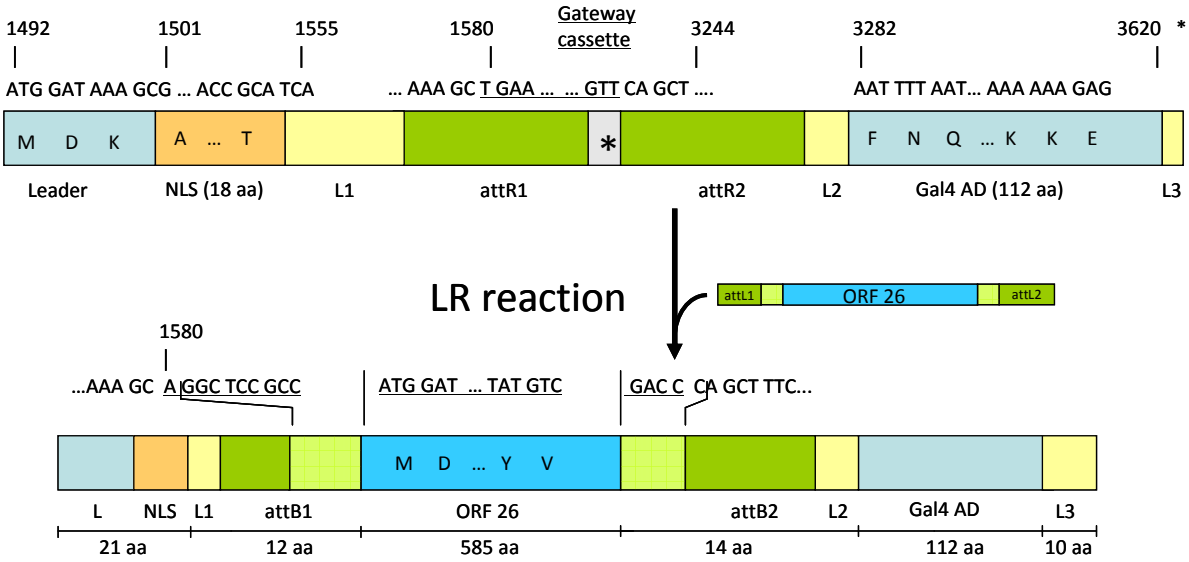


Figure 20: Example: Recombination of VZV-ORF26 into pGADCg
The prey vector pGADCg generates a C-terminal fusion of the Gal4 AD (Stellberger et al., 2010). The cloned ORF (here: VZV ORF 26) is preceded by a 33 amino acid sequence that contains the nuclear localization signal (NLS). A 14 amino acid linker separates the ORF from the Gal4 AD. The HA tag of pGADCg has been shortened to seven amino acids so that it may not be recognized by anti-HA antibodies. The LR-recombination recognizes attR (Destination vector) and attL sites (Donor-vector). The resulting attB sites, which were part of the attL sites (shown in light green), create short AA linkers at the N- and C-Terminus of the insert, ORF26. Underlined sequences differ between vector and final product and thus represent the inserted ORF plus a few additional bases introduced by the clone originating from attL1/2. For further technical details see the Gateway manuals at http://www.invitrogen.com.

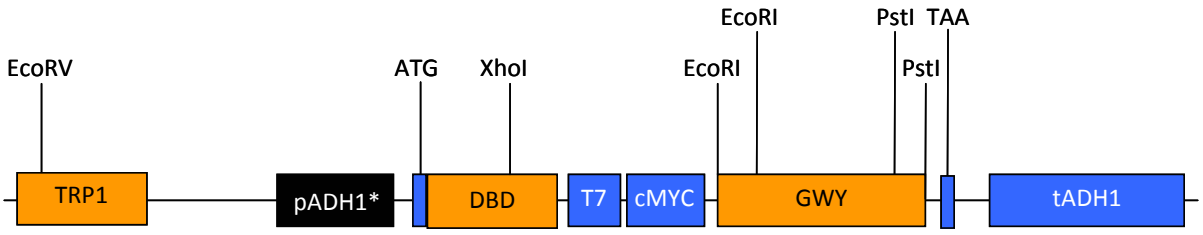


Figure 21: Sector of pGBKT7g implicated in the conversion to pGBKCg.
This is an illustration of the vector region which is relevant for transcription and the expressed bait-constructs. pADH1*: Truncated yeast ADH1 promoter; ATG: Start codon; DBD: Gal4 DNA binding domain; T7: Bacteriophage T7 promoter; MYC: Coding for human cMYC; GWY: Gateway-cassette, reading frame B (Invitrogen); TAA: Stop codon; Relevant restriction nuclease recognition sites are indicated.

- EcoRI leaves a 4 nucleotide 5´-TTAA overhang while PstI-restriction results in a 4 nt 3´-ACTG overhang. This resulted in an 8 nt gap on the sense-strand which was closed by oligonucleotide-ligation of the pGBKC_8mer with the sequence: 5´-AATTTGCA-3´.

- The reaction was transformed into TOP10-cells. Loss of the Gateway-cassette was confirmed by colony-PCR using the primers pADH1_FW and GAL4_DBD_REV. The resulting plasmid was the first progenitor of pGBKCg (see Figure 22A).
- Insertion of a restriction site upstream of the DBD

To generate a restriction site suitable for insertion of the recombination cassette in frame between the promoter region and the DNA binding domain I used a fusion-PCR strategy. The goal of this step was to obtain a PCR-product carrying the generated KpnI restriction site and spanning the neighbored sequences until unique restriction sites. These are subsequently used to exchange the PCR-product with the corresponding vector sequence.
Three separate PCR reactions were performed (see also Figure 22B):

- The first PCR was performed with pGBKT7g as template using the primers pGBKC_A (A) and pGBKC_B (B). The reverse primer B contains a non-specific sequence bearing the KpnI restriction site (dotted line).
- A second PCR reaction, using primers pGBKC_C (C) and pGBKC_D (D) replicates the sequence downstream of the designated KpnI site with and fusing it to the PCR-product with 28 nt homology to the first PCR product (dotted line).
- The third PCR reaction, using primers pGBKC_A (A) and pGBKC_D (D) produced the desired DNA-molecule by fusing the gel purified PCR-products from PCR #1 and #2 which were both used as templates generating the desired DNA molecule (Figure 22C).

The DNA resulting from the fusion-PCR was double digested with EcoRV/XhoI as well as the progenitor vector which was additionally dephosphorylated. The Backbone of the digested progenitor was gel purified and ligated with the column-purified digest of the PCR. The resulting transformants were checked by pADH1_FW/GAL4_DBD_REV and subsequently sequenced resulting in the intermediate pGBKC vector.

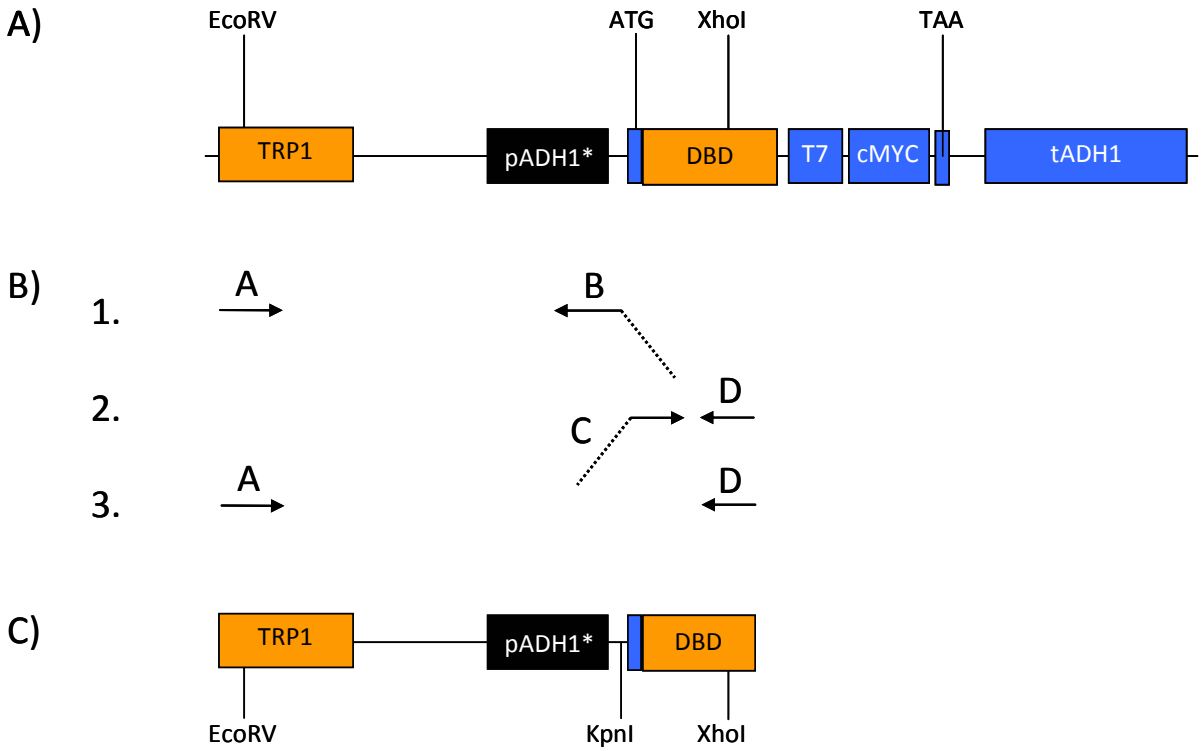


Figure 22: Fusion-PCR strategy to generate KpnI restriction site.
A) First progenitor of pGBKCg. To generate bait constructs with a C-terminal DBD-fusion the Gateway cassette had to be inserted between pADH1* and the start codon ATG. B) A Fusion-PCR strategy was used and three consecutive PCR steps led to the DNA molecule (C) for EcoRV/XhoI cloning in the progenitor vector.

2) Insertion of the Gateway-cassette in front of the Gal4 DBD.

The insertion of the recombination cassette was performed as described for pGADCg in chapter 3.1.2.1 yielding the vector pGBKCg which was finally verified by PCR and sequencing (Figure 23).

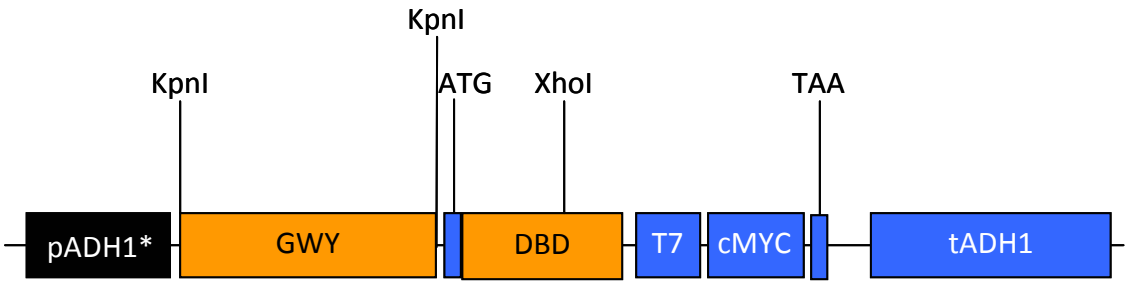


Figure 23: Resulting vector pGBKCg
The vector has the Gateway-cassette inserted between the truncated ADH1-promoter (pADH1*) and the original start codon which now has to be generally replaced by an intrinsic start codon of the test constructs which are later on recombinated in the destination vector pGBKCg.

The vector was amplified in DB3.1 cells and purified by plasmid-DNA maxipreparation. By LR-recombination reaction every protein coding DNA sequence can be cloned into this vector from a Gateway entry-clone with other than a kanamycin resistance marker. A detailed view of the recombination event is given in Figure 24.

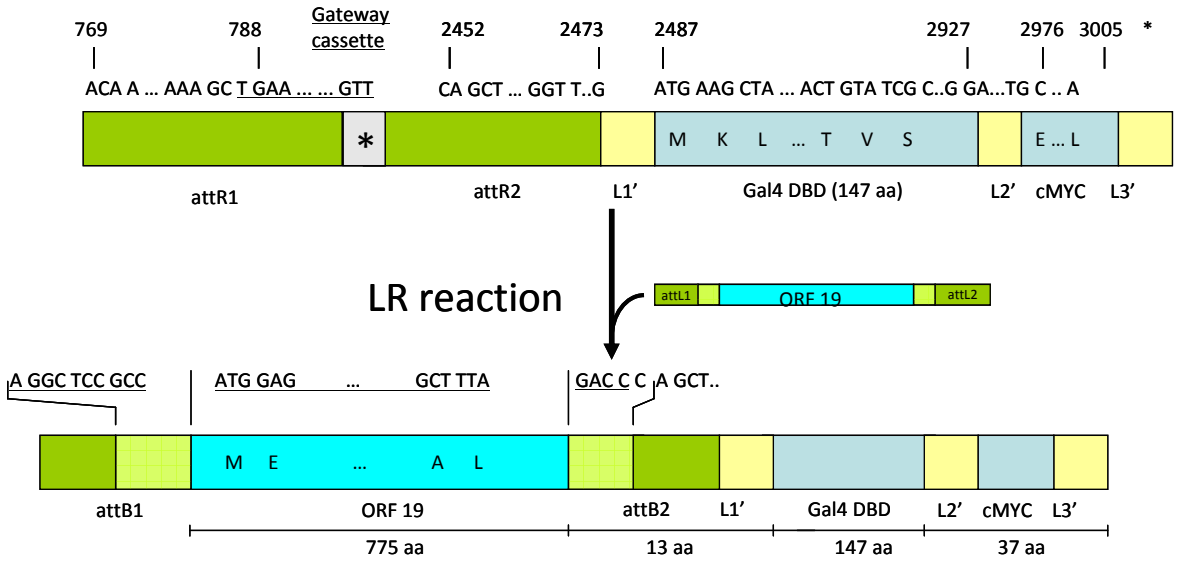


Figure 24: Example: Recombination of VZV-ORF19 into pGBKCg
The vector pGBKCg generates a C-terminal fusion of the Gal4 DBD separated by a 13 amino acid linker from the N-terminally fused ORF (here: VZV ORF 19) (Stellberger et al., 2010). The NLS of pGBKC is part of the DBD. A MYC tag is embedded in a C-terminal 37 amino acid tail. The LR-recombination recognizes attR (destination vector) and attL sites (donor-vector). The resulting attB sites, which were part of the attL sites (shown in light green), create short AA linkers at the N- and C-Terminus of the insert. Underlined sequences differ between vector and final product and thus represent the inserted ORF plus a few additional bases introduced by the clone originating from attL1/2.

3.1.2.3 Resources of pGBKCg and pGADCg

The vectors pGBKCg and pGADCg are available from Dr. Peter Uetz, Addgene (http://www.addgene.com) or myself. The sequences have been deposited with GenBank under accession numbers FJ696409 (pGBKCg) and FJ696408 (pGADCg). The vectors pGBKT7g and pGADT7g are available from Dr. Peter Uetz.

3.1.3 Combinatorial Y2H screening of the VZV ORFeome

The goal of this project was to generate a complete as possible protein interaction map of an organism at the example of the Varicella zoster virus. The big advantage of this model is was the availability of a good quality collection of entry clones. Moreover, the collection contains 96 full-length and domain constructs of 71 distinct

open reading frames, which can be handled very convenient in an array-based Y2H-screen.

The first step was to clone each ORF into the bait vectors pGBKT7g and pGBKCg and the prey vectors pGADT7g and pGADCg. An overview about the VZV ORFs and the annotated proteins is shown in Table 4. The resulting clones were transfected into the Y2H-reporter yeast strains Y187 (bait-) and AH109 (prey-clones) and placed on a 96 positions array. In an array-based Yeast two-hybrid screen as described in the Methods chapter baits were screened against pGADT7g- and pGADCg prey arrays.

Product Name	Length	Accession	GeneID	Locus	Locus_tag
membrane protein UL56	129	YP_053044.1	1487696	ORF0 (ORFS/L)	HHV3_gp01
membrane protein V1	108	NP_040124.1	1487664	ORF1	HHV3_gp02
myristylated tegument protein CIRC	221	NP_040125.2	1487680	ORF2	HHV3_gp03
nuclear protein UL55	179	NP_040126.1	1487681	ORF3	HHV3_gp04
multifunctional expression regulator	452	NP_040127.1	1487672	ORF4	HHV3_gp05
envelope glycoprotein K	340	NP_040128.1	1487673	ORF5	HHV3_gp06
helicase-primase, primase subunit	1083	NP_040129.1	1487676	ORF6	HHV3_gp07
tegument protein UL51	259	NP_040130.1	1487677	ORF7	HHV3_gp08
deoxyuridine triphosphatase	396	NP_040131.1	1487671	ORF8	HHV3_gp09
envelope glycoprotein N	87	YP_068406.1	4711773	ORF9A	HHV3_gp10.5
tegument protein VP22	302	NP_040132.1	1487674	ORF9	HHV3_gp11
transactivating tegument protein VP16	410	NP_040133.1	1487675	ORF10	HHV3_gp12
tegument protein VP13/14	819	NP_040134.1	1487654	ORF11	HHV3_gp13
tegument protein VP11/12	661	NP_040135.1	1487655	ORF12	HHV3_gp14
thymidylate synthase	301	NP_040136.1	1487659	ORF13	HHV3_gp15
envelope glycoprotein C	560	NP_040137.1	1487660	ORF14	HHV3_gp16
envelope protein UL43	406	NP_040138.1	1487652	ORF15	HHV3_gp17
DNA polymerase processivity subunit	408	NP_040139.1	1487653	ORF16	HHV3_gp18
tegument host shutoff protein	455	NP_040140.1	1487714	ORF17	HHV3_gp19
ribonucleotide reductase subunit 2	306	NP_040141.1	1487715	ORF18	HHV3_gp20
ribonucleotide reductase subunit 1	775	NP_040142.1	1487716	ORF19	HHV3_gp21
capsid triplex subunit 1	483	NP_040143.1	1487685	ORF20	HHV3_gp22
tegument protein UL37	1038	NP_040144.1	1487686	ORF21	HHV3_gp23
large tegument protein	2763	NP_040145.1	1487704	ORF22	HHV3_gp24
small capsid protein	235	NP_040146.1	1487705	ORF23	HHV3_gp25

Product Name	Length	Accession	GeneID	Locus	Locus_tag
nuclear egress membrane protein	269	NP_040147.1	1487693	ORF24	HHV3_gp26
DNA packaging protein UL33	156	NP_040148.1	1487665	ORF25	HHV3_gp27
DNA packaging protein UL32	585	NP_040149.1	1487694	ORF26	HHV3_gp28
nuclear egress lamina protein	312	NP_040150.2	1487666	ORF27	HHV3_gp29
DNA polymerase catalytic subunit	1194	NP_040151.1	1487712	ORF28	HHV3_gp30
single-stranded DNA-binding protein	1204	NP_040152.1	1487713	ORF29	HHV3_gp31
DNA packaging terminase subunit 2	770	NP_040153.1	1487661	ORF30	HHV3_gp32
envelope glycoprotein B	931	NP_040154.2	1487662	ORF31	HHV3_gp33
protein V32	143	NP_040155.1	1487663	ORF32	HHV3_gp34
capsid maturation protease	605	NP_040156.1	1487717	ORF33	HHV3_gp35
capsid scaffold protein	302	YP_068407.1	4711772	ORF33.5	HHV3_gp35.5
DNA packaging tegument protein UL25	579	NP_040157.1	1487687	ORF34	HHV3_gp36
nuclear protein UL24	258	NP_040158.1	1487688	ORF35	HHV3_gp37
thymidine kinase	341	NP_040159.1	1487667	ORF36	HHV3_gp38
envelope glycoprotein H	841	NP_040160.1	1487668	ORF37	HHV3_gp39
tegument protein UL21	541	NP_040161.1	1487706	ORF38	HHV3_gp40
envelope protein UL20	222	NP_040162.2	1487707	ORF39	HHV3_gp41
major capsid protein	1396	NP_040163.1	1487708	ORF40	HHV3_gp42
capsid triplex subunit 2	316	NP_040164.1	1487669	ORF41	HHV3_gp43
DNA packaging terminase subunit 1	747	NP_040165.1	1487719	ORF42	HHV3_gp44
DNA packaging tegument protein UL17	676	NP_040166.1	1487670	ORF43	HHV3_gp45
tegument protein UL16	363	NP_040167.1	1487718	ORF44	HHV3_gp46
tegument protein UL14	199	NP_040168.1	1487720	ORF46	HHV3_gp47
tegument serine/threonine protein kinase	510	NP_040169.1	1487678	ORF47	HHV3_gp48
deoxyribonuclease	551	NP_040170.1	1487679	ORF48	HHV3_gp49
myristylated tegument protein	81	NP_040171.1	1487656	ORF49	HHV3_gp50
envelope glycoprotein M	435	NP_040172.1	1487658	ORF50	HHV3_gp51
DNA replication origin-binding helicase	835	NP_040173.1	1487657	ORF51	HHV3_gp52
helicase-primase subunit	771	NP_040174.1	1487721	ORF52	HHV3_gp53
tegument protein UL7	331	NP_040175.1	1487722	ORF53	HHV3_gp54
capsid portal protein	769	NP_040176.1	1487723	ORF54	HHV3_gp55
helicase-primase helicase subunit	881	NP_040177.1	1487682	ORF55	HHV3_gp56
nuclear protein UL4	196	NP_040178.2	1487683	ORF56	HHV3_gp57
protein V57	71	NP_040179.1	1487684	ORF57	HHV3_gp58

Product Name	Length	Accession	GeneID	Locus	Locus_tag
nuclear protein UL3	221	NP_040180.1	1487690	ORF58	HHV3_gp59
uracil-DNA glycosylase	305	NP_040181.1	1487691	ORF59	HHV3_gp60
envelope glycoprotein L	159	NP_040182.1	1487692	ORF60	HHV3_gp61
ubiquitin E3 ligase ICP0	467	NP_040183.1	1487698	ORF61	HHV3_gp62
transcriptional regulator ICP4	1310	NP_040184.1	1487699	ORF62	HHV3_gp63
regulatory protein ICP22	278	NP_040185.1	1487700	ORF63	HHV3_gp64
virion protein US10	180	NP_040186.1	1487701	ORF64	HHV3_gp65
membrane protein US9	102	NP_040187.1	1487702	ORF65	HHV3_gp66
serine/threonine protein kinase US3	393	NP_040188.1	1487703	ORF66	HHV3_gp67
envelope glycoprotein I	354	NP_040189.1	1487689	ORF67	HHV3_gp68
envelope glycoprotein E	623	NP_040190.1	1487709	ORF68	HHV3_gp69
virion protein US10	180	NP_040191.1	1487710	ORF64	HHV3_gp70
regulatory protein ICP22	278	NP_040192.1	1487711	ORF63	HHV3_gp71
transcriptional regulator ICP4	1310	NP_040193.1	1487695	ORF62	HHV3_gp72

Table 4: Annotated VZV proteins and their corresponding ORFs.
Protein annotations were derived from the NCBI genome database. Reference sequence (Refseq) accession no.: NC_001348 (http://www.ncbi.nlm.nih.gov/nuccore/NC_001348). Table adapted from the associated table of protein coding genes (http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genome&Cmd=Retrieve&dopt=Protein+Table&list_uids=10044). Note that ORFs # 63 and # 64 are duplicated (loci HHV3_gp64 and HHV3_gp71 are ORF63 and HHV3_gp65 and HHV3_gp70 both contain ORFs # 64.

3.1.3.1 Generation of an extended PI network

The starting point of this study was a set of 182 intraviral VZV interactions (173 non-redundant) previously identified in our group (Uetz et al., 2006). This study was performed with standard N-terminal fused Y2H-constructs in the vectors pGBKT7g and pGADT7g which were also part of this study.

Combining this data with the interactions of this study, in four distinct screens testing intraviral interactions in the bait-prey combinations:

- pGBKT7g-pGADT7g (NN),
- pGBKT7g-pGADCg (NC),
- pGBKCg-pGADT7g (CN) and
- pGBKCg-pGADCg (CC)

Each screen was performed twice, resulting in a large number of new and overlapping interactions. Taken together, 87 of 96 test constructs showed a total of 569 individual interactions between full-length (FL) ORFs and/or domains (see Supplementary Table S4). 164 of these interactions were redundant in terms of

fusion tag permutation, which means that they were detected in two, three or four of the performed screens, e.g. NN, CN or NC, CN, CC. This leaves a total of 405 PPIs which are listed in the Supplementary Table S1. These 405 interactions still contain redundancies in certain ways. The interactions in this list are all different in terms of interacting constructs. For example, the two interactions between ORF1 and ORF60C (Interactions #6 and #7) are listed separately because they involve different constructs (full-length ORF1 in #6 and an N-terminal fragment of ORF1 in #7). Such redundancies are indicated by "red" in the column PPIs. All interacting protein pairs irrespective of constructs are listed in VZV_pairs. This column is redundant because the same protein pairs may occur multiple times (as in interactions #6 and #7). The PPI network of individual constructs comprising 405 interactions is shown in Figure 25.

To create a non-redundant PPI network, I had to exclude two more sources of redundancy. In the following step of data processing I eliminated redundancies caused by overlaps between different constructs of the same ORF. That means that I mapped each interaction that was detected in bait-prey direction to their full-length (FL)-ORF, for example the interactions #6 and #7 in Supplementary Table S1 (ORF1-ORF60C and ORF1-ORF60 interaction) were mapped to the same FL interaction (ORF1-ORF60). This led to a reduction of seven percent of the interactions from 405 to 377 PPIs. In the last step I removed identical protein pairs detected as bait-protein Y / prey-Protein X and bait-protein X / prey-Protein Y. This was the case in 7.7 % of the remaining 377 interactions leading to a non-redundant PPI set of 348 unique intraviral VZV protein interactions (Figure 26). That means that a two-fold increase compared to previously described PPIs was accomplished with the improved Y2H system (173 * 2.01 = 348). Summarized, PPIs of 67 of the 71 distinct VZV ORFs showed at least one interaction (94 %). This is an extremely high coverage, compared to the traditional screening method. In the initial NN-screens only 55 of the 71 ORFs interacted at least once as FL- or domain- bait or prey (77 %). That means that 17 % more ORFs were accessible by the Y2H system when steric hindrance effects can be overcome by changing the fusion-tag topology.

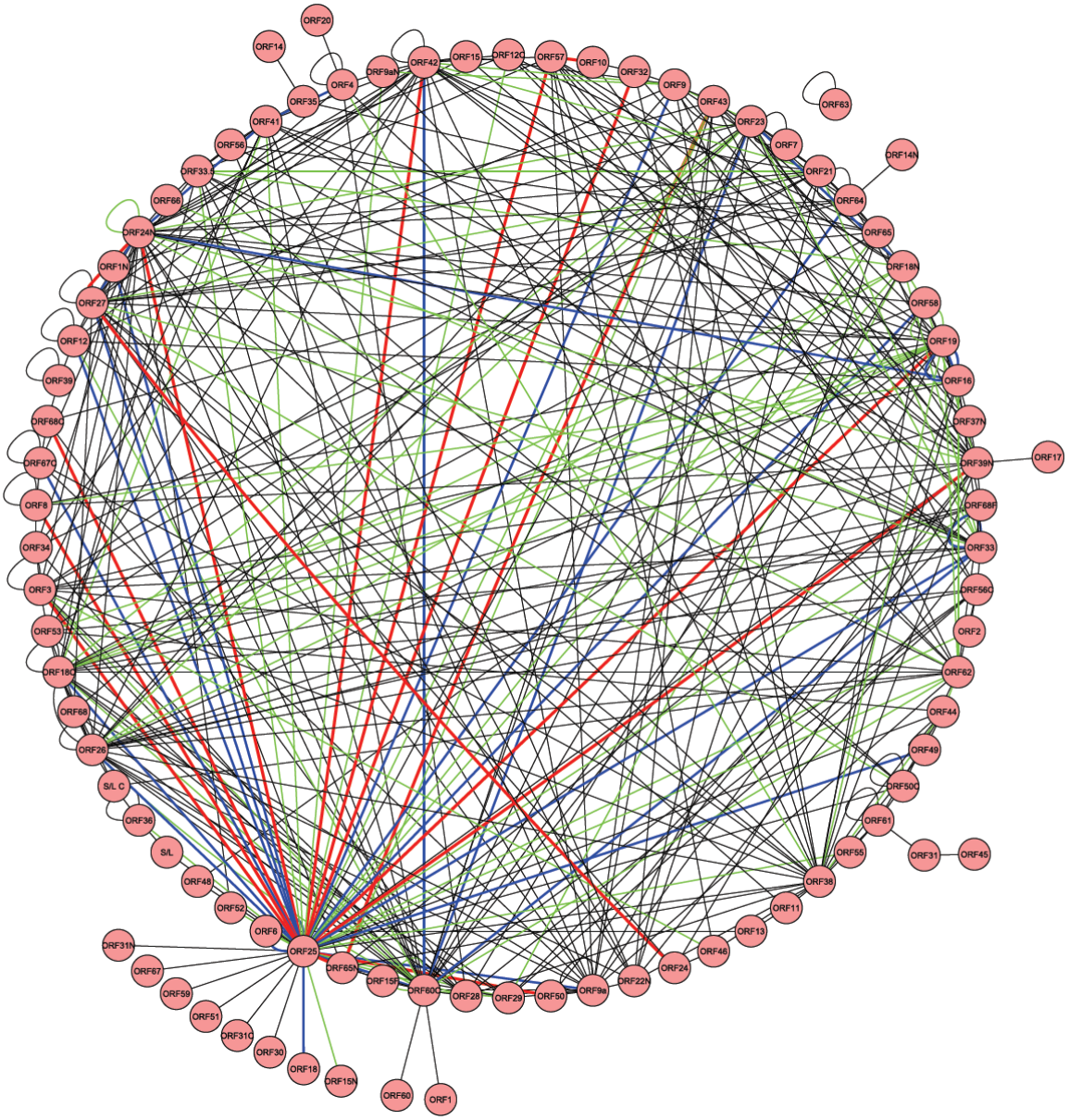


Figure 25: Combinatorial VZV PPI network
The PPI network comprises 405 protein-protein and protein-domain interactions. The individual distribution of tag-topology combinations an individual interaction was found with is excluded in this figure. The number of tag-topology combinations is indicated by the edge colour. **Black**: Interaction detected in one screen, e.g. NN; **Green**: detected in two screens; **Blue**: detected in three screens; **Red**: detected in all 4 screens. The network was generated using Cytoscape software environment v2.5.2 (Shannon et al., 2003).

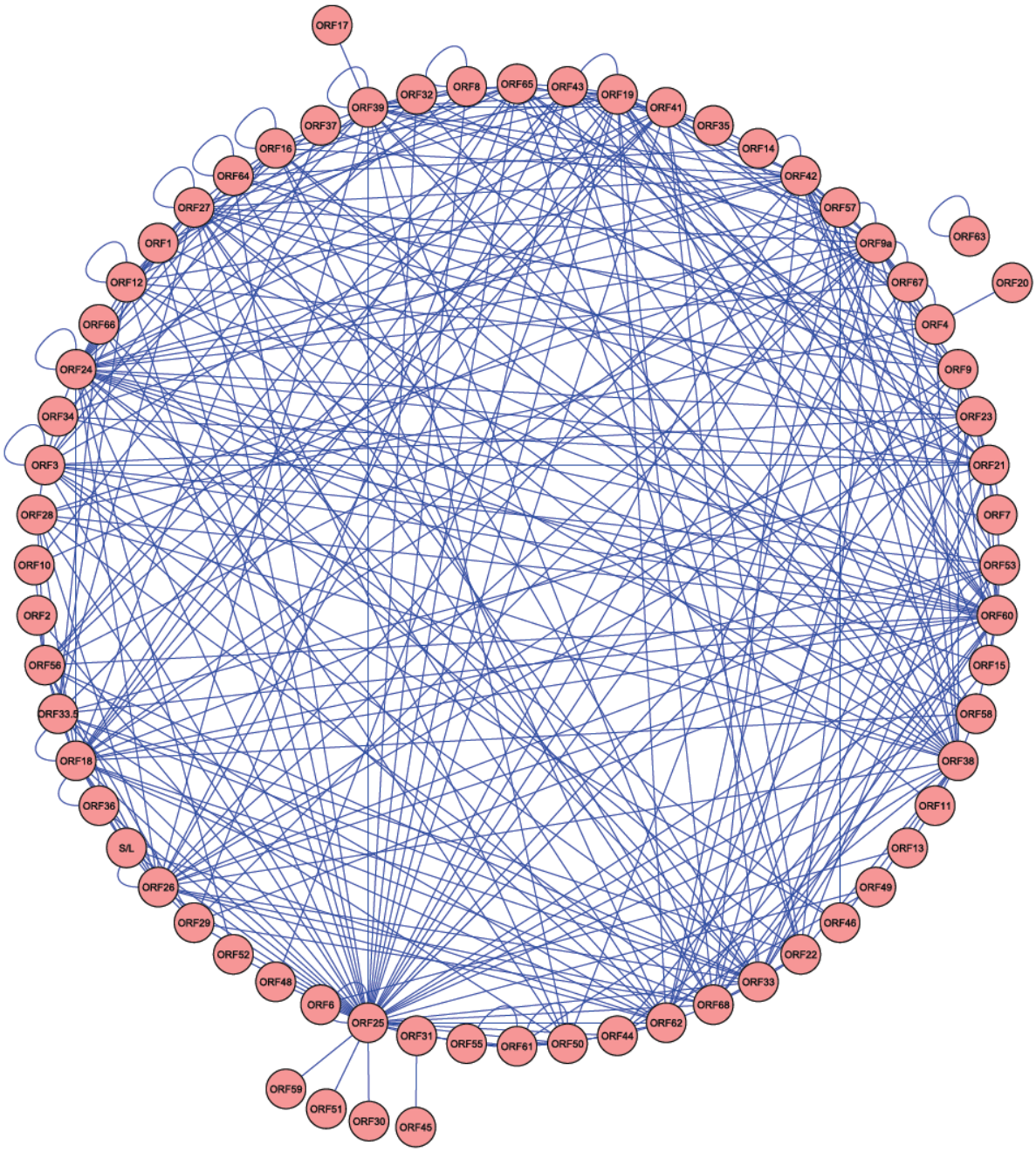


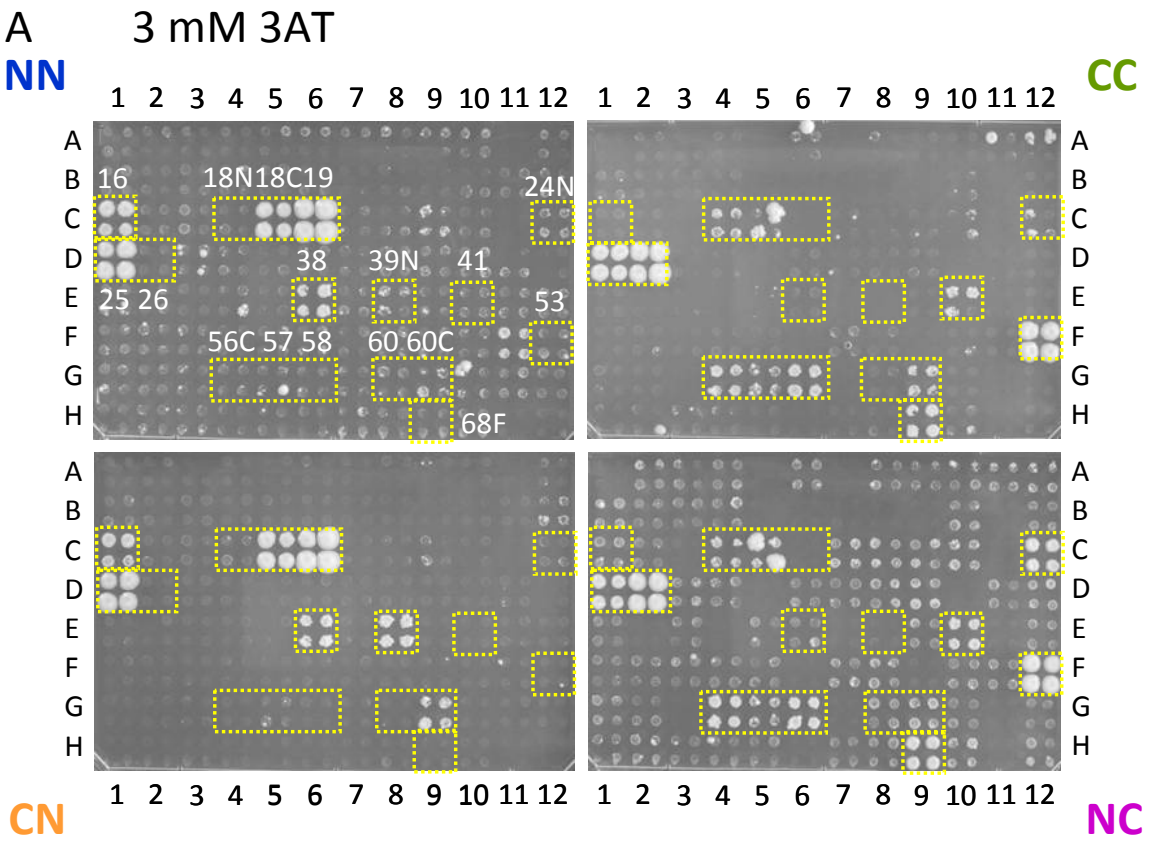
Figure 26: Non redundant VZV PPI network.
This PPI network comprises 348 identified physiological protein-protein interactions after removal of 29 (7.7 %) redundant interactions of the 377 FL network; Compared to the dataset taken as a basis of this study the number of PPIs was increased by two fold (Uetz et al., 2006).

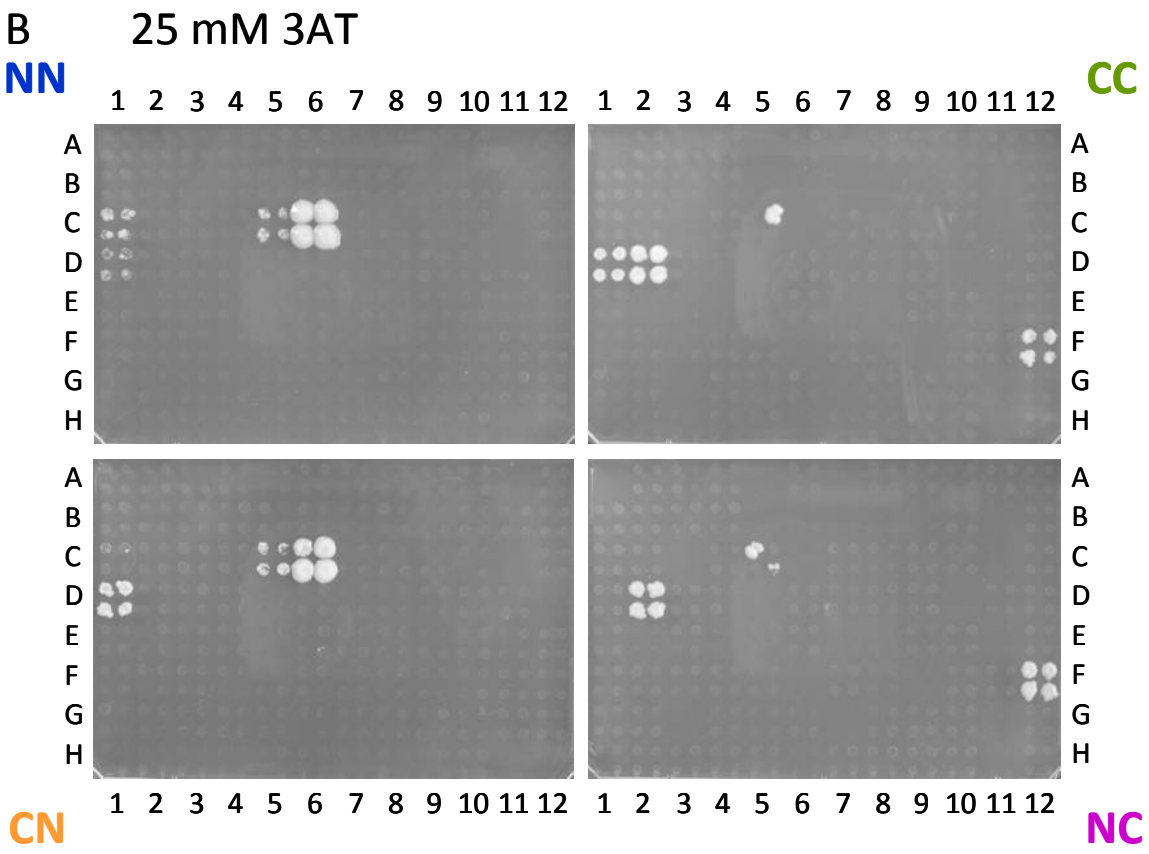
3.1.3.2 Interactions resulting from different vector combinations

Each of the four bait/prey vector combinations produced a significantly different interaction subset. For example, the screens with bait ORF19, the large subunit of ribonucleotide reductase, produced a total of 17 interactions (of 15 distinct proteins), of which only two were found in all four combinations (namely ORF25 and ORF18C), see Figure 27A. Five interactions were found with the N-terminal fusions (in pGBKT7g and pGADT7g) while 11 (3 strong + 8 weak) interactions resulted from the screens with the C-terminal fusions (in pGBKCg and pGADCg). The C-N and N-C combinations generated seven and eight interactions, respectively.

The interactions detected also depended strongly on the selection pressure: typically all screens were initially carried out without 3-aminotriazole (3AT), a competitive inhibitor of imidazoleglycerolphosphate dehydratase (Hilton et al., 1965), which is the His3 reporter enzyme used in our assays. If baits turned out to be autoactivating under these conditions, we raised the 3AT concentrations up to 50 mM in steps of 1, 3, 10, 25, 40 and 50 mM. ORF19 was an activator at 0, 1, and 3 mM and clear results were only obtained at 10 or 25 mM (Figure 27). Most interactions disappeared at 25 mM (Figure 27B). In general, the results reported here were obtained at a 3AT concentration that clearly differentiated between signal and noise. Note that ORF19 was a stronger activator as N-terminal DBD-fusion than as C-terminal fusion (background in Figure 27A). However, this was not generally true: in 20 cases the N-terminal bait fusion autoactivated while 21 of the C-terminal baits did so. In seven cases both fusions were autoactivators at 3 mM 3AT or higher concentrations but only one (ORF46) required more than 25 mM in both cases and was thus not interpretable (see Supplementary Table S4).

Figure 27: N- and C-terminal vectors detect different interactions.
Y2H screens of the four different vector combinations showing the differences on 3 mM (A) and 25 mM 3AT (B). The same bait, ORF19 (Uniprot accession P09248) was used as bait with N-terminally and C-terminally fused DNA-binding and activation domains and screened against a whole-genome array of Varicella Zoster Virus (VZV). The N-terminal bait and prey constructs (in pGBKT7g, pGADT7g, NN) show markedly different interaction patterns compared to the C-terminal constructs cloned into pGBKCg and pGADCg (CC). NC and CN combinations show yet different interactions. Preys are indicated by their ORF number, e.g. the bait ORF19 is the large ribonucleotide reductase (RNR) subunit which is known to interact with itself and the small RNR subunit (ORF18 = Uniprot P09247). Note that N and C labels near yeast colonies indicate N- and C-terminal protein fragments, not AD or DBD fusions (e.g. 18C and 18N are N- and C-terminal domains of ORF18). A complete list of interactions is provided in Supplementary Table S1. The sequences of all proteins are listed in Supplementary Table S5.


A
3 mM 3AT
NN
CC
16
18N18C19
24N
38
39N
41
25 26
53
56C 57 58
60 60C
68F
CN
NC


B
25 mM 3AT
NN
CC
CN
NC

3.1.3.3 Distribution of interactions identified in combinatorial screens

For the present study, I have conducted more than 27,000 individual Y2H tests, each performed two times to gain a systematic insight in the nature of the Yeast two-hybrid testing system which I have developed (96 × 96 pairwise combinations for three of four permutations, including multiple constructs of 18 of the 71 proteins) (see Supplementary Table S4). 112 interactions were exclusively found with the N-terminal combination. No combination turned out to be superior to the other three, although the number of N-C pairs was somewhat lower (90) than the NN, CN, and CC pairs (ranging from 146 to 182, see Figure 29). Absolute numbers lie within a normal distribution interval of $\mu = \pm 2\ \sigma$ indicating that no screen is an outlier. Overall, adding three additional combinations to any of the four increased the number of additional interactions by 2.2 to 4.5-fold (see Table 5).

combination	PPIs	other comb.	fold increase
NN	182	223	2.2
NC	90	315	4.5
CN	151	254	2.7
CC	146	259	2.8
∑	569		
mean	142.25		
σ	38.30		

Table 5: Number of PPIs found in individual screens.
Individual screened permutations and the associated interaction data. PPIs are the number of interactions found with this combination. Other combinations are all interactions not found with combination. For example, NN detected 182 PPIs, all other combinations found 223 in addition to those found by NN (405 PPIs without tag-topology redundancies). The "fold increase" is the ratio of other combinations over PPIs, i.e. the fold increase of interactions that are found when additional combinations are used in addition to the one indicated. PPIs + other combinations always add up to 405. ∑: Sum of overall detected interactions; **mean**: Arithmetic mean of detected interactions in individual screens; **σ**: Standard deviation.

If we look at the distribution of interactions in regard to the sheer number of screens in which the interactions were detected I observed that the majority was detected merely for one time. While only 64 of 569 redundant interactions were found in all four runs, a rising number was found in decreasing numbers of screens (see Table 6 and Figure 28).

# screens	# Interactions	% Interactions
1	298	52.37
2	132	23.20
3	75	13.18
4	64	11.25
∑	569	100.00

Table 6: Distribution of the number of times individual PPIs were found.
In this table the number of screened permutations (1-4) and the respective numbers of interactions and the corresponding fraction are listed. The amount of interactions decreases with rising overlaps between distinct screens.

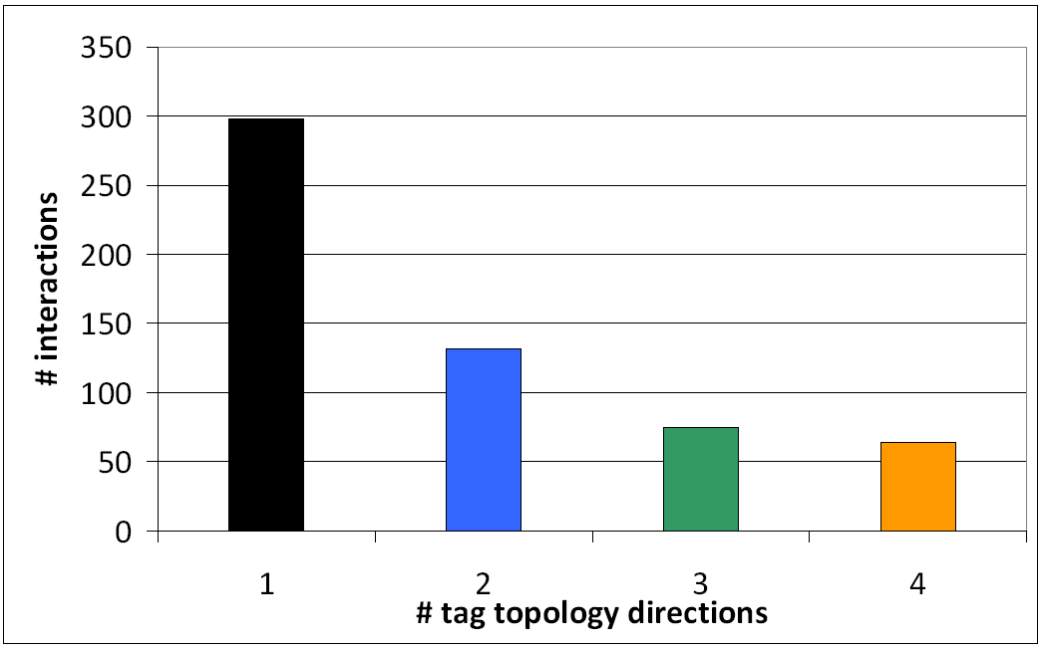


Figure 28: Distribution of interactions detected one to four times.
This graph shows the distribution of interacting constructs detected one, two, three or four times irrespective of the nature of the individual combination. **4** = NN, NC, CN and CC; **3** = e.g. NN, CN, CC or CN, NC, CC et cetera. The sum of interactions is 569.

3.1.3.4 Overlap between vector combinations

As shown in Figure 29, increasing numbers of vector combinations detected decreasing numbers of overlapping results. This was generally true: while the N-terminal fusions produced a total of 182 interactions, only 16 of them were also found with the other three combinations.

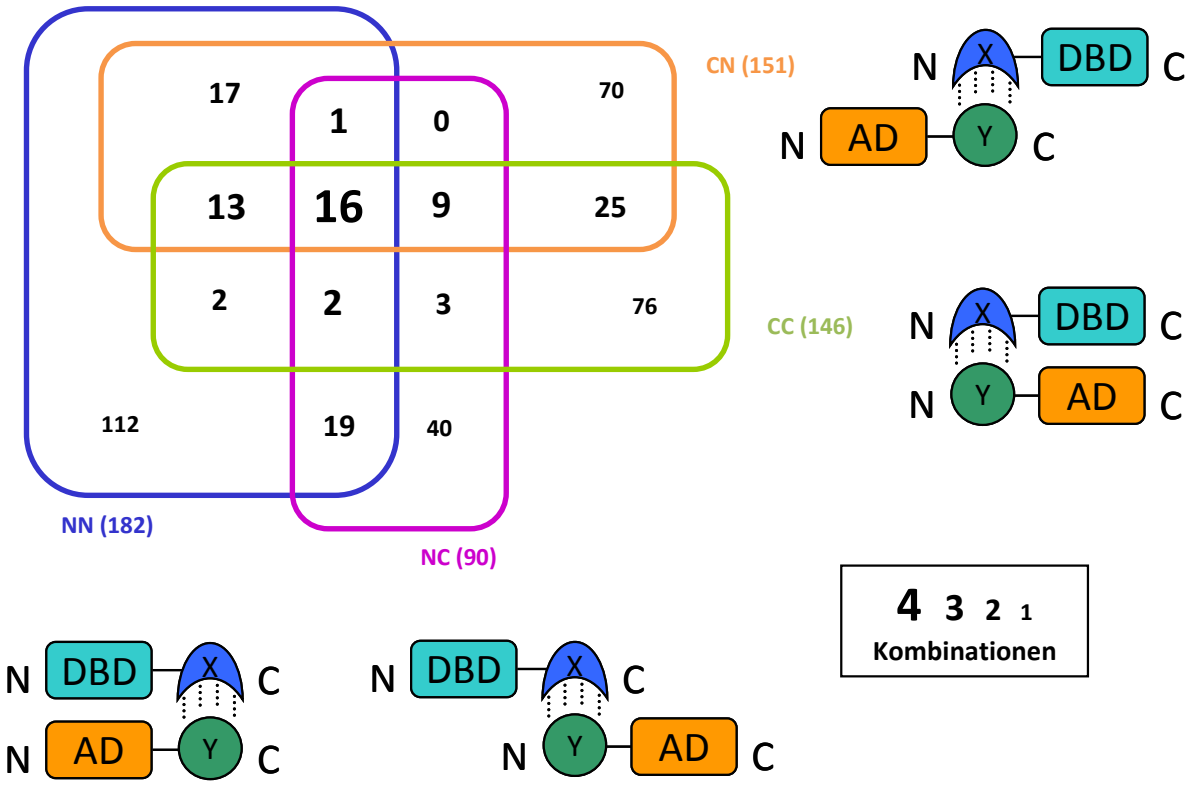


Figure 29: Overlaps between tag-topology combinations.
This Venn diagram shows the distribution of the 405 VZV interactions found in either one or more tag-topologies. 182 interactions were detected in the traditional screen (NN); 90, 151 and 146 PPIs were additionally identified in the new combinatorial screens (NC, CN and CC respectively). The differences in the tag topologies are sketched aside of the corresponding topology abbreviations. The total number of interactions in each screen is enfolded by a box in the respective colour. The overlaps between different vector systems are specified in Table 7. The number of overlaps is indicated by the thickness of numbers decreasing from four to one.

	NN	NC	CN	CC
NN	182	38 (21/42 %)	47 (26/31 %)	33 (18/23 %)
NC	—	90	26 (29/17 %)	30 (33/21 %)
CN	—	—	151	63 (42/43 %)
CC	—	—	—	146

Table 7: Overlap between screens.
Identical interactions found with multiple screens of different combinations. For example, the NN and NC screens found 182 and 90 interactions, respectively, of which 38 were identical (i.e. the protein pairs were identical), corresponding to 21 % of NN- and 42 % of NC- interactions.

3.2 Quality assessment of PPI-data

In the next chapter I will investigate the quality of the Y2H interaction data. This is a very crucial step to allow an estimation of the quality of the present dataset as well as for future users willing to use this technique. The two main questions that have to be addressed are the degree of reduction of false-negative and false-positive interactions.

3.2.1 Improvement of the assay sensitivity

The assay sensitivity of Y2H screens (fraction of all biophysical interactions that can be detected by a given assay), has been estimated to cover only 17 % of all actual PPIs taking place in one certain test system (Venkatesan et al., 2009). This low coverage is a major stumbling block in the attempt to gain a complete dataset which is necessary e.g. for systems biological attempts towards the understanding of pathomechanisms (Dreze et al., 2009).
In regard to the present testing system I was primarily interested in how many interactions can be verified by any additional evidence. Therefore I created a set of high-confidence interactions which meet at least one of a set of certain predefined demands, the Gold Standard interactions.

3.2.1.1 Gold Standard interactions

To gather a preferably large dataset of Gold Standard interactions I used three different parameters. The first one was to compare the Y2H interaction data to published PPI-data, so called literature-curated (LC) interactions. Second, I predicted VZV-interactions from four large scale screens of related herpesviruses (HSV-1, MCMV, EBV and KSHV), which were previously performed in our group (Fossum et al., 2009). At last, I took those interactions into account which appeared more than one time within the VZV screens, represented by different constructs of the same protein pair. Comparison of the literature curated interactions and Y2H interologs with the VZV interaction data from the combinatorial screening was performed with the help of databases generated with “FileMakerPro”, v.9.5 (http://www.filemaker.de).

3.2.1.1.1 Literature-curated interactions

From a recent publication I could make use of an overview about all published protein interaction data of VZV to date (Fossum et al., 2009). This literature-curated dataset is made up of various small-scale and large-scale studies and was consequently used as gold-standard set. Since not many VZV interactions have been published in small-scale studies so far I included small- and large scale derived interactions from four more members of the *Herpesviridae* family, viz. HSV-1, MCMV, EBV and KSHV (Rozen et al., 2008) and checked the VZV interactions also for interologous interactions thereof (see Supplementary Table S2).

3.2.1.1.2 Orthologous prediction of LC VZV interactions

Fossum et al. listed 91 literature-curated interactions from five different herpesviruses including nine in VZV. 67 of the 91 interacting pairs had orthologs in VZV (Davison, 2004) and could therefore be used for orthologous prediction. A List of orthologous proteins in VZV, HSV-1, MCMV, EBV and KSHV can be found in the supplementary information of Fossum et al., 2009.

3.2.1.1.3 Orthologous prediction of Y2H interactions

In the same way as described in the previous passage, I compared published Y2H interactions from large scale studies (Fossum et al., 2009; Uetz et al., 2006) to the 405 VZV-interactions found in this study. Altogether, 60 interactions could be found in this way. The individual ratio is constantly rising from 11.1 % within interactions found in one vector combination to 43.8 % found in all four combinations. Altogether the number of Y2H verified interactions is corresponding to 14.8 % of the whole dataset (see Table 8).

3.2.1.1.4 Redundancies within the screen

The third indication for high confidence interactions is the appearance of redundant interactions within the dataset, e.g. identical protein pairs are represented by different constructs by an interaction of full-length proteins reproduced by a full-length protein and a protein fragment. See for example interaction number 18 and 19 in Supplementary Table S1. Or an interaction is found in the combinations bait-protein X/ prey-protein Y and bait-protein X/ prey-protein Y, e.g. the interactions #147 and #148.

3.2.1.2 Verification of the dataset using Gold Standard interactions

The raw data was systematically screened for Gold Standard interactions with the help of a database created with FileMaker version 8.5 (FileMaker GmbH, Unterschleißheim). Taken together 47 % of all interactions could be verified by at least one quality feature (41 % of the redundant dataset without tag-topology redundancies and 32 % of the non-redundant dataset). The individual distribution of gold-standard interactions is listed in Table 8 and shown in Figure 30.

# combinations	1	2	3	4	∑
verified PPIs	76 (107)	17 (35; 70)	9 (14; 42)	9 (12; 48)	111 (168; 267)
all PPIs	267 (298)	48 (66; 132)	20 (25; 75)	13 (16; 64)	348 (405; 569)
% verified	28 % (36 %)	35 % (53 %)	45 % (56 %)	69 % (75 %)	32 % (41 %; 47 %)

Table 8: Verification by additional evidence.
Verifications include literature-curated interactions, large-scale Y2H assays, or interactions found with alternative constructs. For example, there are 13 unique interactions that have been found in all 4 combinations (NN, NC, CN, CC) (redundant interactions are those that are found among the same proteins but using different constructs, e.g. between ORF24 and ORF27 as well as between ORF24N and ORF27). Redundant interactions are labeled "red" in column PPIs in Supplementary Table S1. Verified interactions have been verified by (1) an elsewhere published interaction (indicated by the PubMed ID in Supplementary Table S1), (2) a large-scale Y2H experiment (Fossum et al., 2009; Uetz et al., 2006), or by an independent construct (e.g. a fragment of a protein in addition to a full-length protein). Numbers apply to non-redundant interactions (348 PPIs), numbers in brackets refer to the redundant dataset (405 PPIs w/o-; 569 PPIs with tag-topology overlaps).

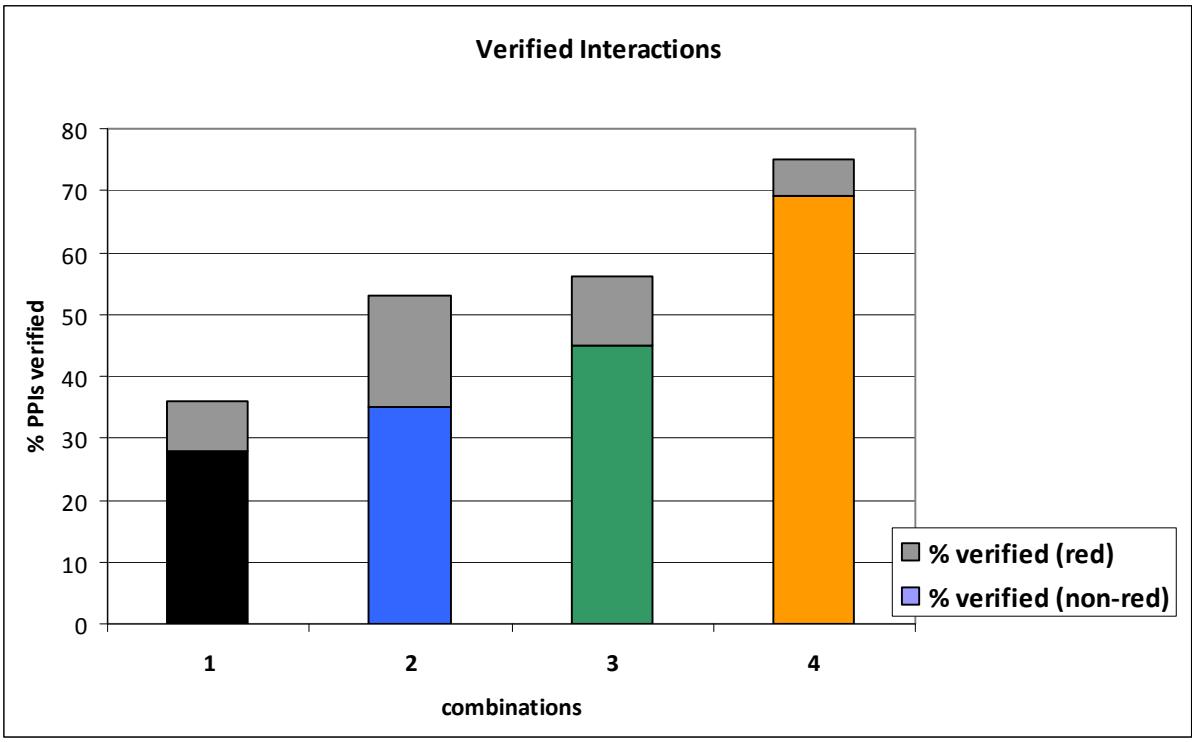


Figure 30: Distribution of verified VZV interactions.
Fraction of gold-standard verified interactions in subject to the number of tag-permutations they were detected with. Values refer to Table 8; the percentage of verified interactions referring to non-redundant interactions (non-red) is topped by the value within the redundant dataset (red).

3.2.1.3 Is the data-quality between tag-permutations equal?

A crucial question is how good the quality of interactions is distributed when different vector combinations are used. Comparison of the number of Gold-Standard verified interactions to the overall interaction number of the respective screen revealed ratios between 45 % and 50 % for any of the four independent screens (see Table 9).

	NN	NC	CN	CC
verified	83	45	71	68
total	182	90	151	146
% verified	45 %	50 %	47 %	47 %
mean	47.25 %			
σ	2.06 %			

Table 9: All permutations generate data of equal quality.
The interactions found for each permutation were checked for their fraction of verified interactions as in the previous chapter. No permutation seems to be superior to any other permutation. Note that the percentages are larger than in Table 8 because redundant interactions were counted as well. Such redundant interactions are generated by different constructs of the same protein all of which may confirm an interaction of another permutation. **σ**: Standard deviation.

3.2.2 Conservation of PPIs between orthologous proteins

It has been shown that the conservation of proteins positively correlates with the conservation of protein interactions (Fox et al., 2009). In this paragraph I want to investigate if the VZV interaction data supports this observation.

To investigate which conserved PPIs have been found in which screen and for how many times, I extracted literature-curated and Y2H interologous interactions from the raw interaction data (Supplementary Table S1) and apportioned the remaining interactions. Altogether, 74 interactions are in the VZV dataset which were previously reported in at least one more herpesvirus and are therefore called interologous interactions. The overlaps of interologous interactions between the Y2H and LC dataset with VZV interactions detected in this study are graphed in Figure 31. 15 of these interactions imply one core-ortholog, five take place between non-core proteins, while 54 take place between two core-orthologs. This indicates a high degree of interaction conservation between closely related species. On the other hand I found 136 interactions between 38 of 41 core-proteins which are conserved in HSV1, MCMV, EBV, or KSHV as well (see Figure 32). Subsequently, 212 interactions take place between one core-ortholog and a non-core protein or between two non-core proteins. Putting these numbers together, 54 of 136 (39.7 %) core-interactions are conserved between five herpesvirus species while only 20 of 212 (9.4 %) are conserved between less conserved proteins. This meets with a 4.2-fold increase of interaction conservation between conserved proteins in comparison to non-conserved interaction partners. Compared to the whole dataset (348 PPIs), 21.2 % of all VZV interactions are conserved among HV-species, 15.5 % are conserved within core-orthologs and 5.7 % outside of the core interactions. For detailed numbers see Table 10.

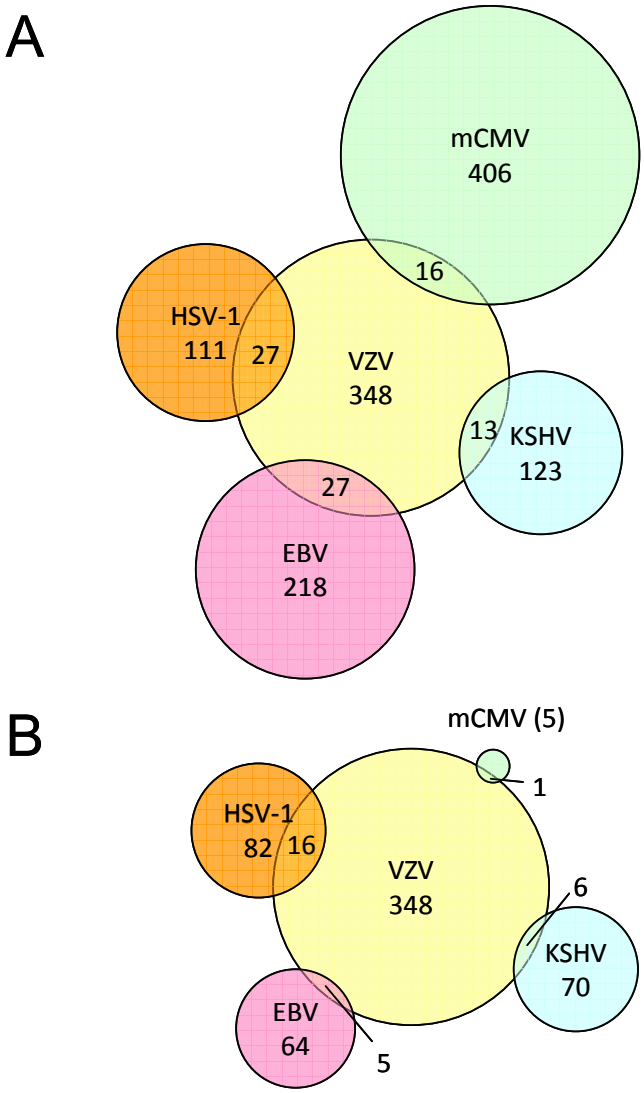


Figure 31: Overlaps between Y2H-data and VZV PPI network.
This Venn diagrams show the number of PPIs detected in **(A)** large-scale Y2H-screens (Fossum et al., 2009; Stellberger et al., 2010; Uetz et al., 2006) and in **(B)** extensive investigation of small-scale studies which make up the literature-curated dataset (Fossum et al., 2009) and the interologous overlaps with the VZV-network. In aid of clarity overlaps between HSV-1, MCMV, KSHV and EBV were omitted.

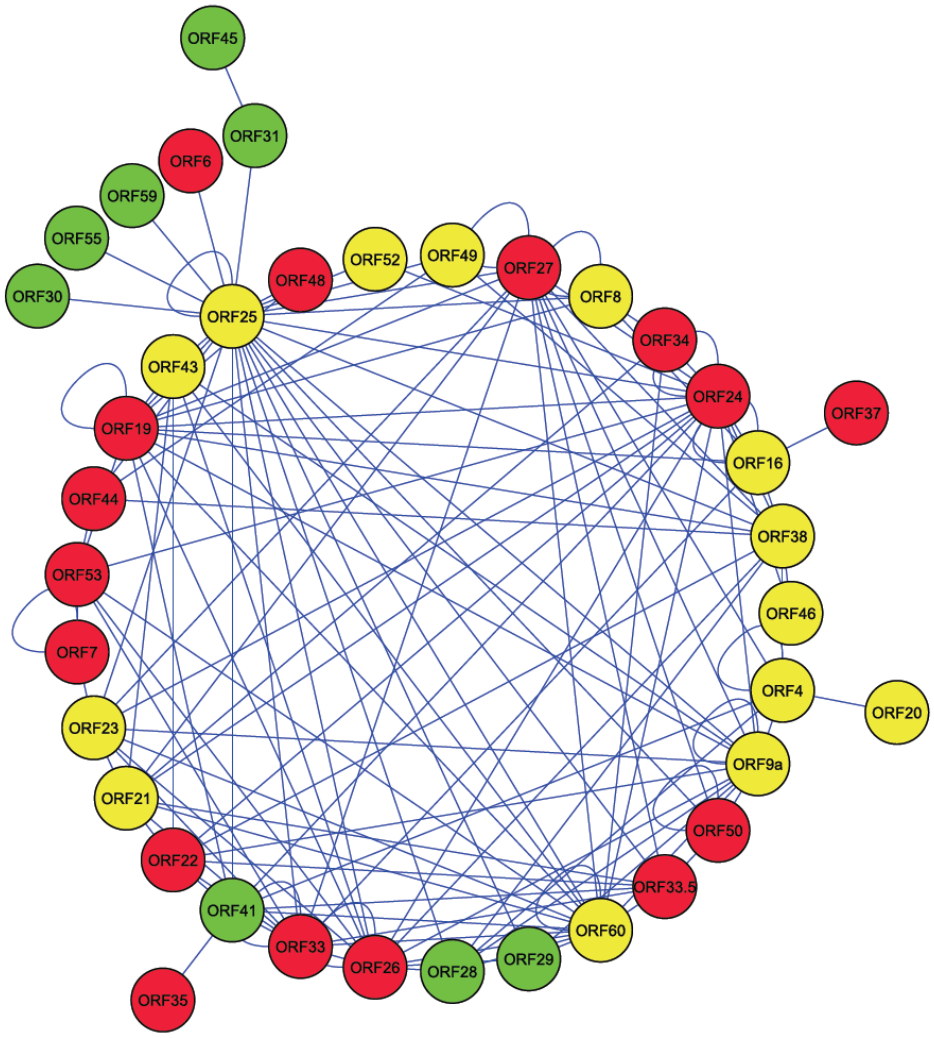


Figure 32: VZV interactions among core-proteins.
The core-network comprises 136 binary protein-protein interactions between 38 of 41 conserved herpesviral proteins. Colors indicate the degree of sequence conservation between the orthologs in VZV, HSV1, MCMV, EBV, or KSHV. Yellow: Sequence conservation < 30 %; Red: < 40 %; Green: > 40 %.

3.2.3 Introducing an intrinsic quality score

The existence of false-positive interactions within a PPI-network is almost inevitable and it is very difficult to prove the existence of false-positives, if this can be done at all. In this study, proteins are investigated which are expressed in the same cell and one can not be 100 % sure that interactions detected in Y2H assays do not take place in a human host cell. In this chapter I provided circumstantial evidence that interactions found in multiple combinations are of higher quality than interactions found in a single screen. Thus, I can allege that interactions found with multiple protein fragments support this interaction (as long as the fragments are not exclusive as in non-overlapping N- and C-terminal fragments).

combinations	1	2	3	4	∑
LC interologs (core)	11	0	1	1	13
Y2H interologs (core)	26	11	6	5	48
verified core PPIs (NR)	31	11	7	5	54
PPIs total (NR)	267	48	20	13	348
% verified core (136)	22.8 %	8.1 %	5.1 %	3.7 %	**39.7 %**
% verified core total (348)	11.6 %	22.9 %	35.0 %	38.5 %	**15.5 %**
LC interologs (non-core)	9	1	0	0	10
Y2H interologs (non-core)	8	0	1	2	11
verified non-core PPIs (NR)	16	1	1	2	20
% verified non-core (212)	7.5 %	0.5 %	0.5 %	0.9 %	**9.4 %**
% verified non-core total (348)	6.0 %	2.1 %	5.0 %	15.4 %	**5.7 %**
% verified core and non-core	17.6 %	25.0 %	40.0 %	53.9 %	**21.2 %**

Table 10: Interologous interactions among core and non-core proteins.
This table summarizes the fractions of interologous interactions found by one to four fusion tag permutations. Interologs are divided in core (upper half of the table) and non-core interactions (lower half). Additionally each group is subdivided into literature-curated interologs (LC interologs) and those identified in large-scale Y2H screens (Y2H-interologs). As there are overlaps between those two subgroups the number of verified PPIs is lower than the sum of both. This is caused by interactions verified by literature and Y2H data from distinct HV-species, e.g. interaction #117 in supplementary Table S1. Percentual fractions of verified core and non-core interactions are calculated in relation to the number of non-redundant (NR) intra-core (136) or intra non-core (212) interactions. Total numbers correspond to the fraction of interactions found one to four times in the whole non-redundant dataset, e.g. combination = 1; # verified core PPIs (NR) = 31; PPIs total (NR) = 267; consequential the ratio of verified interaction in total is 31/267 = 11.6 %.

Based on these findings I defined three quality scores, high-, medium- and basic quality interactions. Quality scores were defined as:

- "3" (high) when found in three or four different screens and/ or verified by literature curation
- "2" (medium) when found one or two times in the own Y2H screens and having either an interologous Y2H interaction in another herpesvirus (Fossum et al., 2009; Uetz et al., 2006) and/ or being verified by redundancy within the permutated screens, and
- "1" (basic) quality interactions detected in one or two permutations and not further confirmed.

The distribution of the quality scores seems intuitively to make sense as the fractions almost double from the higher quality score to the next lower one (see Table 11).

quality score	PPI attributes	# interactions	fold increase
3 (high)	perm_count = 3 or 4 and/or LC verified	65	-
2 (medium)	perm_count 1 or 2 + Y2H-Interolog and/or redundancy	118	1.82
1 (basic)	perm_count 1 or 2	222	1.88
∑		405	

Table 11: Quality scores assigned based on data verification.
Distribution of interactions within the three quality characteristics defined. perm_count: Number of permutations this protein pair was found in, e.g. 2 when found as NN and CC; Y2H-Interolog: Interologs from previous large-scale Y2H screens by our group (Uetz et al. 2006 and Fossum et al. 2009); LC verified: Published interaction or interologs of published interactions; redundancy: Interactions detected in this study by independent constructs of the same protein pair.

3.2.4 Network analysis

In this chapter I want to compare the three most robust network parameters of the traditionally screened VZV interactome (N-terminally fused bait and prey tags) and the one derived by permutated screening. Those measures of network topology are degree distribution, clustering coefficient and the average path length.

Network parameters were calculated using the "NetworkAnalyzer" plugin (Assenov et al., 2008) for Cytoscape network visualization software (Shannon et al., 2003). Protein interaction networks are commonly treated as undirected networks meaning that an edge (protein interaction) between two nodes (proteins) describes an alternating binding relationship without predetermined direction. For example a protein X interacts with protein Y and vice versa (Barabasi and Oltvai, 2004). It would be possible to introduce a direction by concerning the bait-prey direction of interactions but this is without relevance to the biological system which is investigated and is therefore omitted.

3.2.4.1 Node degree distribution and average degree

The degree of a node, abbreviated *k*, tells how many links a node has to other nodes. In undirected networks, the node degree of a node *n* is the number of edges linked to *n*. A self-loop of a node is counted like two edges for the node degree (Diestel, 2005). An undirected network with *N* nodes (proteins) and *L* edges (links, interactions) is characterized by an average degree $\langle k \rangle = 2*L/N$ while self-interactions are subtracted.

The average degree of the initial network is $\langle k \rangle = 2*(171-11)/55 = 5.6$.

After introduction of C-terminal tagged test constructs <*k*> = 2*(348-23)/67 = 9.7.
This means that each protein has in average 1.7 times more interactions.
The node degree distribution gives the number of nodes with degree *k* for *k* = 0, 1… and is used to distinguish between random and scale-free networks. Both networks degree distributions are plotted in Figure 33, see Supplementary Table S7 for the associated data. As reported for cellular protein interaction networks (Bonifazi et al., 2009), the VZV networks appear to be scale-free. Scale-free networks are complex networks which have no characteristic number of edges per node. The node-degree distributions follow power law functions of the form $y = a*x^k$ implicated in the respective graphs of Figure 33. Comparison of the Coefficients of Determination (R^2) which are a measure of agreement between the observed values and the modeled power-law values shows a higher agreement for the extended VZV protein interaction dataset. R^2 rises from 0.516 to 0.619 in the permutated screening for 0 = no correlation and 1 = perfect fit between the observed values and the computed regression line.

3.2.4.2 Network clustering coefficient

The clustering coefficient of a node is the number of triangles (3-loops) that pass through this node, relative to the maximum number of 3-loops that could pass through the node. In undirected networks, the clustering coefficient Cn of a node n is defined as Cn = 2en/(kn(kn-1)), where kn is the number of neighbors of n and en is the number of connected pairs between all neighbors of n (Barabasi and Oltvai, 2004; Watts and Strogatz, 1998). The clustering coefficient is a ratio N / M, where N is the number of edges between the neighbors of n, and M is the maximum number of edges that could possibly exist between the neighbors of n. The clustering coefficient of a node is always a number between 0 and 1.
The average clustering coefficient distribution gives the average of the clustering coefficients for all nodes n with k neighbors for k = 2 …

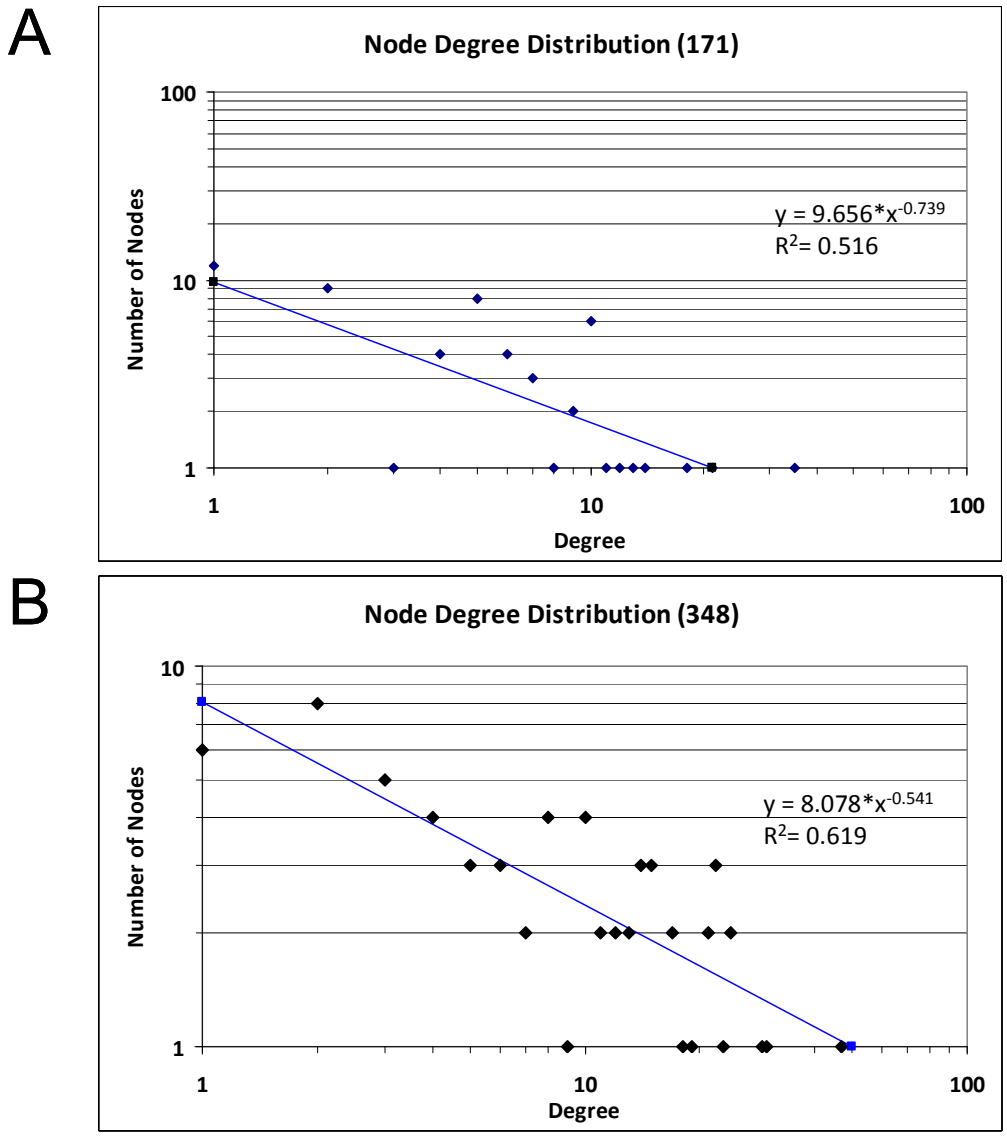


Figure 33: VZV node degree distribution of the primary and the extended VZV-network.
Node degree distributions indicate power-law distribution of the protein-interactions. The initial dataset comprising 171 protein interactions (A) has an average node degree <*k*> of 5.6 and fits the calculated regression with a Coefficient of Determination (R^2) of 0.516. For the extended interaction network <*k*> = 9.7 and R^2 = 0.619. The regressions following power-law distributions are marked on the respective plot.

The network clustering coefficient is the average of the clustering coefficients for all nodes in the network. Here, nodes with less than two neighbors are assumed to have a clustering coefficient of 0.

The network clustering coefficient of the 348 PPI network is 0.311, the one of the 171 PPI network is 0.236. That means that each node has 1.32 times more clusters (triangular relationships between proteins) after introduction of C-terminally tagged test constructs.

3.2.4.3 Characteristic path length and attack tolerance of VZV networks.

The characteristic path length (CPL) is for example the number of clicks which will lead you from one website to another, in this case the number of interactions leading from one protein to another on an average within the whole network (average distance between any two proteins). The characteristic path length distinguishes an easily acessible network from one which is complicated and inefficient, with a shorter CPL being more desirable, for example when it comes to transfer of information or signalling in a biological system. The network itself might have some very dislodged connected nodes and many nodes which are neighbors of each other, and the CPL describes what the path length will most likely be.

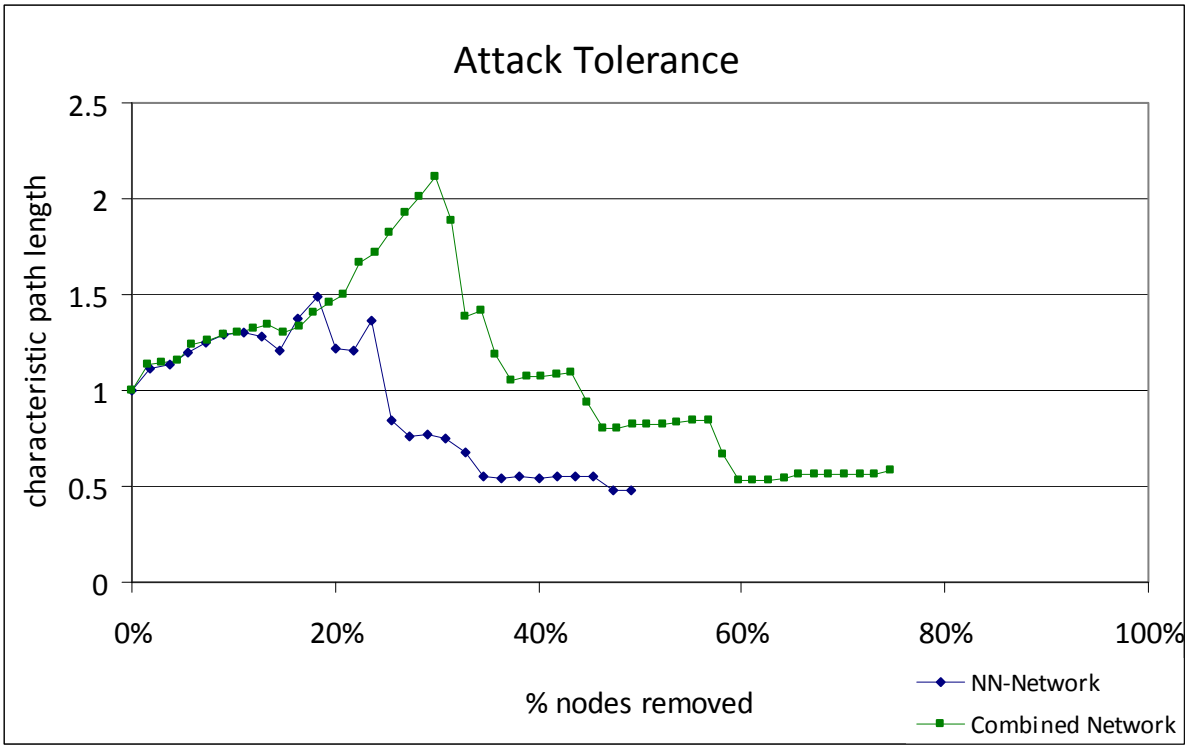


Figure 34: Attack Tolerance of the primary and the extended VZV-network.
The VZV networks were attacked by step-by-step removal of the most connected nodes (hubs). Characteristic path length was recalculated after each step and put in relation to the initial value which was set as 1. Deterioration of the networks organization reflects in an increase of the characteristic path length. Removal of 18 % of the most highly connected hub proteins (node degree 9 for the NN- and 18 for the extended network) makes no difference (t-value = 0.44) while removal of 19 % to 50 % of the more complex network significantly increases the attack tolerance (t-value = $1.3*10^{-6}$).

The robustness of the networks is investigated by a step-by-step assault strategy and subsequent recalculation of the characteristic path length. The most highly connected nodes are removed in decreasing order. After each node is removed, the new network CPL of the remaining network is plotted as a multiple or fraction of the

original parameters (Figure 34 & Supplementary Table S6). A higher attack tolerance is ascertained when the increase in path length is considerably smaller. The more complex network of 348 intraviral VZV interaction shows a lower attack tolerance after removal of more than 18 % of the highest connected proteins. Those so called hub-proteins share the same alteration of the characteristic path length (t-value = 0.44) compared to the NN-network, while removal of less connected nodes induces a highly significant increase of the CPL in the combined network (t-value = $1.3*10^{-6}$***).

3.2.5 Combinatorial screening of human reference sets

In order to be able to make estimations about the assay sensitivity and the rate of false-positives, we tested human reference sets of binary protein-protein interactions into the Yeast two-hybrid vectors used for the newly developed combinatorial screening system. Both a reference set of positive interactions and very unlikely interactions were tested. The positive reference set (hsPRS-v1; PRS) and the random reference set (hsRRS-v1; RRS) were kindly provided as entry clone collections by Pascal Braun (Braun et al., 2009). This project was performed in collaboration with Yu-Chi Chen and Seesandra Venkatappa Rajagopala, PhD, two members of our split research group at the J. Craig Venter Institute in Rockville, Maryland, USA. The work has been published recently (Chen et al., 2010).

The reference set ORFs were recombinated from the entry vectors by LR-reaction into the prey vectors pGADT7g and pGADCg as well as into the bait vectors pGBGT7g and pGBKCg (see Figure 16). The pGADCg and pGBKCg test constructs were generated by myself, while the pGADT7g and pGBKT7g- constructs, the yeast test-strains and subsequent screening was performed by Yu-Chi Chen under supervision of Seesandra Venkatappa Rajagopala, PhD.

Both reference sets consist of 92 protein pairs. That means, for each set two times 92 entry clones were recombinated into four destination vectors. Altogether, 736 LR-recombination reactions were necessary for the PRS and again the same number for the RRS. Bait vectors were transformed into the reporter yeast strain AH109 (a) and the prey vectors into the complementary strain Y187 (α). For a detailed list of the hsPRS-v1 and hsRRS-v1 see Supplementary Table S8. In addition to each vector pair, each protein was tested both as activation (prey) and DNA binding domain fusion (bait), as well as C- and N-terminal fusions. This way, each protein pair was tested in eight different configurations.

3.2.5.1 Estimations on the assay sensitivity and false-positive rate

The resulting bait and prey arrays were mated against each other and the resulting diploid test strains were grown on -LTH-agar plates for seven days with rising stringency (0, 1, 3, 10, 25 mM 3-AT) (Yu-Chi Chen, Seesandra Venkatappa Rajagopala). The interactions detected in the single tag-topology combinations and taken together reflect the assay sensitivity of the screening method applied to the hsPRS-v1 set. Furthermore, it can give evidence on the general assay sensitivity that can be obtained by Y2H screening with permutated tag topology combinations.

3.2.5.1.1 hsPRS-v1 interactions

The 92 hsPRS-v1 interactions were tested pairwise in combinatorial screens of rising stringency (Yu-Chi Chen, Seesandra Venkatappa Rajagopala). Every assay was accompanied by a control assay with the respective bait-array mated against the empty prey vector to monitor autoactivation of the bait constructs. From 92 possible interactions, 74 could be detected in at least one screen representing an assay sensitivity of 80.4 %. The results are visualized in detail as an interaction matrix in Figure 35. The distribution of interactions referring to the fusion-tag orientations is showed in Table 12. To assess, if the interaction-distribution between the human positive reference set and the VZV interactome are comparable, I plotted the fusion-tag distributions of either screens in parallel and could notice that there is probably no significant difference (see Figure 36A). In order to make a statistical comparison, I calculated the relative screen performance of both datasets. The number of total non-redundant interactions was normalized and the different tag-topology combinations were calculated to their proportion of the total value (see Figure 36B). Relative values of NN, NC, CN and CC-detected interactions (listed in Table 13) were pairwise compared by a two-tailed t-test for heteroskedastic samples. The resulting p-value is 0.11, which means that there is no significant difference between the tag-topology distributions of both datasets.

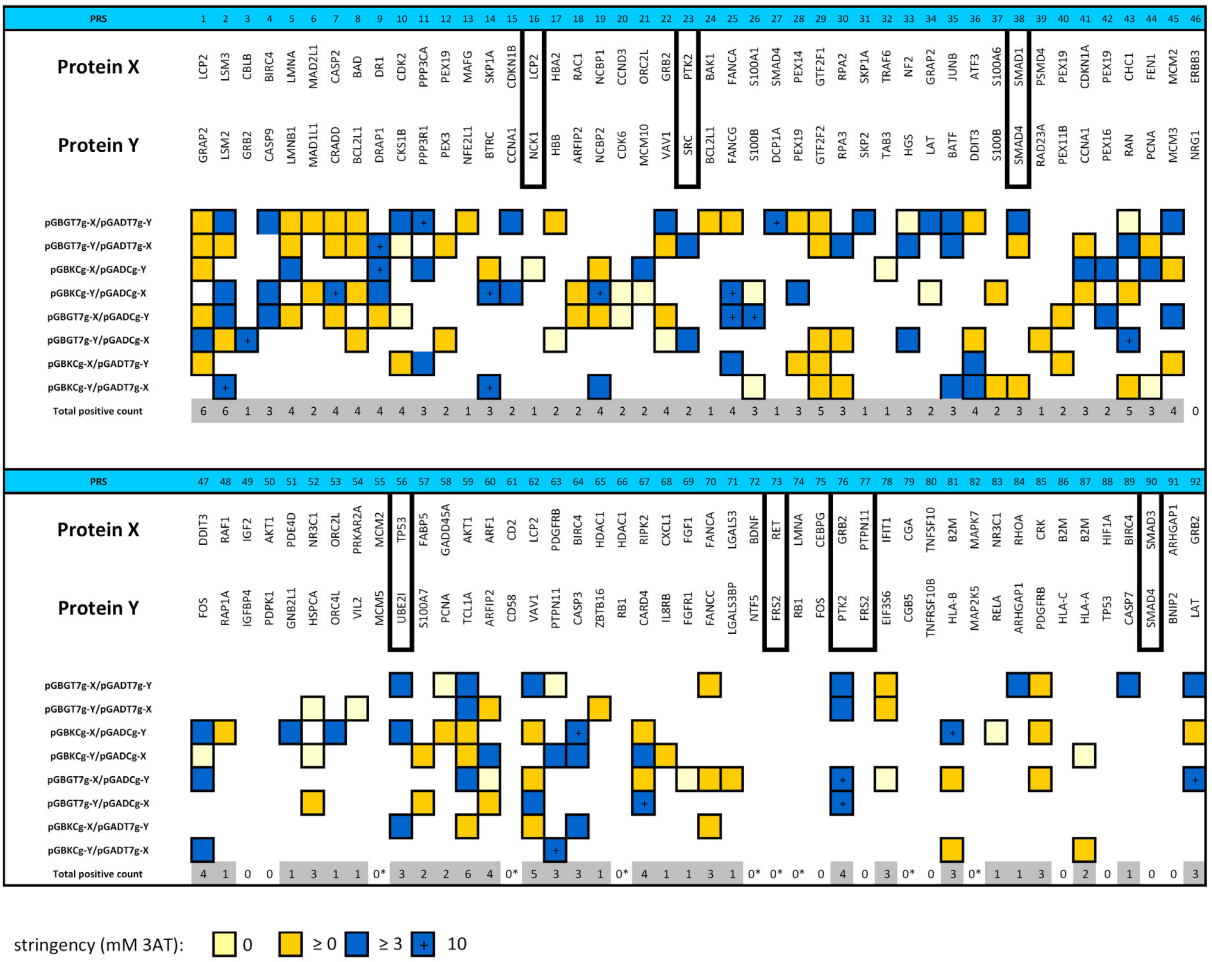


Figure 35: hsPRS-v1 interactions reproduced in combinatorial screens.
Matrix of interacting protein pairs and the bait-prey vector combinations they were assayed with. Interactions are marked by colored boxes according to the highest stringency they were still detectable (see legend). Interacting pairs in boxes indicate interactions which are depending on phosphorylation, e.g. interactions #16 and #23. A subset of those post-translational modification-dependent interactions can be detected in Y2H assays (5 of 8 interactions). Adapted from: Chen et al., 2010.

tag permutation	# interactions detected	assay sensitivity
NN	48 of 92	52.17 %
CC	47 of 92	45.65 %
NC	42 of 92	31.52 %
CN	29 of 92	51.09 %
combined	74 of 92	80.44 %

Table 12: hsPRS-v1 interactions broken down to the single tag permutations.
Number of Interactions detected in the single tag permutation combinations (data from Figure 35) and the respective assay sensitivity regarding the hsPRSv1 set of pairwise interacting proteins.

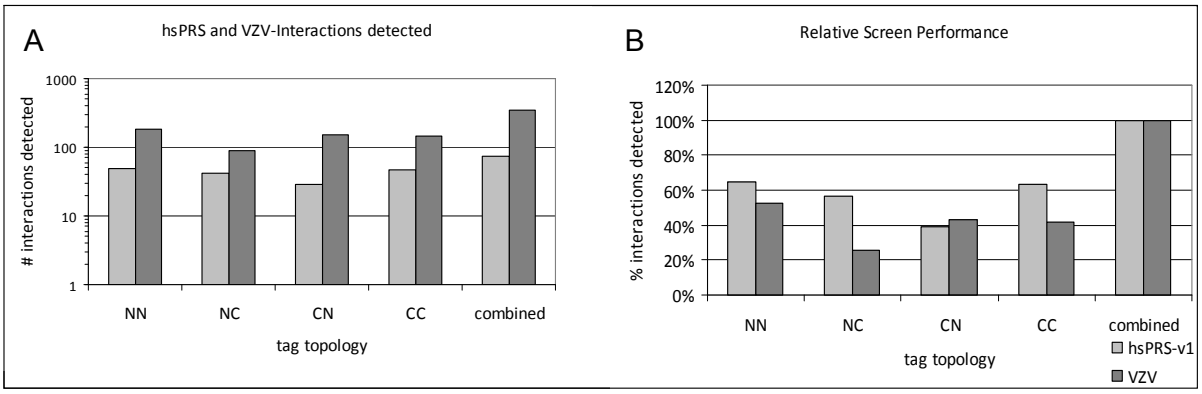


Figure 36: Correlation between hsPRS-v1 and VZV data.
A) Number of reproduced hsPRSv1 interactions and the respective VZV interactions. The hsPRS-v1 interactions from four different tag-topology combinations (see Table 12) and the respective data from the VZV screen (c.f. Table 7) were put in the same histogram. In regard of the different sample sizes the absolute number of interactions was plotted on a logarithmic y-axis. The combined interaction numbers refer to non-redundant interactions (74 of the hsPRS-v1 set and 348 VZV-interactions). **B)** Relative proportions of hsPRS-v1 and VZV interactions. The number of non redundant interactions from all four tag-topology combinations was set to 100 % and the individual interaction numbers from the different screens were calculated as a proportion of all interactions. For the detailed numbers see Table 13.

tag permutation	% of all detected hsPRS-v1 interactions	% off all detected VZV interactions
NN	65	52
NC	57	26
CN	39	43
CC	64	42
combined	100	100
p-value	0.10879	

Table 13: Relative contributions of single tag permutations are equal.
Number of Interactions detected in the single tag permutation combinations (data from Table 7 & Table 12) normalized to the total number of non-redundant interactions. Proportions of NN, NC, CN and CC interactions of the hsPRS-v1 and VZV-dataset were compared to each other using a two-tailed t-test for heteroskedastic samples. The resulting p-value of 0.11 is not sufficient to decline the null hypothesis h_0. It can be declined if the p-value is equal or smaller 0.05 (R-Development-Core-Team, 2004).

3.2.5.1.2 hsRRS-v1 interactions

The 92 hsRRS-v1 interactions were tested against each other in combinatorial screens of rising stringency as described for the hsPRS-set (Yu-Chi Chen, Seesandra Venkatappa Rajagopala). A maximum of 16 non-redundant interactions were detected in at least one screen which would account for a maximum false-positive (FP)-rate of 17.4 %. By increasing the assay-stringency with 3-AT the false-negative rate quickly declines to zero. A detailed interaction matrix is shown in Figure

37. The FP-rates according to the fusion-tag orientations are showed in Table 14. The FP-rates were subsequently plotted against the assay stringency where a rapid decrease could be observed with increasing stringency (see Figure 38).

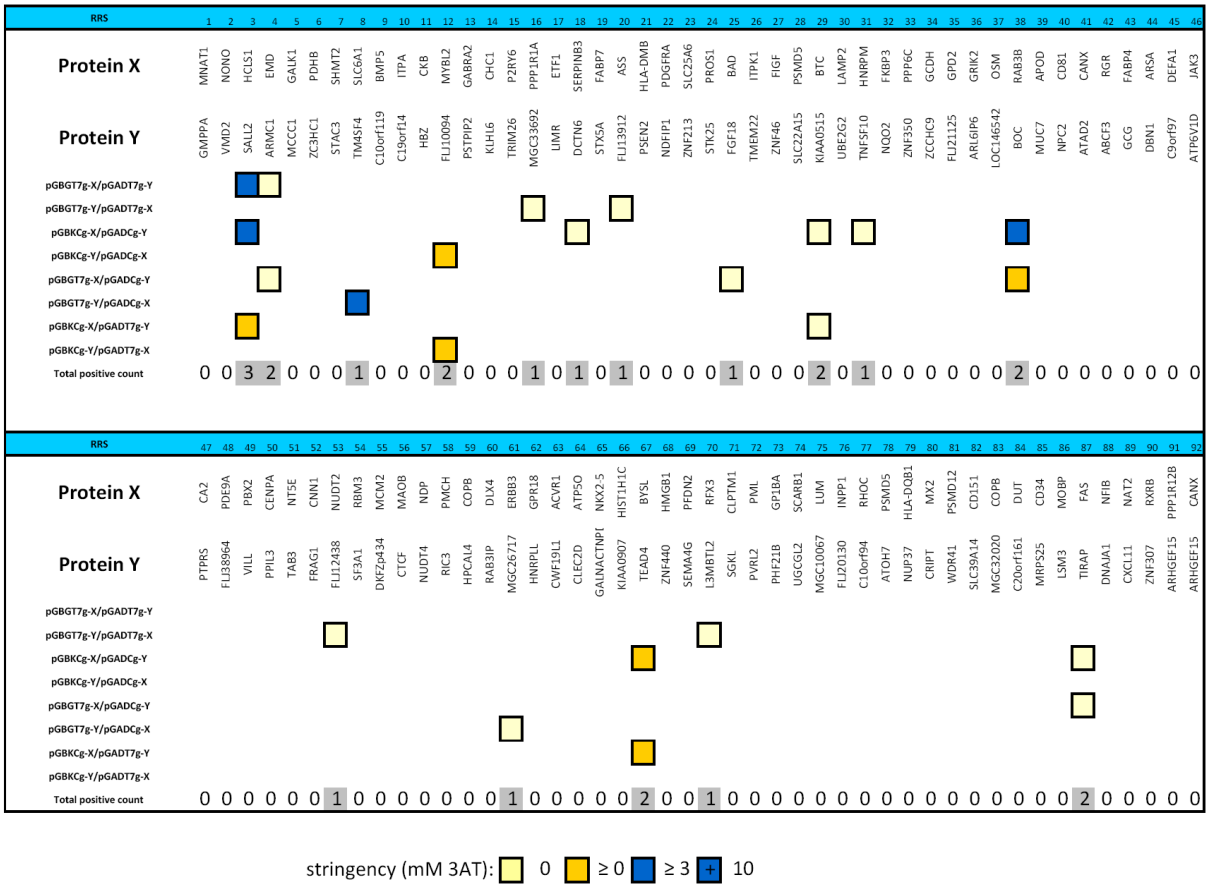


Figure 37: hsRRS-v1 interactions detected in combinatorial screens.
Matrix of interacting hsRRS-v1protein pairs according to the bait-prey vector combinations they were detected with as described before the hsPRS-v1 interactions (see Figure 35). Adapted from: Chen et al., 2010.

tag topology	stringency [mM 3-AT]			
	0	≥ 0	≥ 3	10
	% false-positives			
NN	5.4 %	0.0 %	1.1 %	0 %
CC	4.3 %	2.2 %	2.2 %	0 %
NC	4.3 %	2.2 %	1.1 %	0 %
CN	1.1 %	3.3 %	0.0 %	0 %
combined	12.0 %	4.3 %	3.3 %	0 %

Table 14: hsRRS-v1 interactions depending on the assay stringency.
Number of Interactions detected in the combinatorial hs-RRS-v1 screens depending on the concentration of 3-AT they were detected with (see Figure 37).

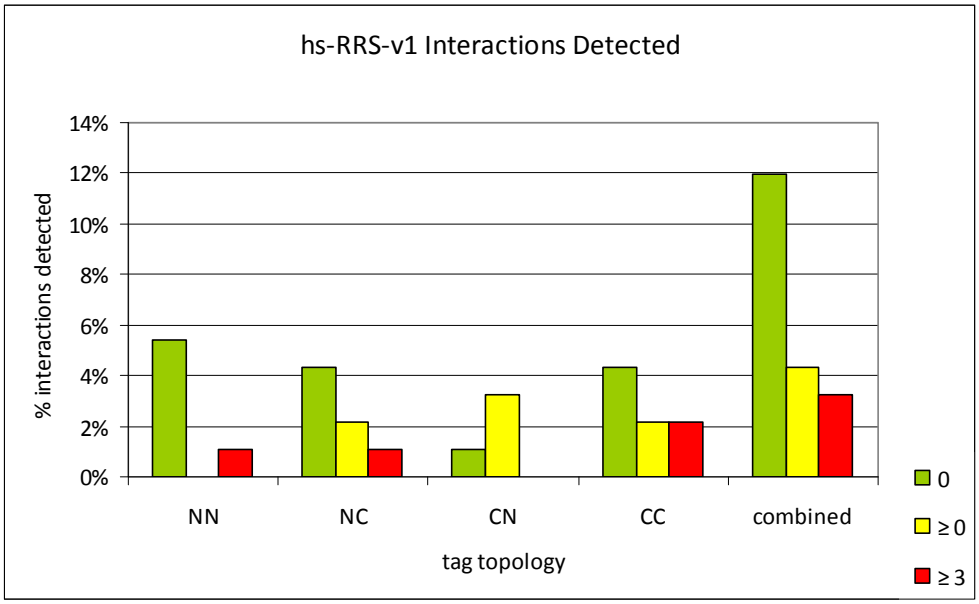


Figure 38: hsRRS-v1 interactions detected in combinatorial screens.
The rate of false-positive interactions decreases with higher stringency applied to the Y2H assay. Assays were repeated with rising concentrations of 3-AT on the Y2H readout-medium (see figure legend, c = [mM 3-AT]). Interactions were counted positive when unspecific background due to possible autoactivation had disappeared. At a 3-AT concentration of 10 mM no more false positive interactions were detected.

3.3 VZV Terminase complex retest

A more detailed look was taken on the interactions between VZV DNA-encapsidation gene products, putative terminase complex subunits (Visalli et al., 2007). In herpesviruses and many tailed double-stranded DNA (dsDNA) bacteriophages, the newly synthesized viral genomes in the infected host cell are present as concatemeric DNA that is comprised of multiple copies of the genome. The concatemeric DNA is packaged into the procapsid through a channel formed by a portal protein complex embedded in a unique vertex of the procapsid (Rao and Feiss, 2008).

For the retests, the respective entry vectors of full-length ORFs and fragments available from the VZV library were sequenced and newly recombined by LR-reaction into the Y2H bait and prey vectors for N- and C-terminally fused test domains and subsequently transformed into the reporter yeast strains Y187 and AH109. All topological bait-prey combinations were subsequently tested. Figure 39 shows the test-plates of ORF42 screened as bait against the potential terminase subunits. The interaction data is listed in Table 15 and was subsequently integrated

into an interaction matrix which gives an overview about all unique interactions detected in the complete library screens, predictable from four more HV-species and found in the retest-screen (see Figure 40). Based on this matrix I designed an interaction network of the DNA-packaging proteins providing a graphical overview of the relationships between these components and their functions, if known (see Figure 56).

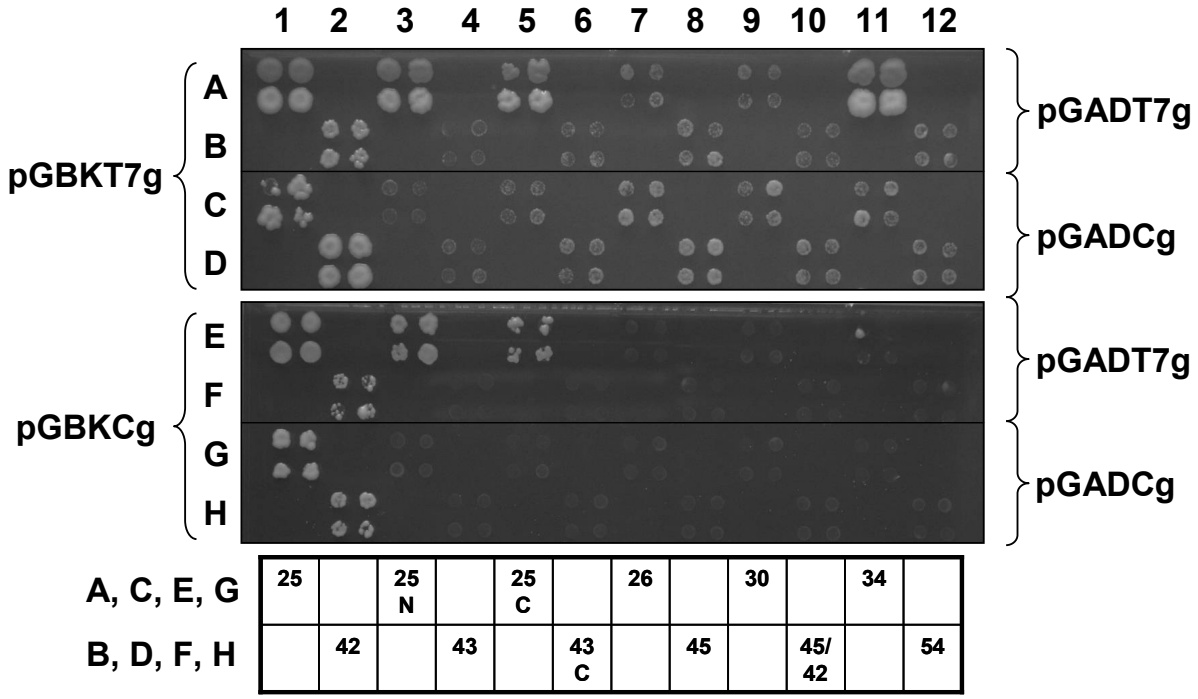


Figure 39: Systematic Y2H-retest of the VZV terminase complex.
Terminase subunits were systematically retested against each other. For example, ORF42 bait constructs were mated against the terminase prey array (see box below, referring to the respective ORF numbers) containing all putative terminase subunits as N-terminal AD-fusion (pGADT7g) and C-terminal AD-fusion (pGADCg). Baits were screened as N- and C-terminal DBD-fusions (pGBKT7g and pGBKCg) on appropriate 3-AT concentrations (0 and 1 mM 3-AT, respectively) for seven days on readout medium lacking histidine. Note that the interactions detected as N- and C-terminally tagged baits are the same, except the one between ORF25 and ORF34 which is only detected in N-N conformation (position A12).

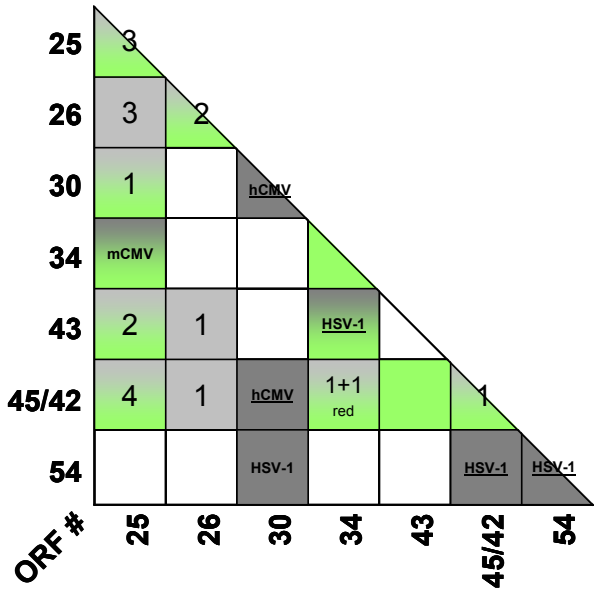


Figure 40: Combined interaction matrix of putative terminase complex members.
This interaction-matrix contains 19 non-redundant interactions derived from Table 15. Light grey boxes indicate interactions from the full-library screens; ciphers indicate the respective number of tag-topology combinations the respective interaction was found with (as described in Chapter 3.1). The interaction of ORF34 and ORF45/42 (represented by the C-terminal ORF42) was found redundant in both bait-prey directions, each time in one permutation-combination, indicated as "1+1 red". Dark grey boxes indicate interologous predicted interactions from the labeled HV-species, from Y2H screens performed in our group or literature-curated from small scale studies (underlined). Green boxes represent interactions found in the retest-screen, semi-green boxes for verified Y2H or interologous interactions. Plain-green boxes represent novel interactions found in the retest-screen.

bait	prey	Y2H tag topology	comments and references	source
ORF25	ORF25	NN_NC_CN	HSV-1 interolog (Chapter 3.1) (Visalli et al., 2009)	Terminase retest
ORF25	ORF25C	NN	HSV-1 interolog (Chapter 3.1)	Terminase retest
ORF25N	ORF25	NN	HSV-1 interolog (Chapter 3.1)	Terminase retest
ORF25N	ORF25N	NN	HSV-1 interolog (Chapter 3.1)	Terminase retest
ORF25N	ORF25C	NN	HSV-1 interolog (Chapter 3.1)	Terminase retest
ORF25C	ORF25	NN_NC	HSV-1 interolog (Chapter 3.1)	Terminase retest
ORF25C	ORF25C	NN	HSV-1 interolog (Chapter 3.1)	Terminase retest
ORF26	ORF26	NN	EBV interolog (Chapter 3.1)	Terminase retest
ORF30	ORF25	NN	HSV-1 interolog (Jacobson et al., 2006) (Visalli et al., 2009)	Terminase retest
ORF34	ORF25	NN_NC	MCMV interolog confirmed (Fossum et al., 2009)	Terminase retest
ORF34	ORF34	NN	new identified	Terminase retest
ORF34	ORF42	NC	Chapter 3.1, redundant direction	Terminase retest

bait	prey	Y2H tag topology	comments and references	source
ORF42	ORF25	NN_NC_CN_CC	Chapter 3.1 (Visalli et al., 2009)	Terminase retest
ORF42	ORF25N	NN_CN	Chapter 3.1	Terminase retest
ORF42	ORF25C	NN_CN	Chapter 3.1	Terminase retest
ORF42	ORF34	NN	Chapter 3.1	Terminase retest
ORF42	ORF42	NN_NC_CN_CC	HSV-1 interolog (Abbotts et al., 2000)	Terminase retest
ORF43C	ORF25	NN_NC_CN_CC	Chapter 3.1 (Visalli et al., 2009)	Terminase retest
ORF43C	ORF25N	CN	Chapter 3.1	Terminase retest
ORF43C	ORF34	NN	HSV-1 interolog (Trus et al., 2007)	Terminase retest
ORF43C	ORF42	NC	new identified	Terminase retest
ORF25	ORF25	NN_NC_CC	HSV-1 interolog (Chapter 3.1) (Visalli et al., 2009)	Full screen+Interologs
ORF26	ORF25	NC_CN_CC	Chapter 3.1	Full screen+Interologs
ORF26	ORF26	NN_CC	EBV interolog (Chapter 3.1)	Full screen+Interologs
ORF30	ORF25	NN	HSV-1 interolog (Jacobson et al., 2006) (Visalli et al., 2009)	Full screen+Interologs
ORF30	ORF30	-	mCMV interolog (Savva et al., 2004)	-
ORF34	ORF25	-	mCMV interolog (Fossum et al., 2009)	-
ORF42	ORF25	NN_NC_CN_CC	Chapter 3.1 (Visalli et al., 2009)	Full screen+Interologs
ORF42	ORF26	CC	Chapter 3.1	Full screen+Interologs
ORF42	ORF34	NN	Chapter 3.1	Full screen+Interologs
ORF42	ORF54	-	HSV-1 interolog (White et al., 2003)	-
ORF42	ORF42	CN	HSV-1 interolog (Abbotts et al., 2000)	Full screen+Interologs
ORF43	ORF25	CN_CC	Chapter 3.1 (Visalli et al., 2009)	Full screen+Interologs
ORF43	ORF26	CC	Chapter 3.1	Full screen+Interologs
ORF45/42	ORF30	-	EBV interolog (Fossum et al., 2009)	-
ORF54	ORF30	-	HSV-1 interolog (Yang et al., 2007)	-
ORF54	ORF54	-	HSV-1 interolog (Trus et al., 2004)	-

Table 15: Interaction data of putative terminase subunits.
The interaction data in this table is derived from the full-library screens, interologous interactions from three additional herpesviruses, both Y2H- and LC-derived (see references), and the interactions from the retest.

3.4 Mapping of the ORF25 homomerization interface.

ORF25 encodes is the most connected protein among the DNA-packaging proteins and it was proven to be essential for VZV replication in a collaboration study at the Max-von-Pettenkofer-Institut at the Ludwig-Maximillians-Universität (LMU), Munich (Vizoso Pinto et al., Manuscript in preparation). Moreover, this study suggests ORF25 protein being a chaperone, which provides an explanation for the high number of interaction partners, also beyond those proteins which are involved in DNA-packaging. Six of the seven annotated DNA-packaging proteins interact with ORF25, including a self interaction, indicating that ORF25 protein forms multimeric complexes.

3.4.1 Yeast two-hybrid and Peptide array mapping

The ORF25 amino acid sequence was spotted as a peptide-scan of 15mers with a three AA shift on a cellulose membrane and probed with purified MBP-ORF25 protein. Interacting peptides were detected immunologically and visualized by Enhanced Chemiluminescence (ECL) on light sensitive film as described in Chapter 2.2.5. ORF25 was expressed with an N-terminal MBP-tag from the vector pETG-40K and subsequently purified according to Chapter 2.2.4.4.

Three interacting sequences were identified (see Figure 41). The results obtained with the peptide arrays were compared to Y2H- interactions found among the ORF25 full-length construct and N- and C-terminal fragments thereof (see Figure 42). Yeast two-hybrid analysis could show that beyond the interaction between the full-length ORF25, the N- and C-termini interact with each other, as well as the N- with the C-terminus. These results are supported by the peptide arrays, as N- and C-terminal halves contain interacting peptides.

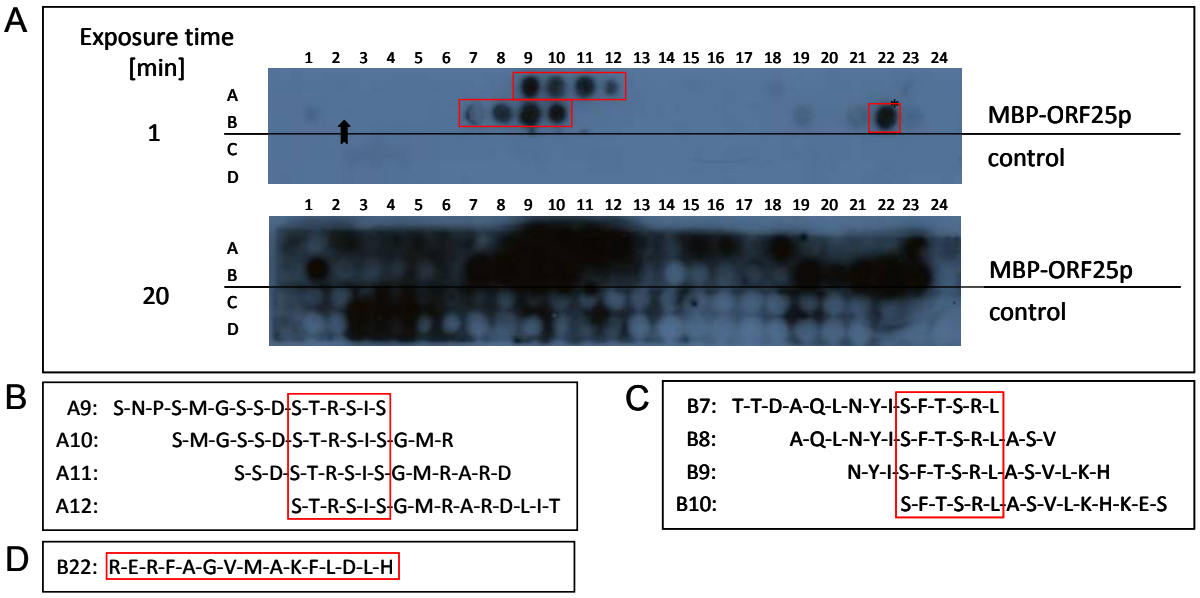


Figure 41: Mapping of the ORF25 homomerization interface.
The ORF25 AA sequence was spotted as 15mers with a three AA shift on a β-Alanine esterfied membrane support. Peptide positions are defined by row (A-D) and column (1-24). Positions A1 to B24: ORF25 array hybridized with MBP-ORF25p [10 µg/ml]. Black spots indicate peptides with bound ORF25p, detected immunologically by the MBP-tag. The respective control array (C1 to D24) was hybridized with 10 µg/ml purified MBP and processed in parallel to the test-array. The Arrow indicates the border between spots bearing sequences unique to Y2H-constructs ORF25N (A1-B2) and ORF25C (B3-B24) which were used for Y2H-mapping of the ORF25p self-interaction.

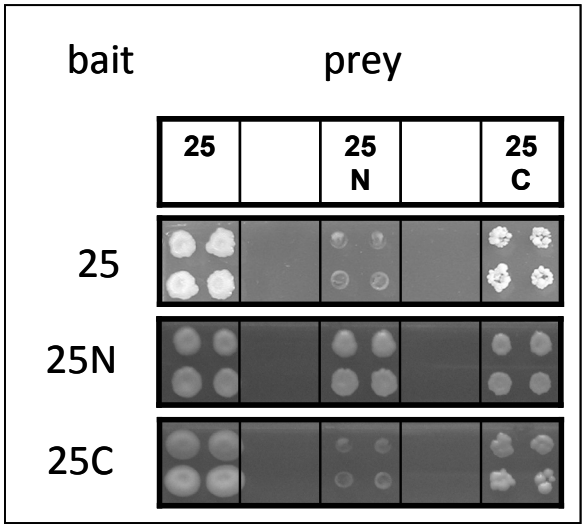


Figure 42: Y2H-mapping of the ORF25 self-interaction.
Extracts from the VZV terminase retest showing the interactions between ORF25 constructs (see Chapter 3.3). Interactions were detected by pGBKT7g bait-, and pGADT7g prey vectors. Overlapping results using C-terminally tagged bait and prey constructs are not shown. This results show that ORF25p N- and C-terminal constructs interact with each other as well as the N- and C-terminal constructs interact with themselves.

3.4.2 Bioinformatical analysis of the ORF25 mapping results.

To investigate whether the interacting peptides of ORF25 are conserved, I performed a multiple sequence alignment between the eight human pathogenic herpesviruses. The sequences orthologous of the ones detected in VZV (HHV-3) were highlighted in red boxes and consecutively given the numbers I, II and III, according to their position in the protein from N- to C-terminus (as shown in Figure 41 B-D). It turned out that the interacting sequence I is conserved in HHV-1 to HHV-4 but absent in HHV-5 to HHV-8. The central and C-terminal sequences II and III are conserved throughout all herpesvirus subfamilies (see Figure 43).

3.4.2.1 Comparison of ORF25 and its HSV-1 ortholog UL33.

To proof the *in vivo* relevance of interaction interfaces that I obtained from the peptide arrays, I compared the results of a recently published study where the HSV-1 ortholog of VZV ORF25, UL33, was systematically mutated. Mutations of UL33 were obtained by random transposon-mediated insertion of five codons resulting in 15 mutants in 14 distinct regions throughout the whole ORF. The resulting mutants were tested for viral growth, viral genomic DNA-packaging and further, mutated UL33 proteins were tested for their ability to interact with the large terminase subunit, UL28p by Co-IP (Beilstein et al., 2009).

Comparison of the pairwise aligned sequences shows a remarkable correlation between the sequences I to III, identified to be involved in self-interaction of ORF25 and the orthologous sequences of UL33, which show a replication and DNA-packaging phenotype in HSV-1 (Figure 44). While the mutation in the proximity of sequence “I”, which is not conserved throughout the herpesvirus family, has only a mild phenotype in HSV-1, both the central and C-terminal mutations, which overlap with the interacting sequences II and III completely inhibit viral growth and DNA-processing. In terms of the interaction with the large terminase subunit UL28, the N- and C-terminal interaction sites play no or just less important roles, respectively and may therefore be in reverse more important for the ORF25p self-interaction.

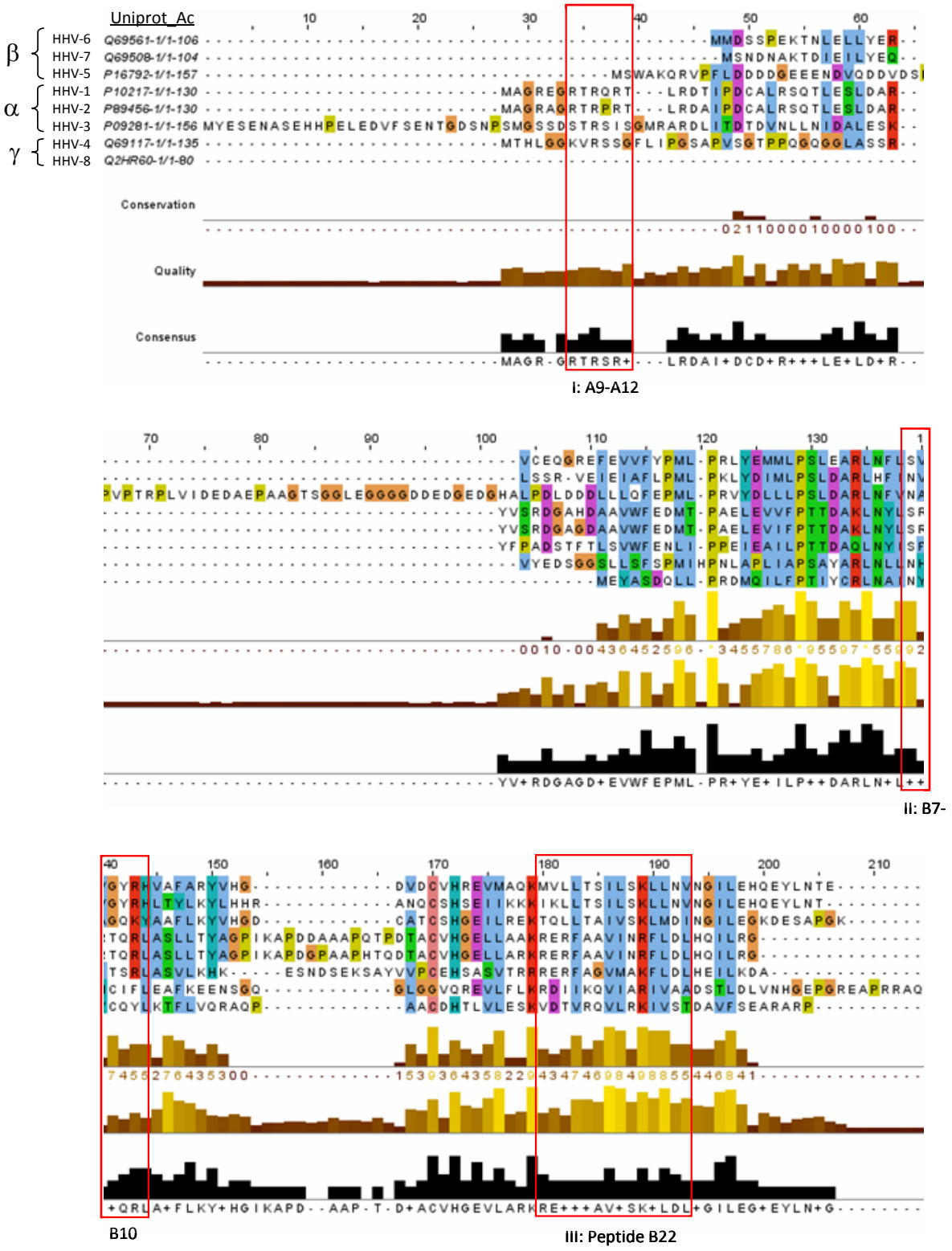


Figure 43: Multiple sequence alignment- ORF25 orthologs in human pathogenic herpesviruses. Amino acid sequences of the eight human pathogenic herpesvirus species, HHV-1 to HHV-8 were retrieved from the Uniprot Knowledgebase (UniProt-Consortium, 2010). Sequences were aligned directly from the Uniprot website (www.uniprot.org) using ClustalW version 2 online via the EBI ClustalW server and visualized by default ClustalX settings (Larkin et al., 2007). See Table 3 for the applied colour scheme. HHV-1 - 8: Human Herpesvirus species 1 to 8. VZV = HHV-3. Indicated at the left is the respective subfamily within the *Herpesviridae*, *Alpha-*, *Beta-*, or *Gammaherpesvirinae*. Red boxes mark the orthologous sequences of interacting peptides identified in Figure 41.

Nevertheless, the central domain of UL33 and probably ORF25 plays a versatile role in protein interactions and can be regarded as the most promising candidate for drug targeting of this protein. To identify the essential residues in the AA sequences “I” and II, I spotted Alanine substitution arrays and hybridized them once more with MBP-ORF25p. In sequence II, the central interacting interface, exchange of the Arginine residue at position 104 in ORF25p (S-F-T-S-**R**^{104}-L) lead to the loss of the protein-peptide binding (Figure 45B). Sequence “I” showed a higher sensitivity against amino acid exchange which might reflect a lower binding affinity to this sequence. Single substitutions of the three consecutive residues 36-38 (S-T-**R**^{36}-**S**^{37}-**I**^{38}-S) avoided binding of MBP-ORF25p (Figure 45A). Comparing both sequences, a high similarity of both sequences is observable. The common binding motif can be described as S-[FT]-[TR]-S-[RI]-[LS], while AAs in braces are either in sequence “I” or sequence II. Remarkable is the alternation of the polar, hydrophilic amino acids Serine (S) and Threonine (T) and the hydrophobic residues Isoleucine (I), Leucine (L) and Phenylalanine (F) in both sequences, which is interrupted by a single basic Arginine (R) residue, which plays apparently a pivotal role in the binding. Significantly, Arginine residues and commonly basic residues can play essential roles in protein-protein and protein-DNA binding (Leung et al., 2010).

Last step of investigating the ORF25p self-interaction was to specify the C-terminal interacting sequence III, which was detected only as a single interacting peptide on the peptide array (see Figure 41D). Longer exposure times of the light sensitive film at MBP-ORF25p detection revealed additional signals, probably unspecific, which were included in the following analysis.

To localize the interaction epitope, I synthesized a peptide array of the signaling amino acid sequence in a "Hidden-epitope assay". The 29 AA sequence was first spotted as 8mers with a two amino acid shift, and then repeated as 7mers, 6mers, 5mers and 4mers. In doing so, the minimum interacting epitope can be identified. Figure 46 shows the test array and the respective control, identifying the amino acid sequence V^{142}-M-A-K-F-L as minimum requirement for MBP-ORF25p binding. Additionally, a sequence nine amino acids closer to the N-termnus, T^{133}-R-R-R-E, gave a weak but specific signal. However, the adjacent AA sequence R^{136}-E-R-F was bound very strong by the MBP control protein which makes it very difficult to make definite predictions on this interaction (results of this assay are additionally summed up in Table 16).

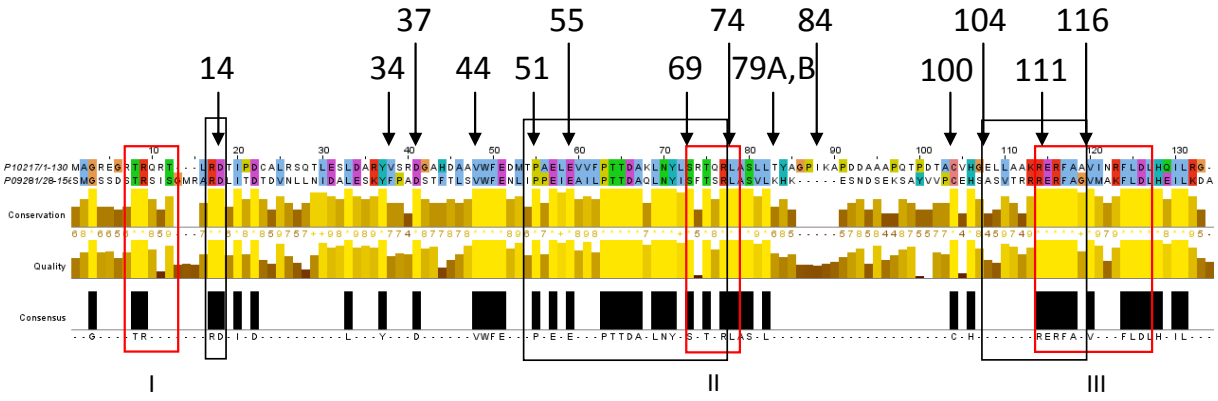


Figure 44: Pairwise sequence alignment of HHV-1 UL33 and HHV-3 ORF25.
Red boxes indicate interacting sequences I to III derived from HHV-3 (VZV ORF25, UniProt Accession P09281) protein peptide interaction assays and the orthologous sequences of HHV-1 (HSV-1 UL33, UniProt Accession P10217). Numbers and arrows point to five amino acid insertion sites from the UL33 mutagenesis study (Beilstein et al., 2009). 79A, B: Two insertion mutants in the same codon, differing in the inserted AAs and showing the same phenotypes. Black boxes span sequences within which mutations caused defects in viral replication and DNA-packaging. The 27 N-terminal AAs of ORF25 with no homology to UL33 were removed for improving figure clarity.

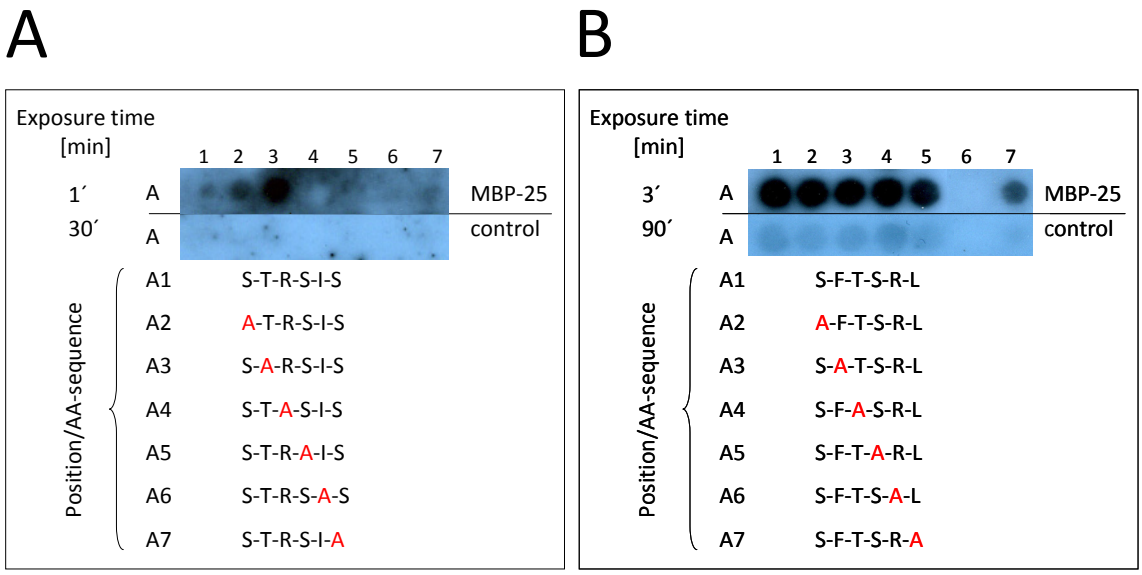


Figure 45: Alanine substitution of the ORF25 interacting peptides I & II.
The interacting sequence I and II of ORF25 were spotted on a β-Alanine esterfied cellulose membrane. Single amino acids were successively substituted by alanine (A). The arrays were hybridized with MBP-ORF25p [10 µg/ml] and bound protein immunologically detected. As a control, the membranes were stripped (see Chapter 2.2.5.1.4) and hybridized once more with 10 µg/ml purified MBP and subsequently processed as in the test assay before. **A)** Alanine substitution of sequence I showed that the residues exchanged in the spots A4 to A6 are essential for the MBP-ORF25p binding to the peptide S-T-R-S-I-S (R^{36}, S^{37}, I^{37} of ORF25p), according to the missing signals at the positions A4 to A6; **B)** Substitution analysis of sequence II revealed that the Arginine residue (R^{104} of ORF25p) is essential for the MBP-ORF25p binding to the peptide S-F-T-S-R-L as binding disappears at the spot at position A6.

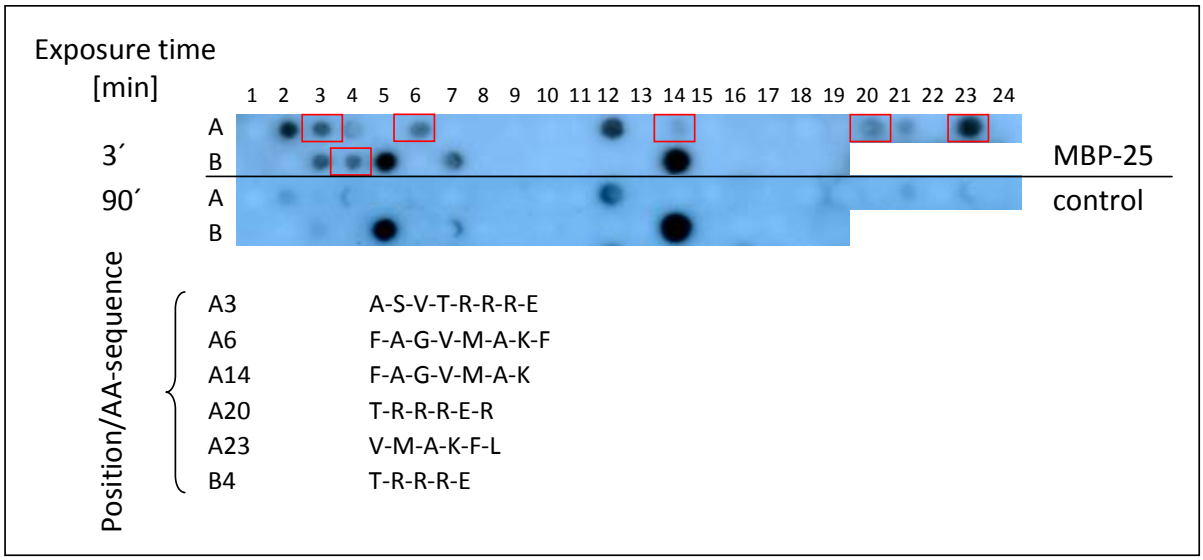


Figure 46: Hidden-epitope assay of the ORF25 interacting sequence III.
The C-terminal interacting domain was spotted on a β-Alanine esterfied cellulose membrane. The 29 AAs spanning sequence within the spots B18 to B22 in Figure 41, was synthesized with a two AA shift as 8mers, then as 7mers and in decreasing manner up to 4mers to identify the minimum interacting epitope hidden in the ORF25 C-terminus. The array was hybridized with MBP-ORF25p [10 µg/ml] and bound protein immunologically detected as already described. As control, the membrane was stripped (see Chapter 2.2.5.1.4) and hybridized once more with 10 µg/ml purified MBP and processed as in the test assay before. Specific signals are highlighted in red boxes and respective positions and AA sequences are shown.

position on array	AA-sequence	assay	control	position on array	AA-sequence	assay	control
A1	V-P-C-E-H-S-A-S	-	-	B1	V-P-C-E-H	-	-
A2	E-H-S-A-S-V-T-R	++	+	B2	E-H-S-A-S	-	-
A3	**A-S-V-T-R-R-R-E**	+	-	B3	A-S-V-T-R	+	+
A4	T-R-R-R-E-R-F-A	+	+	B4	**T-R-R-R-E**	+	-
A5	R-E-R-F-A-G-V-M	-	-	B5	R-E-R-F-A	++	++
A6	**F-A-G-V-M-A-K-F**	+	-	B6	F-A-G-V-M	-	-
A7	V-M-A-K-F-L-D-L	-	-	B7	V-M-A-K-F	+	+
A8	M-A-K-F-L-D-L-H	-	-	B8	K-F-L-D-L	-	-
A9	V-P-C-E-H-S-A	-	-	B9	F-L-D-L-H	-	-
A10	E-H-S-A-S-V-T	-	-	B10	V-P-C-E	-	-
A11	A-S-V-T-R-R-R	-	-	B11	E-H-S-A	-	-
A12	T-R-R-R-E-R-F	++	++	B12	A-S-V-T	-	-
A13	R-E-R-F-A-G-V	-	-	B13	T-R-R-R	-	-
A14	**F-A-G-V-M-A-K**	+	-	B14	R-E-R-F	+++	+++
A15	V-M-A-K-F-L-D	-	-	B15	F-A-G-V	-	-
A16	M-A-K-F-L-D-L	-	-	B16	V-M-A-K	-	-
A17	V-P-C-E-H-S	-	-	B17	K-F-L-D	-	-
A18	E-H-S-A-S-V	-	-	B18	L-D-L-H	-	-

position on array	AA-sequence	assay	control	position on array	AA-sequence	assay	control
A19	A-S-V-T-R-R	-	-				
A20	**T-R-R-R-E-R**	+	-				
A21	R-E-R-F-A-G	+	+				
A22	F-A-G-V-M-A	-	-				
A23	**V-M-A-K-F-L**	++	-				
A24	K-F-L-D-L-H	-	-				

Table 16: Hidden-epitope assay of the ORF25 C-terminal dimerization domain.
Results from the Hidden-epitope assay (c.f. Figure 46) were listed and evaluated by the signal strength according to the following scheme: **+** = weak-; **++** = robust-; **+++** = strong- and - = no signal. Sequences specific for MBP-ORF25p interaction are highlighted bold. Strong unspecific MBP-binding was detected at R-E-R-F containing sequences, whereas specific signals contain the minimum-epitope V-M-A-K-F-L and T-R-R-R-E, the latter one overlapping the unspecific R-E-R-F sequence (spot at position B14).

4 Discussion

4.1 Combinatorial Y2H screening with permutated fusion tags

The major goal of this study was to develop, apply and to evaluate a Yeast two-hybrid system with permutated fusion tags that may improve the performance of the assay under standardized conditions. It was shown that steric constraints do play a role in the detectability of protein-protein interactions in certain cases, and I performed a systematic analysis of this impact on high-throughput Yeast two-hybrid screens.

As a prerequisite, I have generated two new Yeast two-hybrid vectors, pGBKCg and pGADCg, which allow the C-terminal fusion of the GAL4 DNA-binding and activation domain. Together with the parental vectors for N-terminal fusions, four different bait-prey fusion tag topology combinations are now possible to be screened: NN, CC, NC, and CN. In the next step I tested all approximately 9000 pairwise combinations between 96 Varicella zoster virus proteins and fragments thereof for pairwise interactions in the now possible combinations. About 27,000 individual Y2H tests, each performed twice, resulted additionally to the previously reported 182 NN-interactions (Uetz et al., 2006) in 90 NC, 151 CN, and 146 CC interactions. This study was the first systematic analysis of such N- and C-terminal Y2H vectors.

4.1.1 Structural influence of tag-topologies

The overlap between the four permutation combinations lies between 17 % (NC-CN) and 43 % (CN-CC), regarding the raw data (c.f. Table 7). The highest data-overlaps were observed between bait-prey combinations of the same bait permutation (NN-NC, 42 % and CC-CN, 43 %). This indicates that the orientation of the bait-hybrid, which is sitting as a dimer on the GAL1-UAS (Upstream Activating Sequence) plays the major role in permission of subsequent reporter gene transcription (Figure 47). In the converse argument, the PPI overlaps which are achieved between different bait permutation screens (NN-CN, NN-CC, NC-CN and NC-CC) are smaller (between 17 % and 31 %), thus detecting more different interaction data subsets.

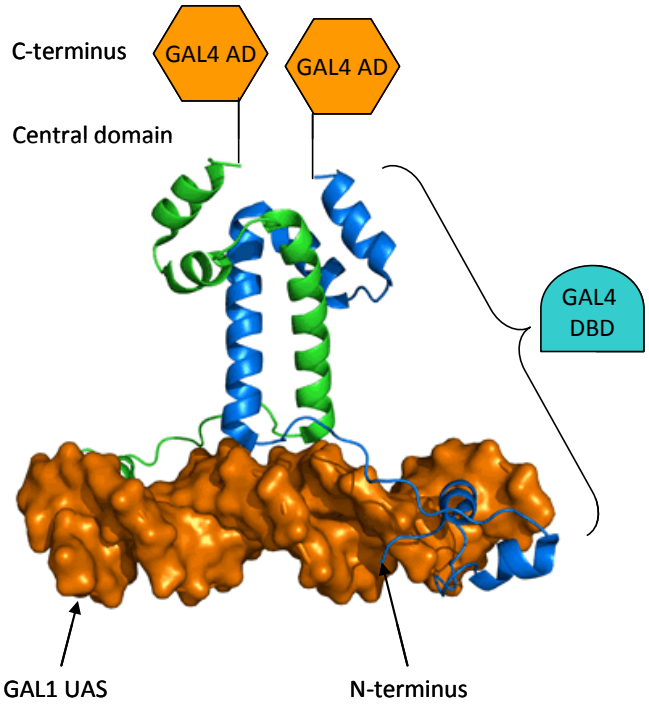


Figure 47: Structure of the Gal4-DNA complex.
Gal4 is a transcription factor for galactose-induced genes (GAL1, GAL2, GAL7, GAL10, and MEL1). These genes encode enzymes which convert galactose to glucose. pGal4 recognizes a 17 base-pair long sequence in the upstream activating sequence (uas-g) of these genes. pGal4 binds to the DNA as a homodimer, most likely forming a complex with GAL11. The DBD is located at the N-terminus, the AD lies at the C-terminus of pGal4. The structure was resolved by Hong and coworkers in 2008 (Hong et al., 2008).

The prey orientation seems to play a minor role. An explanation for this observation could be that if prey binding does occur, the elasticity of the DNA strand allows transactivation of downstream reporter elements. Taken together, these findings support the thesis that steric restrictions are a source of false-negative interactions in Yeast two-hybrid screens.

4.1.2 Reduction of false-negatives and data quality

Screening of four bait-prey tag permutations instead of one resulted in almost exactly twice as much non-redundant interactions (348 compared to 173, regarding full-length ORFs) and thus fewer false-negative interactions. Increasing numbers of interactions are detected by decreasing numbers of tag permutations.
In addition, interactions that are found in multiple combinations confirm each other and thus provide an intrinsic quality score. Moreover, the fraction of interactions which could be confirmed by the Gold-Standard set is rising within the subsets of interactions which were detected in multiple screens. While 36 % of non-redundant interactions could be verified, found in exclusively one tag-topology, the ratio is rising

up to 75 % within the interactions detected in all four combinations. Furthermore, I could show that the unique permutations produce data of equal quality, as the ratio of Gold-Standard interactions lies between 45 % and 50 %.

4.1.3 Conservation of interactions among herpesviruses

Evaluation of the VZV interaction data with regard to the conservation of interactions among five herpesvirus species throughout the alpha-, beta- and gamma-subfamily (c.f. Chapter 3.2.2) could confirm that there is an enrichment of interactions between core-proteins. Regarding the VZV interaction data of this study, there is a 4.2-fold enrichment of core-core interactions within the interactions of core-proteins, compared to literature-curated data in five different herpesviruses (HSV-1, HCMV, MCMV, EBV and KSHV).

4.1.4 Network analysis

The comparison of the VZV network before and after combinatorial screening showed that both networks are scale-free. Both degree distributions follow power law functions, while the extended network shows a better correlation to the fitted power law. Due to the higher sample size the probability of an even node-degree distribution was given. Similarly, the attack tolerance was higher in the smaller NN-derived network, reflecting that this viral interactome only represented a minor part of the complete interactome. These observations confirm the evolutionarily conserved topology of herpesviral PPI networks on the one hand (Fossum et al., 2009), and on the other hand I could proof that no randomization of the network did occur with the newly generated data.

4.1.5 Novel VZV interactions: ORF10-ORF57

The VZV interaction data bears 15 protein-protein interactions, which were found in three or four permutations and have never reported in the literature (see Supplementary Table S1). They should be interesting candidates for further studies of VZV biology.

One example, which was described by Dr. Armin Baiker at the Max-von-Pettenkofer Institut, is the interaction between pORF57 and pORF10 (Stellberger et al., 2010), which we found in all four Y2H tag-permutations. The ORF10 protein is the HSV-1 VP16 ortholog. Both proteins have been shown to be structural components of the

virion tegument, but also appear to activate viral immediate-early (IE) protein promoters (Moriuchi et al., 1993). In contrast to HSV-1, where ORF10 is essential, VZV ORF10 is dispensable for viral replication *in vitro* (Cohen and Seidel, 1994). However, a more detailed analysis of VZV ORF10 deletion mutants in SCIDhu skin xenografts revealed that ORF10 is a virulence factor for the pathogenesis of VZV in skin (Che et al., 2007; Che et al., 2006). VZV ORF10 mutants are characterized by decreased viral titers and decreased cutaneous lesions within skin xenografts. Electron microscopy pictures showed that VZV-infected epidermal cells had significantly fewer DNA containing nucleocapsids and extensive aggregates of cytoplasmic viral particles (Che et al., 2006). Little is known about the function of the VZV ORF57 protein. This protein has been shown to localize to the cytoplasm of infected cells and to be dispensable for the replication of VZV in vitro (Cox et al., 1998). However, UL3.5, the homologue of VZV ORF57 in Pseudorabies virus (PrV), is essential for virus replication *in vitro*, playing an important role in viral egress. Electron microscopic pictures showed that cells infected with PrV UL3.5 deletion mutant virus exhibited accumulated cytoplasmic viral capsids (Fuchs et al., 2007; Fuchs et al., 1996). Accordingly, deletion mutants of VZV ORF10 and the UL3.5 homologue of VZV ORF57 in PrV exhibit identical phenotypes, the cytoplasmic aggregation of viral particles, and thus suggest that both proteins play an essential role in virus egress. However, their precise molecular role in this process is not known.

4.2 Combinatorial screening of human reference sets

4.2.1 False-negative interactions

A main issue working with the Yeast two-hybrid system is the high rate of false-negatives (Braun et al., 2009; Rajagopala et al., 2007). With rising numbers of binary protein interaction studies, dataset quality has been questioned. A first study that compared several interaction data sets to a Gold Standard of high-quality protein complexes suggested that high-throughput Y2H data is of poor quality (von Mering et al., 2002). A more recent analysis showed that protein complexes are inappropriate for evaluating Y2H data and that high-throughput Y2H data are of high quality when compared against a Gold Standard of directly interacting proteins (Yu et al., 2008).

The assay sensitivity is the fraction of all biological occurring interactions (true-positives) that can be identified by an assay which is performed under specific experimental conditions. Like other protein interaction detection systems, the Y2H system is not able to detect all PPIs that occur in vivo (Braun et al., 2009). An assay sensitivity of 25 % means conversely that 75 % of biological relevant interactions are missed, those are false-negative interactions.

A first approach to evaluate the outcome of Yeast two-hybrid screens was performed with the help of a positive reference set of binary yeast protein-protein interactions, comprising 116 binary PPIs which were curated in more than four publications (Yu et al., 2008). In this case, an assay sensitivity of 20 % was achieved, while a second method, an YFP-based protein fragment complementation assay (PCA) (see Figure 48) detected 18 % PRS-interactions, which in addition detected 5 % of probably false-positive interactions. This was investigated in a similar way, with the use of a second reference set of 116 randomly chosen protein-protein combinations, a random reference set (RRS) (False positive interactions will be discussed separately).

New vector systems as the one presented here can decrease the number of false-positive interactions. It was shown recently, that alternative N-terminal vector systems can reduce the number of false-positives (Rajagopala et al., 2009) and this study expands this strategy with C-terminal fusions and their combination with classical N-terminal vectors.

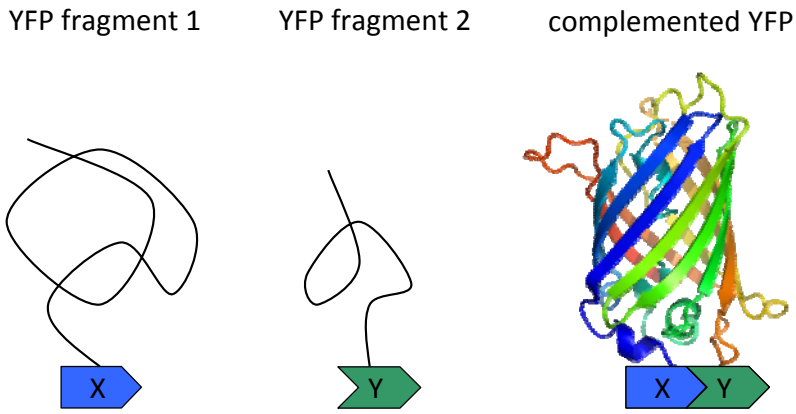


Figure 48: YFP-PCA system.
Principle of the Yellow fluorescent protein-protein complementation assay (YFP-PCA). YFP fragments 1 (amino acids 1–158) and 2 (amino acids 159–239) are fused to a protein X and a protein Y. Protein interaction of the proteins X and Y brings the two fragments of YFP into close proximity and leads to complementation into functional fluorescent YFP by folding into an active 3D structure. Subsequently, fluorescence can be detected by FACS-analysis or fluorescence-microscopy. (Adapted from: Nyfeler et al., 2005)

4.2.1.1 hsPRS-v1 screening

In the following, a systematic analysis of protein-protein interactions of a positive reference set of well described human protein pairs (hsPRS-v1) by Braun and coworkers allowed the calculation of the assay sensitivity of five different detection methods, including the Yeast two-hybrid system (Braun et al., 2009). For the given standard assay conditions of the group (low copy plasmids, one of two distinct reporters activated), the Yeast two-hybrid system lay in the middle position at place three, with an assay sensitivity of 25 % (23 out of 92 interactions detected), including 25 % of the assayed phosphorylation-dependent interactions (two out of eight interactions). The highest assay sensitivity was reported for the LUMIER-system with 36 % (33/92) of the PRS-interactions detected. LUMIER is an automated high-throughput technology designed for the systematic mapping of dynamic protein-protein interaction networks in mammalian cells (Barrios-Rodiles et al., 2005), see Figure 49. However, when they used an assay layout that is more comparable to our classical Y2H-system (high-copy plasmids, one reporter gene) 40 % (34/92) of the PRS interactions were found.

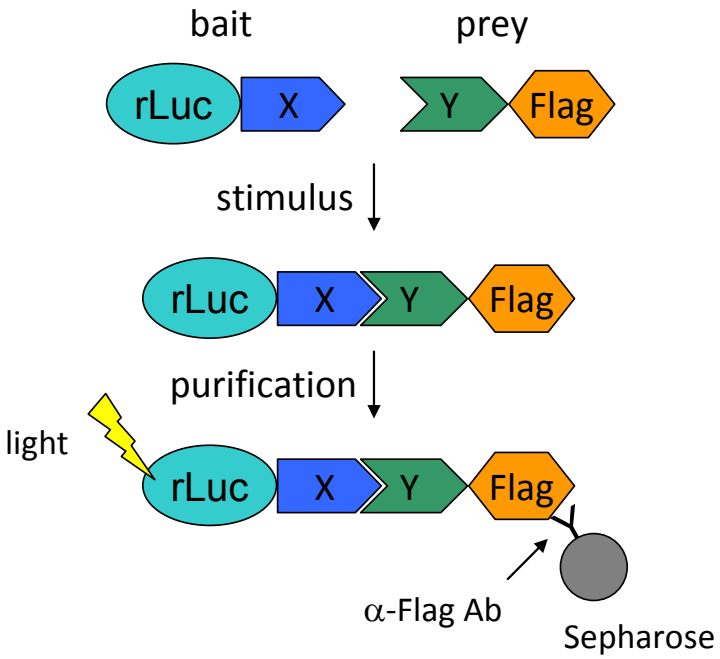


Figure 49: LUMIER system.
Renilla luciferase tagged bait co-expressed in mammalian cells with a Flag-tagged prey is detected in immunoprecipitates (purification) enzymatically as light emission. It was designed to identify PPIs in dynamic signaling pathways that can be induced by a defined stimulus. (Adapted from Barrios-Rodiles et al., 2005).

Using the classical Yeast two-hybrid system with N-terminal fused domains which is well established in our research group, more than half of the hsPRS-v1 interactions (52 %, 48 out of 92 interactions) could be detected. That means, that the classical layout detected already more than twice as much of the interactions reported by Braun et al., 2009. However, the assay conditions and the vectors used in both studies are not equal and though give rise to explanation possibilities of these differences, even though the same tag-topology was used.

4.2.1.2 Plasmid copy number and expression level

The copy number of a plasmid has an influence on the sensitivity of a Y2H assay. Braun and coworkers have shown that the use of high copy plasmids instead of low-copy plasmids increases the assay sensitivity. High copy plasmids carry a 2µ origin of replication and exist in 50 to 100 copies per cell, while ARS/CEN- (autonomously replicating/centromeric sequence) based plasmids occur only as one or two copies in a yeast cell (Van Criekinge and Beyaert, 1999). The elevation of the copy number presumably leads to a higher transcription rate of reporter genes. But the protein expression level has been shown to be not equivalent to the plasmid copy number, the difference in protein levels is only 20- to 30-fold (Van Criekinge and Beyaert, 1999).

34 interactions of the PRS were found with high-copy plasmids by Braun et al., 2009, 25 thereof were reproduced by our standard high-copy plasmids pGBGT7g and pGADT7g (74 %). Regarding the low-copy results, 19 of 23 interactions were reproduced (83 %). This indicates that the low copy number bait- and prey-plasmids mainly detect a subset of the high copy number plasmids due to lower assay stringency caused by a higher abundance of hybrid proteins. But obviously the plasmid copy number alone does not solely result in different Y2H sensitivity, because the CEN/ARS based vectors don't produce a simple subset of 2µ based interactions, which can be considered to be stronger interactions or differently described, interactions between proteins with a high affinity for each other. But this is not the only explanation, as there are still differences that may be a result of other differences in the used vector systems.

A second factor that has an impact on the expression level of the hybrid proteins is the promoter that drives their expression. The ADH1 promoter is a strong constitutively expressing promoter of the *S.cerevisiae* Alcohol dehydrogenase 1

(Grimm et al., 1991) and is used in most GAL4-based two-hybrid systems (Van Criekinge and Beyaert, 1999). However, there are two types of the ADH1 promoter in use, the full-length nucleotide sequence and a truncated version. The bait vectors that I used for instance, pGBKT7g and pGBKCg use a truncated version of the promoter, denoted ADH1*-promoter. In contrast to the full-length promoter the expression level is lower, which compensates for the higher expression level of high-copy plasmids compared to low copy plasmids. For the above mentioned comparison of PRS interactions found with high and low copy plasmids, however, this still holds true as both systems use ADH1* promoters (Vidal et al., 1996; Vidalain et al., 2004).

4.2.1.3 Cloning sites and linker sequences

Even if the Y2H assay layout is the same, with bait and prey fusions using GAL4 DBD and -AD, respectively, there are differences between vector systems which are reflecting in the amino acid sequence of the test constructs. Figure 50 shows some examples of bait and prey vectors which are commonly used in our research group. Bait vectors differ mainly in the length of the linker sequence between the DBD and the test insert ("variable" sequence). Secondly, additional C-terminal fused peptides appear, if the insert contains no stop codon (c.f. Figure 50A pDEST32 and pGBKT7g). Conversely, the C-terminus reflects the native situation, if test constructs are cloned with stop codon (see for example Figure 50B, pLP-GADT7). So, omitting stop codons has the disadvantage that the native C-termini get lost. Depending on the vector used, C-terminal peptides arise, depending on when the next in-frame stop-codon appears in the downstream vector sequence. The advantage of generating test-libraries without stop codons, e.g. as Gateway entry vectors like the VZV ORFeome that was used in this study, is the more versatile use that can be made of those libraries, because they can be cloned into destination vectors for N- and C-terminal fusions, like for example pGBKT7g and pGBKCg.

4.2.1.4 Sampling sensitivity

A source for false-negative interactions in HTP-screens is the rate of detectable interactions, which can be found in a single independent run of a screen. Repetitive screens can produce partially overlapping data until a level of saturation is reached. The ratio of interactions detected at saturation of an assay is equivalent to the assay sensitivity. However, this holds true for library screens, while it is assumed that

pairwise mating experiments operate at or near full sampling sensitivity since such experiments overcome losses due to pooling, limited selection of positives and sequencing. Moreover, one-on-one screens are performed as duplicates or even quadruplicates.

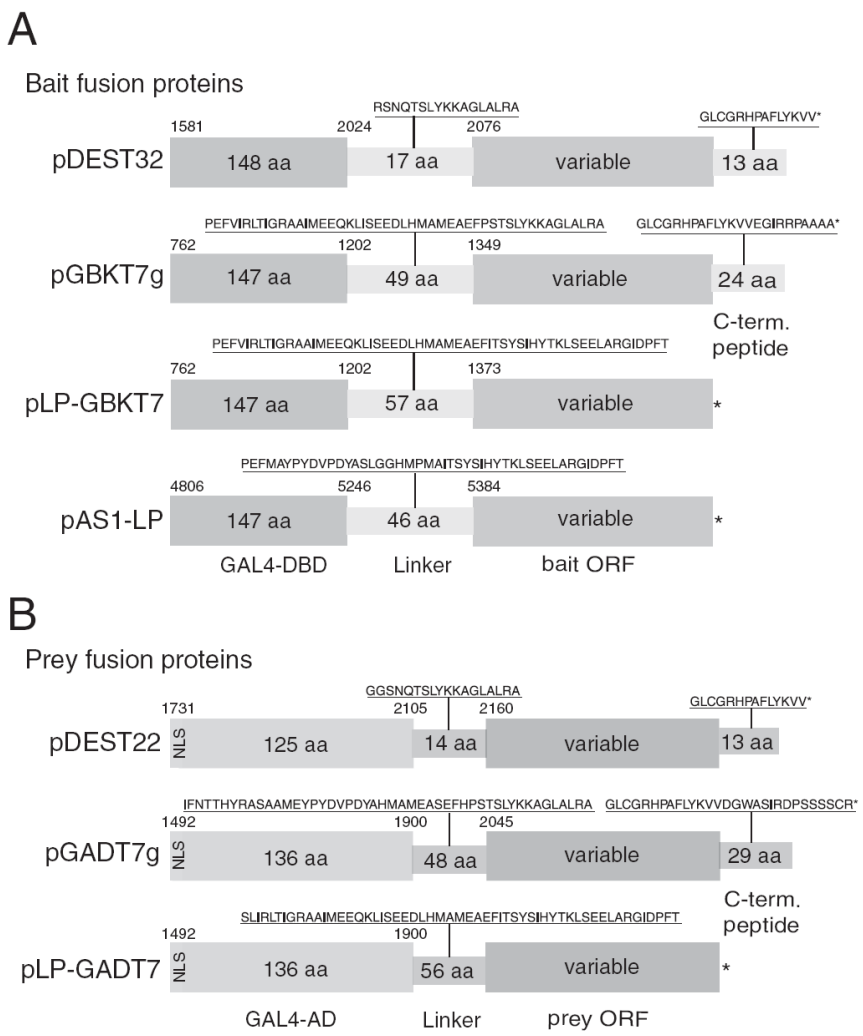


Figure 50: Comparison of the design of different Y2H vectors.
A) Bait and **B**) prey-fusions of different Y2H vectors were compared by Rajagopala and coworkers (Rajagopala et al., 2009), including the vectors for N-terminal fusions used in this study, pGBKT7g or pGBGT7g which are identical except of their auxotrophic marker and pGADT7g. All compared bait and prey vectors are based on GAL4-DBD and -AD. The amino acid sequence of the linker region between Gal4 AD or DBD and the prey or bait ORF is indicated above the respective box. **A**) Bait vectors differ in the length of the linker sequence between the DBD and the test insert ("variable" sequence). Secondly, additional C-terminal fused peptides appear, depending on the nature of the inserts. If the insert contains a stop codon, the C-terminus reflects the native situation. Test constructs without stop codon are on the other hand more versatile as they can be cloned into destination vectors for N- and C-terminal fusions. **B**) Prey fusions mainly in the linker sequences between the AD and the prey ORF. Additionally, pDEST22 has a 6 AA truncated GAL4-AD. NLS: nuclear localization signal. A C-terminal peptide also depends on the existence of a stop codon.

4.2.1.5 Steric constraints

Besides the above discussed reasons, false-negative interactions may be traced back to steric hindrance by the fusion tags, preventing physical interaction by covering interaction sites or preventing subsequent transcriptional activation. So far, no high-throughput accessible set of vectors for C-terminal fusions of GAL4 AD and DBD have been described, even though it was shown that the topology of the fusion site is critical when selected protein pairs were tested (Brown and MacGillivray, 1997). In addition to that, I developed a system of compatible bait and prey vectors, which allow crosswise combination of N- and C-terminal fusion tags.

4.2.1.6 Comparison of tag-topology combinations

- **NN-topology in this study and Braun data**

Regarding the results from the combinatorial screens of the hsPRS-v1 (see Figure 35) the NN-conformation has found, as described above, 52 % (48 out of 92 interactions). 34 interactions were detected by Braun et al., 25 thereof could be also detected by pGBGT7g/pGADT7g that is coverage of 74 % of the Braun data and an overlap of 52 % with the NN-dataset (25/48 interactions).

- **NN and CC tag-permutation**

Almost the same number was found in the NN and CC bait/prey combination: 48 and 47 interactions, respectively. The overlap between both combinations is 26 common interactions. The overlap between the CC- and NN-dataset is 55 %.

- **NN and NC combination**

The assay sensitivity of the NC-combination was lower with 42 of 92 interactions detected. Those are 46 %. 28 of the interactions are in common, that's an overlap of 58 %.

- **NN and CN combination**

The poorest assay sensitivity was observed using the CN-combination pGBKCg bait and pGADT7g prey. 20 of 29 CN-interactions are in common with the NN-dataset. There is also a relatively small overlap with the NN-data, which is 42 %.

4.2.1.7 Relative overlaps between topologies

Interestingly, the relative overlaps of PRS-interactions are the highest among the both NN-vector systems, underlining the impact of steric factors on the detectability of protein-protein interactions (see Table 17).

	NN (Braun)	CC	NC	CN
PRS-positive interactions	34	47	42	29
overlap with NN (this study)	25	26	28	20
ratio	74 %	55 %	67 %	69 %

Table 17: Relative overlaps of PRS-interactions with the NN-topology.
The overlap between different vector systems and the classical system used here varies between 20 and 28 interactions. The relative overlaps between the NN-topologies are the highest (74 %), even though they were assayed by different Laboratories.

4.2.1.8 Assay sensitivity of combined screens

The major benefit that is achieved by combined tag-topology screens is the high assay sensitivity that can be obtained. Taken together, 74 of the 92 PRS interactions could be detected, which reflects an assay sensitivity of 80 %. The classical NN-topology detected 48 (52 %) of the 92 PRS interactions, additional 26 interactions (28 %) were detected involving at least one C-terminal fusion tag, which reflects an increase of interaction data of 65 %.

4.2.1.9 Additional assays

Additional assays like LUMIER or YFP-PCA usually detect different, partially overlapping subsets of PPIs of the same organism or test space (Rajagopala et al., 2009; Rual et al., 2005). Additionally, verification of Y2H interactions by an additional assay, like e.g. Co-Immunoprecipitation, is widely regarded as a tool to confirm the authenticity of Y2H interactions. Especially in the case of large scale studies, they are time-consuming, expensive, and require additional reagents and protocols. In addition, a large fraction of bona fide interactions may be suitable for one particular assay but not for another and thus confirmatory experiments may still miss up to 80 % of all interactions. When the hsPRS-v1 was assayed by four different methods by Braun and colleagues (Braun et al., 2009), similar to combinatorial Y2H-screening, the coverage of PRS-interactions was extended, to 55 of 92 interactions, or 60 %

(see Figure 51). Hence, the coverage was still 20 % lower than from combinatorial Y2H screening.

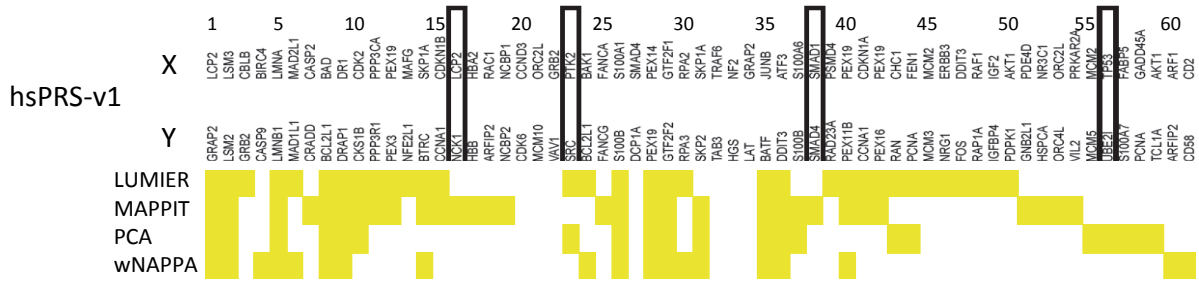


Figure 51: PRS interactions from additional assays.
55 of 92 hsPRS-v1 interactions were detected by four independent assays. LUMIER, MAPPIT (see Figure 52A) and PCA could detect at least one phosphorylation-dependent interaction (black edges), wNAPPA (Figure 52B) did not. Figure modified from Braun et al., 2009.

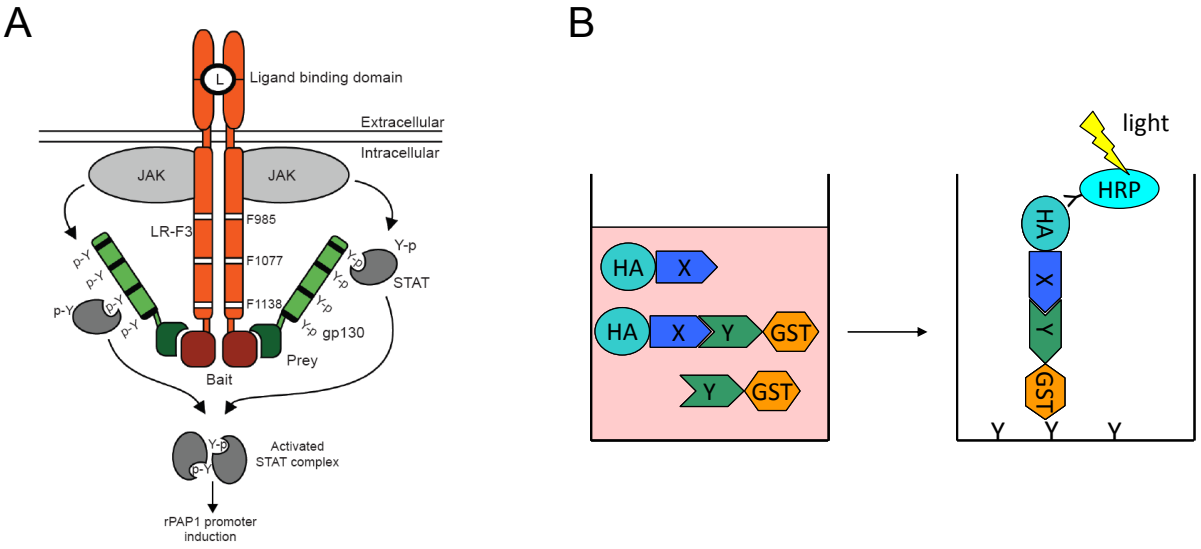


Figure 52: MAPPIT and wNAPPA principle.
(**A**) The mammalian protein-protein interaction trap (MAPPIT) is based on reconstitution of a signaling cascade, activating Stat3 dependent transcription. A bait protein is fused to a hybrid erythropoietin-leptin receptor and the prey is fused to gp130. Upon stimulation with erythropoietin, Janus kinases (JAKs) transphosphorylate each other, and if bait and prey interact, the activated JAKs will phosphorylate gp130, which in turn recruits and subsequently activates STAT3, which then activates transcription of a reporter. Source: Eyckerman et al., 2001. (**B**) Well - nucleic acid programmable protein array (wNAPPA). Plasmids encoding GST-bait and HA-prey fusions are in vitro translated, e.g. using reticulocyte lysate (purple shading). Subsequently, the bait-GST is captured on the bottom of a 96-well plate coated with GST antibody. An interaction recruits the HA-prey fusion which can be immunologically detected (Braun et al., 2009).

4.2.2 False-positives

Yeast two-hybrid interactions have often been considered as unreliable, generating many false-positive and false-negative results (Edwards et al., 2002). False-positives can be suspected because interactions appear implausible to the observer, especially if they lack independent confirmation. However, it is difficult to exclude that proteins which are tested randomly do not interact *in vivo*. But it is anyway an indicator for a running system to evaluate its susceptibility for false-positive interactions.

4.2.2.1 hsRRS-v1 screening

Pairwise testing of hsPRS-v1 interactions detected, taken together all non-redundant interactions of all four screens 16 pairwise interactions, which equals 17 % of the whole RRS. The putative false-positive rates of single permutations were 6.5 % in the NN-, 8.7 % in the CC, 7.1 % in the NC combination. In the CN permutation the rate was somewhat lower with 4.4 %, probably also reflecting its overall weaker performance in the PRS-assays. Taken together, a standard screen contains approximately around seven percent false-positive interactions. Essentially, the rate of false-positives can be minimized to a minimum level, down to zero, see also Table 14 with no RRS-interaction detected using 10 mM 3-AT in the readout medium. But stringency and the resulting loss of sensitivity goes at the expense of loosing true-positive interactions as well (Braun et al., 2009).

Testing the hsPRS-v1 interactions by combinatorial screening was performed under largely comparable conditions as the VZV screen before, pretests were performed with the bait-strains mated against the empty pGADT7g vector and subsequently tested for self-activation. Assays were subsequently scored on minimum inhibitory concentration of 3-AT.

One essential outcome of the RRS-evaluation is that no tag combination produces an elevated number of false-positive interactions compared to the others. What has to be taken in account as well, regarding the assay layout compared to a standard array-based screen is that an array of different bait proteins is screened on one array at a time. This has the disadvantage that the assessment of the specificity by means of the array appearance is lost. So the only evidence of specificity goes back to the self-activation assay that is in addition performed in an independent assay step. In

many cases, the minimum inhibitory concentration of a bait construct stays undecided and two or more concentrations are screened for later evaluation of the results with the help of comparable arrays, see Figure 53 for an example. This strategy helps avoiding false-negative interactions, which in fact is not accounted for in the hsRRS-v1 screen.

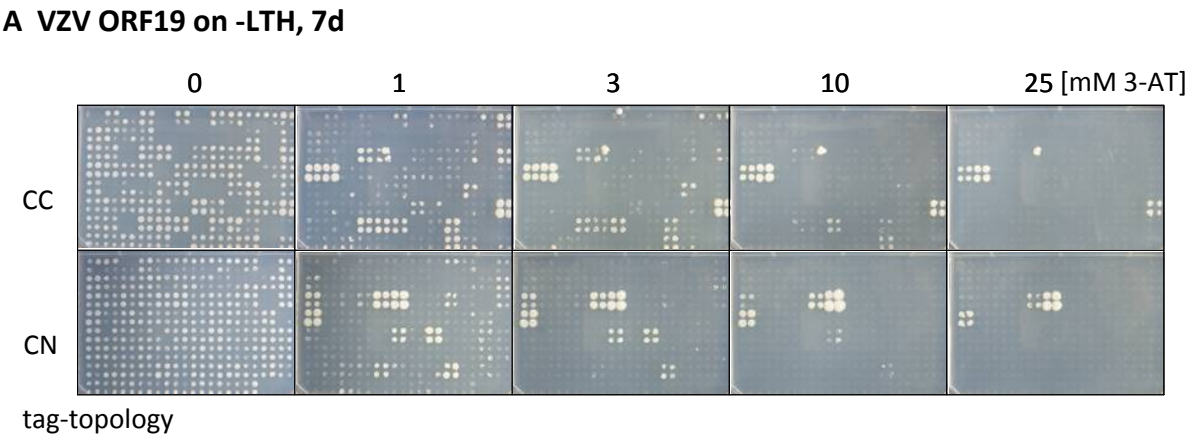


Figure 53: VZV ORF19 in a screen the autoactivation pretest.
A) VZV ORF19 was screened as bait in pGBKCg against the VZV prey array (for the array layout see Table 2). In the screen, there is an unspecific background visible at 1 mM 3-AT that vanishes with rising 3-AT concentration. The unspecific background is of equal strength with C-terminally tagged preys (CC, upper half) and N-terminally tagged preys (CN, lower half) **B**) In the autoactivation pretest of ORF19 in pGBKCg mated against pGADT7g no growth was visible at 1 mM 3-AT. In this case the subsequent scoring of interactions on 3 mM 3-AT reduces false-negative interactions.

4.2.2.2 Additional assays

Like the positive reference set of binary human interactions, the hsRRS-v1 was tested by Braun and colleagues (Braun et al., 2009) using four different methods for PPI detection (Figure 54). Each method detected two to four putative false-positive interactions. Combined, the false-negative rate is 11 % (10 of 92 random-interactions

detected) which is only six percent lower than the calculated FP-rate from the combinatorial Y2H screens. Taking in account the much lower expenses of one single assay and the potential of reducing false-positives from Y2H by array-based screening and elevated assay stringency the extra efforts seem to be not reasonable, especially for high-throughput screening.

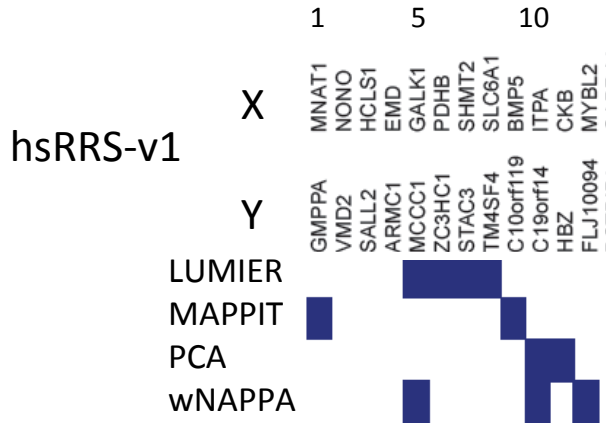


Figure 54: RRS interactions from additional assays.
The least RRS-interactions, two each, were detected by MAPPIT and PCA. One more was detected by wNAPPA and four by LUMIER. Taken together, nine non-redundant putative false-positive interactions were detected. Figure modified from Braun et al., 2009.

4.2.2.3 Quality assessment by LuMPIS

An additional evidence for good Y2H data quality I observed in a collaboration study with the Max-von-Pettenkofer Institut at the LMU Munich. The task was to generate an intraviral Hepatitis-E-Virus PPI network. Cloning and subsequent verification by LuMPIS-assays (Vizoso Pinto et al., 2009), which is a derivative of the LUMIER system, were performed by Andreas Osterman and Dr. Maria Vizoso-Pinto in the research group of Dr. Armin Baiker. I performed an intraviral Y2H screen with permutated fusion tags. Strikingly, the Yeast two-hybrid interactions could be verified to the largest part (88 %). The second outcome of this study is, that the interactions that could not be detected in the LuMPIS system, which uses N-terminal fusions of eGFP-Luciferase and MBP as bait and prey constructs (Figure 55), were exclusively detected involving C-terminal fusion constructs (Osterman et al., Manuscript in preparation). This is an additional evidence for sterical hindrance of PPIs by fusion tags.

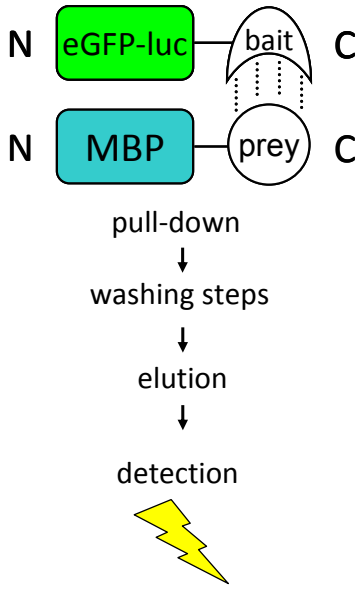


Figure 55: The LuMPIS system.
The eGFP-luciferase tagged bait mixed with a MBP-tagged prey expressed in mammalian cells is detected in affinity purified samples as light emission. It was designed to identify PPIs of organisms with low GC-content, like VZV. The eGFP (enhanced green fluorescent protein) and MBP (maltose-binding protein) fusion tags have a high GC-content which improves recombinant expression in mammalian cells.

4.2.3 Conclusions

Apart from the orientation of fusion tags, various variations of the Y2H system are possible, e.g. copy number, reporter yeast strains, different reporters etc. (Rajagopala et al., 2009). The parental vectors and the two new C-terminal vectors described here are identical in most ways, including their origins of replication (and thus their copy number), their promoters (and thus protein expression levels), as well as the yeast strains in which they were expressed. In addition, the experimental conditions of our Y2H assays were identical (except for adapting the needed 3-AT concentrations). This allows me to conclude that differences in PPIs result from sterical constraints caused by the location of the DBD and AD fusion tags and the associated linker sequences. Considering these findings, other Y2H vector systems including commercial available vectors would benefit from the permutations described here. They should be used routinely in large-scale screens, not least because they provide an intrinsic quality score for the obtained Y2H interactions by providing reproducibility without additional assays.

4.3 VZV ORF25 and the terminase complex

The *Herpesviridae* encode a family of seven orthologous proteins, which cleave concatemeric viral DNA into separate genome units and insert the DNA into preformed capsids in the nucleus of an infected host cell. These DNA-packaging proteins are encoded by the VZV ORFs 25, 26, 30, 34, 43, 45/42 and 54 (Visalli et al., 2007) and orthologs thereof in other herpesviruses. The identification of encapsidation-specific antiviral inhibitors for HSV, HCMV, and VZV suggests that viral DNA encapsidation is a promising antiviral target for herpesviruses (Biron, 2006; Bogner, 2002; Buerger et al., 2001; Di Grandi et al., 2004; Reefschlaeger et al., 2001; Underwood et al., 1998; van Zeijl et al., 2000; Visalli et al., 2003), especially as this is a virus-specific mechanism, which minimizes the risk of side effects of new therapeutic compounds in preclinical development and subsequent clinical trials. New pharmaceuticals are required as the classical antiviral therapy which is mainly based on nucleoside analogs is facing the problem of resistance development. Moreover, resistance against one kind of nucleoside analog is often accompanied by resistance to other nucleoside analog derivatives, so called cross-resistance, based on mutations in the viral thymidine kinase and DNA-polymerase genes (Sauerbrei et al., 2010). That is why a better characterization of the interactions between DNA encapsidation proteins can be used for development of new antiviral compounds based on the prevention of viral DNA packaging.

4.3.1 Terminase complex as drug target - actual state of affairs

So far, only few drugs targeting the human cytomegalovirus viral terminase were or are being developed. BAY-384766 (Tomeglovir®, Bayer AG) and GW-275175X (175X; University of Michigan/GlaxoSmithKline), both had entered Phase I trials but developments were discontinued in November 2000 for unknown reasons. Maribavir (MBV; 1263W94; Camvia®; Viropharma) is a promising agent undergoing Phase III trials. It is a potent member of a new class of drugs called benzimidazole ribosides (Hwang et al., 2009; Lischka and Zimmermann, 2008). A fourth substance, called AIC246 (AiCuris) is undergoing clinical development (Lischka et al., 2010).

4.3.2 Protein interactions among DNA encapsidation proteins

I assembled all available interaction data of the DNA encapsidation proteins to generate a complete-as-possible network of this crucial step in virion morphogenesis. The data was assembled from the combinatorial, array-based Y2H screens of VZV (Stellberger et al., 2010; Uetz et al., 2006), from orthologous prediction from four additional herpesviruses (Fossum et al., 2009) as well as from a systematical re-screen of the seven VZV proteins with re-cloned ORFs, to maximize the quality of the test constructs. Figure 56 gives an overview of the DNA-encapsidation network.

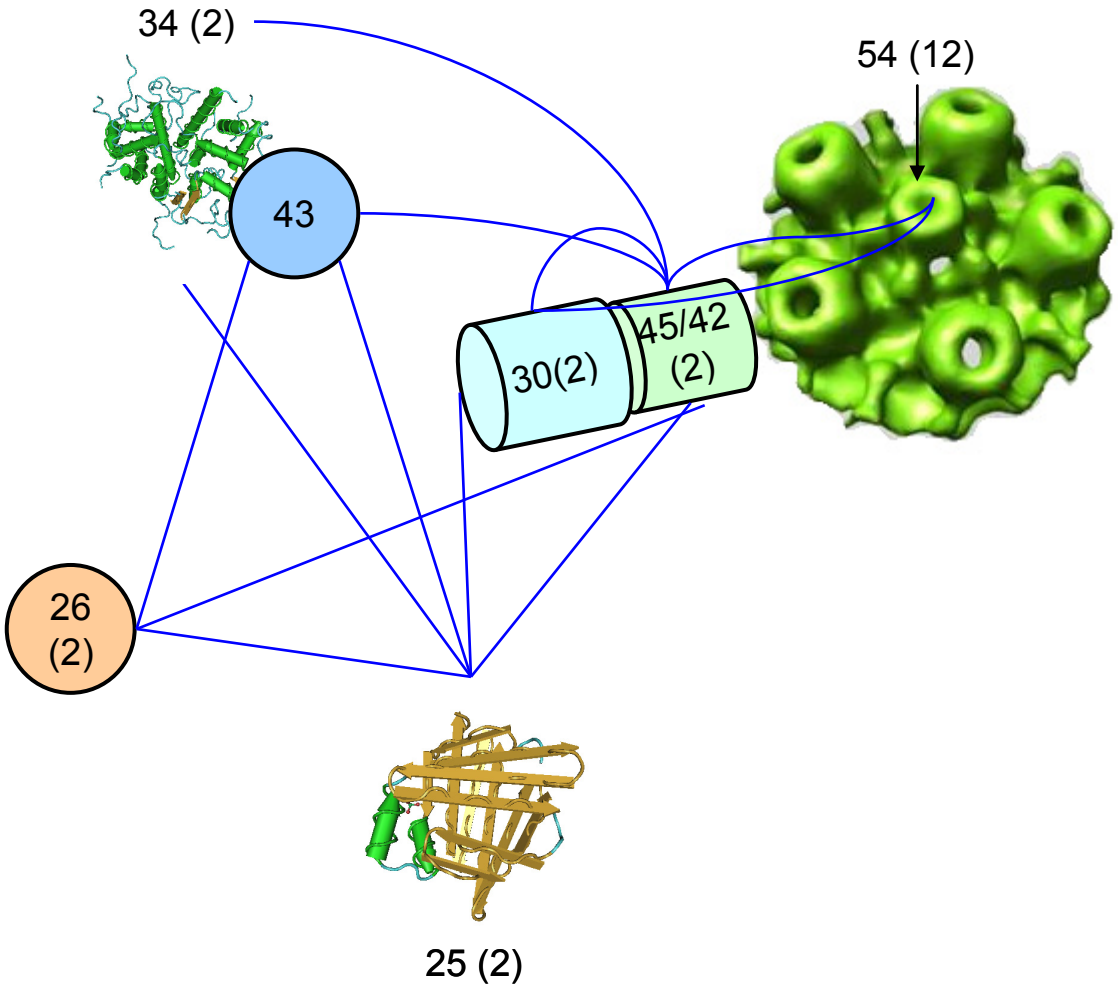


Figure 56: Interaction network of putative terminase complex subunits-
The interaction data from Figure 40 was used to model an interaction network of the VZV proteins involved in DNA-packaging. ORFs are indicated by the respective number, self interactions are indicated by attached numbers in brackets. Structures of ORF25 and ORF34 were derived based on homology from the NCBI Molecular Modeling Database, MMDB (Wang et al., 2007) and visualized using the Cn3D viewer (Wang et al., 2000). The portal vertex image (ORF54 dodecamer) is adapted from Cardone and colleagues (Cardone et al., 2007), showing a detail of the closely related HSV-1 virion. Toroidal structures of the large and small terminase subunit (ORF30 and ORF45/42, respectively) are represented by cylindrical nodes, according to the observations of Scheffzzik and Savva on the HCMV orthologs UL56 and UL89 (Savva et al., 2004; Scheffczik et al., 2002).

4.3.3 Promiscuity of ORF25

The VZV ORF25 encoded protein belongs to the Herpes UL33 protein superfamily (pfam03581) and is highly conserved throughout the *Herpesviridae* (Marchler-Bauer et al., 2007). This very promiscuous protein interacts with almost two thirds of the VZV encoded proteins in the Y2H system. Before the combinatorial Y2H screening of VZV was performed, 33 PPIs were detected with in the NN-topology (Uetz et al., 2006), which is almost half of the VZV encoded proteins. Interestingly, 18 orthologous interactions (56 %) were confirmed in intraviral Y2H screens of HSV-1, EBV, MCMV (Fossum et al., 2009). Additionally, the interaction of ORF25 with the terminase subunit ORF30 and the orthologous interaction in HSV-1, are meanwhile described in the literature (Jacobson et al., 2006; Visalli et al., 2009). This suggests that the interactions of ORF25 protein are mostly specific and not caused by an unspecific stickiness of the VZV ORF25 Yeast two-hybrid constructs used in the screenings. The combinatorial Y2H screens in this study revealed 13 additional interactions involving the newly designed C-terminally tagged Yeast two-hybrid constructs. These are including seven interologous interactions in HSV-1, EBV and MCMV (54 %), which is pointing out an equal data quality derived by the C-terminal fusions.

4.3.4 Role of ORF25

Even though ORF25 protein interacts with a large variety of viral and also host cell proteins (Vizoso Pinto et al., Manuscript in preparation), it was suggested that in HSV-1 the terminase complex is a trimeric complex of the VZV orthologs of ORF45/42, ORF30 and ORF25 (Yang et al., 2007). Recently, this was relativized by Visalli and coworkers. It is now hypothesized that the Herpes UL33 superfamily homologs are not plainly encapsidation proteins. Beyond that they play a role in the assembly of viral proteins and/or in the optimization of viral protein complexes (Visalli et al., 2009). This theory is supported by several studies that include:

- The optimization of the terminase complex of HSV-1 UL33 in HSV infected cells (Yang and Baines, 2006).
- The high number of interactions of VZV ORF25 and its orthologs (Fossum et al., 2009; Stellberger et al., 2010; Uetz et al., 2006)

- The translocation of VZV, HSV-2 and KSHV UL33 orthologs into the nucleus. (Sander et al., 2008; Vizoso Pinto et al., Manuscript in preparation; Yamauchi et al., 2001)
- The interactions of VZV ORF25 protein with most of the DNA encapsidation proteins (Figure 56).

So far, this data suggests the model that the UL33 proteins bring together viral proteins in order to form functional complexes and that the putative DNA encapsidation complex of VZV consists of at least three different proteins analogous to that reported for HSV-1, ORF25, ORF45/42 and ORF30 (Visalli et al., 2007). However, new results suggest that ORF25 acts as a molecular chaperone for viral proteins as it shows characteristic behavior of heat shock proteins like translocation into the nucleus upon cellular stress (Vizoso Pinto et al., Manuscript in preparation). This provides an explanation for the promiscuity of the protein, as well as for the involvement into the DNA-encapsidation complex and the stabilizing effect on protein complexes.

4.3.5 Mapping of the ORF25 homomerization interface

Yeast two-hybrid analysis and peptide scans of the ORF25-ORF25 interaction showed that this involves segments of N- and C-terminal constructs (ORF25N and ORF25C). Moreover, on the peptide array there were two distinct areas on the C-terminal domain identified, which are evolutionary higher conserved based on sequence alignment of the UL33 superfamily in HHV-1 to HHV-8. Based on these findings, I assume that the interaction interface is a discontinuous interaction epitope, with a conserved part made up of central- and C-terminal portions. The N-terminal fraction was lost during herpesviral evolution (see Figure 57), likely due to the fact that it is dispensable for the interaction.

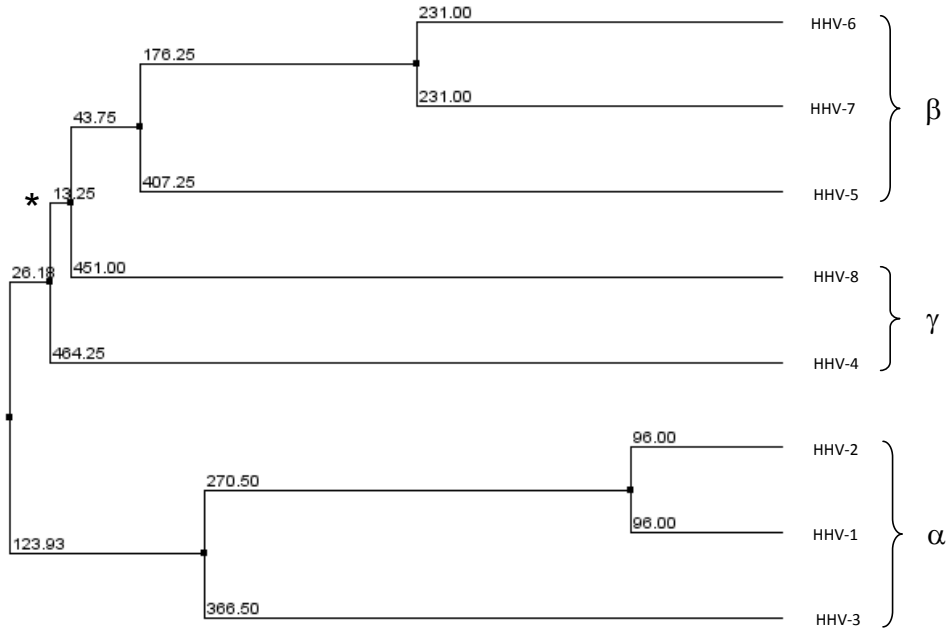


Figure 57: Phylogenetic tree of the UL33 protein superfamily.
The average distance was calculated based on the amino acid sequence alignment of UL33 orthologs in HHV-1 to HHV-8 shown in Figure 43, using the BLOSUM62 algorithm. The star marks the evolutionary event that separated the *Gamma*- and *Betaherpesvirinae*, which led to the loss of amino acids at the N-terminus that is part of the homomerization interface of ORF25 protein and probably the orthologs in HHV-1, -2 and -4.

4.4 Outlook

4.4.1 Combinatorial Y2H screening

Ever since the publication of the novel Yeast two-hybrid vectors pGADCg and pGBKCg, the system has attracted attention in research groups in many parts of the world. To date, I have provided them to laboratories in the United States, China, Hungary, France, Great Britain Switzerland and Germany. I am pleased to be able to provide a tool that has the potential to support protein-protein interaction studies, performed either in large-scale or small-scale, in all fields of basic- or biomedical research.

I screened the intraviral PPIs of Hepatitis-E-Virus as first *de-novo* determined network by combinatorial screening of four tag-permutations and gained very good data quality. Encouraged by this study, I will start in the near future intraviral screening of Human Coronavirus NL63 (HCoV-NL63), as starting point for a postdoctoral project.

4.4.2 Drug design based on PPI blocking

Protein-protein interactions play a central role in biochemical reactions including processes of pathogenesis. Proteins often work in complexes of several macromolecules and small ligands. The structural and functional description of protein-protein interactions is very important for basic and applied research. The interface areas of protein complexes have unique structure and properties, so PPI represent prospective targets for a new generation of drugs.

One of the key targets of PPI inhibitors are oligomeric enzymes, which is of increasing interest within the last years, e.g. the HIV-1 protease (HIVp), which functions as a homodimer or bacterial L-asparaginase (homo-tetramer) (Ivanov et al., 2007).

A prerequisite for drug design based on prevention of essential protein-protein interactions is the identification of interaction sites, like it was determined in this study for VZV ORF25 homomerization. The second basic requirement is the revelation of the three dimensional structure of the target proteins. Supportingly, the field of structural genomics has emerged as one of the -omics (like e.g. genomics, transcriptomics and proteomics) disciplines more than a decade ago, and large scale programs have been launched across the world, for development and application of methods for high-throughput structural biology. As a result, the growing number of high resolution structures of protein drug targets is expected to have a dominant impact on future drug discovery programs (Weigelt, 2010).

Without knowing the structure of VZV ORF25 protein, I started to develop a high-throughput method to screen for antiviral compounds in small-compound libraries, based on the reverse Yeast two-hybrid (rY2H) system. The rY2H allows to screen for mutations which abolish PPIs using the URA3 reporter gene in combination with 5-FOA (Vidal et al., 1996), and thus should have the potential to identify compounds which block protein-protein interactions between promising drug-targets.

5 References

Abbotts, A. P., Preston, V. G., Hughes, M., Patel, A. H. and Stow, N. D. (2000). Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. *J Gen Virol* **81**, 2999-3009.

Akula, S. M., Pramod, N. P., Wang, F. Z. and Chandran, B. (2002). Integrin alpha3beta1 (CD 49c/29) is a cellular receptor for Kaposi's sarcoma-associated herpesvirus (KSHV/HHV-8) entry into the target cells. *Cell* **108**, 407-19.

Alanis, A. J. (2005). Resistance to antibiotics: are we in the post-antibiotic era? *Arch Med Res* **36**, 697-705.

Albers, M., Kranz, H., Kober, I., Kaiser, C., Klink, M., Suckow, J., Kern, R. and Koegl, M. (2005). Automated yeast two-hybrid screening for nuclear receptor-interacting proteins. *Mol Cell Proteomics* **4**, 205-13.

Ammerer, G. (1983). Expression of genes in yeast using the ADCI promoter. *Methods Enzymol* **101**, 192-201.

Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., Feuermann, M., Ghanbarian, A. T., Kerrien, S., Khadake, J. et al. (2010). The IntAct molecular interaction database in 2010. *Nucleic Acids Res* **38**, D525-31.

Arrell, D. K. and Terzic, A. (2010). Network systems biology for drug discovery. *Clin Pharmacol Ther* **88**, 120-5.

Assenov, Y., Ramirez, F., Schelhorn, S. E., Lengauer, T. and Albrecht, M. (2008). Computing topological parameters of biological networks. *Bioinformatics* **24**, 282-4.

Auerbach, D., Thaminy, S., Hottiger, M. O. and Stagljar, I. (2002). The post-genomic era of interactive proteomics: facts and perspectives. *Proteomics* **2**, 611-23.

Bachmair, A., Finley, D. and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179-86.

Bader, G. D., Betel, D. and Hogue, C. W. (2003). BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res* **31**, 248-50.

Bader, G. D. and Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**, 2.

Barabasi, A. L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5**, 101-13.

Barrios-Rodiles, M., Brown, K. R., Ozdamar, B., Bose, R., Liu, Z., Donovan, R. S., Shinjo, F., Liu, Y., Dembowy, J., Taylor, I. W. et al. (2005). High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* **307**, 1621-5.

Bartel, P. L., Roecklein, J. A., SenGupta, D. and Fields, S. (1996). A protein linkage map of Escherichia coli bacteriophage T7. *Nat Genet* **12**, 72-7.

Batterson, W., Furlong, D. and Roizman, B. (1983). Molecular genetics of herpes simplex virus. VIII. further characterization of a temperature-sensitive mutant defective in release of viral DNA and in other stages of the viral reproductive cycle. *J Virol* **45**, 397-407.

Beard, P. M., Taus, N. S. and Baines, J. D. (2002). DNA cleavage and packaging proteins encoded by genes U(L)28, U(L)15, and U(L)33 of herpes simplex virus type 1 form a complex in infected cells. *J Virol* **76**, 4785-91.

Beilstein, F., Higgs, M. R. and Stow, N. D. (2009). Mutational analysis of the herpes simplex virus type 1 DNA packaging protein UL33. *J Virol* **83**, 8938-45.

Beranger, F., Aresta, S., de Gunzburg, J. and Camonis, J. (1997). Getting more from the two-hybrid system: N-terminal fusions to LexA are efficient and sensitive baits for two-hybrid studies. *Nucleic Acids Res* **25**, 2035-6.

Biron, K. K. (2006). Antiviral drugs for cytomegalovirus diseases. *Antiviral Res* **71**, 154-63.
Bogner, E. (2002). Human cytomegalovirus terminase as a target for antiviral chemotherapy. *Rev Med Virol* **12**, 115-27.
Bonifazi, P., Goldin, M., Picardo, M. A., Jorquera, I., Cattani, A., Bianconi, G., Represa, A., Ben-Ari, Y. and Cossart, R. (2009). GABAergic hub neurons orchestrate synchrony in developing hippocampal networks. *Science* **326**, 1419-24.
Braun, P., Tasan, M., Dreze, M., Barrios-Rodiles, M., Lemmens, I., Yu, H., Sahalie, J. M., Murray, R. R., Roncari, L., de Smet, A. S. et al. (2009). An experimentally derived confidence score for binary protein-protein interactions. *Nat Methods* **6**, 91-7.
Breitkreutz, B. J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D. H., Bahler, J., Wood, V. et al. (2008). The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res* **36**, D637-40.
Brown, M. A. and MacGillivray, R. T. (1997). Vectors for expressing proteins at the amino-terminus of an activation domain for use in the yeast two-hybrid system. *Anal Biochem* **247**, 451-2.
Bruckner, A., Polge, C., Lentze, N., Auerbach, D. and Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci* **10**, 2763-88.
Buckmaster, A. E., Scott, S. D., Sanderson, M. J., Boursnell, M. E., Ross, N. L. and Binns, M. M. (1988). Gene sequence and mapping data from Marek's disease virus and herpesvirus of turkeys: implications for herpesvirus classification. *J Gen Virol* **69 (Pt 8)**, 2033-42.
Buerger, I., Reefschlaeger, J., Bender, W., Eckenberg, P., Popp, A., Weber, O., Graeper, S., Klenk, H. D., Ruebsamen-Waigmann, H. and Hallenberger, S. (2001). A novel nonnucleoside inhibitor specifically targets cytomegalovirus DNA maturation via the UL89 and UL56 gene products. *J Virol* **75**, 9077-86.
Cai, X., Lu, S., Zhang, Z., Gonzalez, C. M., Damania, B. and Cullen, B. R. (2005). Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. *Proc Natl Acad Sci U S A* **102**, 5570-5.
Calderwood, M. A., Venkatesan, K., Xing, L., Chase, M. R., Vazquez, A., Holthaus, A. M., Ewence, A. E., Li, N., Hirozane-Kishikawa, T., Hill, D. E. et al. (2007). Epstein-Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci U S A* **104**, 7606-11.
Cardone, G., Winkler, D. C., Trus, B. L., Cheng, N., Heuser, J. E., Newcomb, W. W., Brown, J. C. and Steven, A. C. (2007). Visualization of the herpes simplex virus portal in situ by cryo-electron tomography. *Virology* **361**, 426-34.
Cebrian, J., Kaschka-Dierich, C., Berthelot, N. and Sheldrick, P. (1982). Inverted repeat nucleotide sequences in the genomes of Marek disease virus and the herpesvirus of the turkey. *Proc Natl Acad Sci U S A* **79**, 555-8.
Ceol, A., Chatr Aryamontri, A., Licata, L., Peluso, D., Briganti, L., Perfetto, L., Castagnoli, L. and Cesareni, G. (2010). MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res* **38**, D532-9.
Chautard, E., Ballut, L., Thierry-Mieg, N. and Ricard-Blum, S. (2009a). MatrixDB, a database focused on extracellular protein-protein and protein-carbohydrate interactions. *Bioinformatics* **25**, 690-1.
Chautard, E., Thierry-Mieg, N. and Ricard-Blum, S. (2009b). Interaction networks: from protein functions to drug discovery. A review. *Pathol Biol (Paris)* **57**, 324-33.
Che, X., Berarducci, B., Sommer, M., Ruyechan, W. T. and Arvin, A. M. (2007). The ubiquitous cellular transcriptional factor USF targets the varicella-zoster virus

open reading frame 10 promoter and determines virulence in human skin xenografts in SCIDhu mice in vivo. *J Virol* **81**, 3229-39.
Che, X., Zerboni, L., Sommer, M. H. and Arvin, A. M. (2006). Varicella-zoster virus open reading frame 10 is a virulence determinant in skin cells but not in T cells in vivo. *J Virol* **80**, 3238-48.
Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A., 3rd, Kouzarides, T., Martignetti, J. A. et al. (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr Top Microbiol Immunol* **154**, 125-69.
Chen, J., Ueda, K., Sakakibara, S., Okuno, T., Parravicini, C., Corbellino, M. and Yamanishi, K. (2001). Activation of latent Kaposi's sarcoma-associated herpesvirus by demethylation of the promoter of the lytic transactivator. *Proc Natl Acad Sci U S A* **98**, 4119-24.
Chen, Y., Rajagopala, S. V., Stellberger, T. and Uetz, P. (2010). Exhaustive benchmarking of the yeast two-hybrid system. *Nat Methods* **7**, 667 - 668.
Clarke, P., Beer, T., Cohrs, R. and Gilden, D. H. (1995). Configuration of latent varicella-zoster virus DNA. *J Virol* **69**, 8151-4.
Coen, D. M. and Schaffer, P. A. (2003). Antiherpesvirus drugs: a promising spectrum of new drugs and drug targets. *Nat Rev Drug Discov* **2**, 278-88.
Cohen, J. I. and Seidel, K. (1994). Varicella-zoster virus (VZV) open reading frame 10 protein, the homolog of the essential herpes simplex virus protein VP16, is dispensable for VZV replication in vitro. *J Virol* **68**, 7850-8.
Cohrs, R. J., Barbour, M. B., Mahalingam, R., Wellish, M. and Gilden, D. H. (1995). Varicella-zoster virus (VZV) transcription during latency in human ganglia: prevalence of VZV gene 21 transcripts in latently infected human ganglia. *J Virol* **69**, 2674-8.
Cohrs, R. J. and Gilden, D. H. (2003). Varicella zoster virus transcription in latently-infected human ganglia. *Anticancer Res* **23**, 2063-9.
Cohrs, R. J. and Gilden, D. H. (2007). Prevalence and abundance of latently transcribed varicella-zoster virus genes in human ganglia. *J Virol* **81**, 2950-6.
Cohrs, R. J., Srock, K., Barbour, M. B., Owens, G., Mahalingam, R., Devlin, M. E., Wellish, M. and Gilden, D. H. (1994). Varicella-zoster virus (VZV) transcription during latency in human ganglia: construction of a cDNA library from latently infected human trigeminal ganglia and detection of a VZV transcript. *J Virol* **68**, 7900-8.
Congreve, M., Chessari, G., Tisi, D. and Woodhead, A. J. (2008). Recent developments in fragment-based drug discovery. *J Med Chem* **51**, 3661-80.
Cox, E., Reddy, S., Iofin, I. and Cohen, J. I. (1998). Varicella-zoster virus ORF57, unlike its pseudorabies virus UL3.5 homolog, is dispensable for viral replication in cell culture. *Virology* **250**, 205-9.
Daeschler, E. B., Shubin, N. H. and Jenkins, F. A., Jr. (2006). A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature* **440**, 757-63.
Davison, A. (2004). Compendium of Human Herpesvirus gene names, Reno.
Davison, A. J. and Scott, J. E. (1986). The complete DNA sequence of varicella-zoster virus. *J Gen Virol* **67 (Pt 9)**, 1759-816.
de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M. S., Aublin-Gex, A., Agaugue, S., Meiffren, G., Pradezynski, F., Faria, B. F., Chantier, T. et al. (2008). Hepatitis C virus infection protein network. *Mol Syst Biol* **4**, 230.
Di Grandi, M. J., Curran, K. J., Feigelson, G., Prashad, A., Ross, A. A., Visalli, R., Fairhurst, J., Feld, B. and Bloom, J. D. (2004). Thiourea inhibitors of

herpesviruses. Part 3: Inhibitors of varicella zoster virus. *Bioorg Med Chem Lett* **14**, 4157-60.
Diestel, R. (2005). Graph theory. , (ed. Heidelberg: Springer-Verlag.
Dohner, K., Wolfstein, A., Prank, U., Echeverri, C., Dujardin, D., Vallee, R. and Sodeik, B. (2002). Function of dynein and dynactin in herpes simplex virus capsid transport. *Mol Biol Cell* **13**, 2795-809.
Drees, B. L. (1999). Progress and variations in two-hybrid and three-hybrid technologies. *Curr Opin Chem Biol* **3**, 64-70.
Dreze, M., Charloteaux, B., Milstein, S., Vidalain, P. O., Yildirim, M. A., Zhong, Q., Svrzikapa, N., Romero, V., Laloux, G., Brasseur, R. et al. (2009). 'Edgetic' perturbation of a C. elegans BCL2 ortholog. *Nat Methods* **6**, 843-9.
Edwards, A. M., Kus, B., Jansen, R., Greenbaum, D., Greenblatt, J. and Gerstein, M. (2002). Bridging structural biology and genomics: assessing protein interaction data with known complexes. *Trends Genet* **18**, 529-36.
Eyckerman, S., Verhee, A., der Heyden, J. V., Lemmens, I., Ostade, X. V., Vandekerckhove, J. and Tavernier, J. (2001). Design and application of a cytokine-receptor-based interaction trap. *Nat Cell Biol* **3**, 1114-9.
Fauquet, C. M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (2005). Virus Taxonomy, VIIIth Report of the ICTV (ed. London: Elsevier/Academic Press.
Feire, A. L., Koss, H. and Compton, T. (2004). Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain. *Proc Natl Acad Sci U S A* **101**, 15470-5.
Fields, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-6.
Fossum, E., Friedel, C. C., Rajagopala, S. V., Titz, B., Baiker, A., Schmidt, T., Kraus, T., Stellberger, T., Rutenberg, C., Suthram, S. et al. (2009). Evolutionarily conserved herpesviral protein interaction networks. *PLoS Pathog* **5**, e1000570.
Fox, A., Taylor, D. and Slonim, D. K. (2009). High throughput interaction data reveals degree conservation of hub proteins. *Pac Symp Biocomput*, 391-402.
Frank, R. (1992). SPOT synthesis: an easy technique for the positionally addressble, parallel chemical synthesis on a membrane support. *Tetrahedron* **48**, 9217-9232.
Frank, R. (2002). The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods* **267**, 13-26.
Frank, R. and Overwin, H. (1996). SPOT synthesis. Epitope analysis with arrays of synthetic peptides prepared on cellulose membranes. *Methods Mol Biol* **66**, 149-69.
Fuchs, W., Granzow, H., Klupp, B. G., Karger, A., Michael, K., Maresch, C., Klopfleisch, R. and Mettenleiter, T. C. (2007). Relevance of the interaction between alphaherpesvirus UL3.5 and UL48 proteins for virion maturation and neuroinvasion. *J Virol* **81**, 9307-18.
Fuchs, W., Klupp, B. G., Granzow, H., Osterrieder, N. and Mettenleiter, T. C. (2002). The interacting UL31 and UL34 gene products of pseudorabies virus are involved in egress from the host-cell nucleus and represent components of primary enveloped but not mature virions. *J Virol* **76**, 364-78.
Fuchs, W., Klupp, B. G., Granzow, H., Rziha, H. J. and Mettenleiter, T. C. (1996). Identification and characterization of the pseudorabies virus UL3.5 protein, which is involved in virus egress. *J Virol* **70**, 3517-27.
Furlong, D., Swift, H. and Roizman, B. (1972). Arrangement of herpesvirus deoxyribonucleic acid in the core. *J Virol* **10**, 1071-4.

Geraghty, R. J., Krummenacher, C., Cohen, G. H., Eisenberg, R. J. and Spear, P. G. (1998). Entry of alphaherpesviruses mediated by poliovirus receptor-related protein 1 and poliovirus receptor. *Science* **280**, 1618-20.
Geysen, H. M., Meloen, R. H. and Barteling, S. J. (1984). Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc Natl Acad Sci U S A* **81**, 3998-4002.
Gilden, D., Nagel, M. A., Mahalingam, R., Mueller, N. H., Brazeau, E. A., Pugazhenthi, S. and Cohrs, R. J. (2009). Clinical and molecular aspects of varicella zoster virus infection. *Future Neurol* **4**, 103-117.
Gilden, D. H., Rozenman, Y., Murray, R., Devlin, M. and Vafai, A. (1987). Detection of varicella-zoster virus nucleic acid in neurons of normal human thoracic ganglia. *Ann Neurol* **22**, 377-80.
Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E. et al. (2003). A protein interaction map of Drosophila melanogaster. *Science* **302**, 1727-36.
Goehler, H., Lalowski, M., Stelzl, U., Waelter, S., Stroedicke, M., Worm, U., Droege, A., Lindenberg, K. S., Knoblich, M., Haenig, C. et al. (2004). A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. *Mol Cell* **15**, 853-65.
Goll, J., Rajagopala, S. V., Shiau, S. C., Wu, H., Lamb, B. T. and Uetz, P. (2008). MPIDB: the microbial protein interaction database. *Bioinformatics* **24**, 1743-4.
Granzow, H., Klupp, B. G., Fuchs, W., Veits, J., Osterrieder, N. and Mettenleiter, T. C. (2001). Egress of alphaherpesviruses: comparative ultrastructural study. *J Virol* **75**, 3675-84.
Grimm, C., Schaer, P., Munz, P. and Kohli, J. (1991). The strong ADH1 promoter stimulates mitotic and meiotic recombination at the ADE6 gene of Schizosaccharomyces pombe. *Mol Cell Biol* **11**, 289-98.
Hackbusch, J., Richter, K., Muller, J., Salamini, F. and Uhrig, J. F. (2005). A central role of Arabidopsis thaliana ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *Proc Natl Acad Sci U S A* **102**, 4908-12.
Hilpert, K., Winkler, D. F. and Hancock, R. E. (2007). Cellulose-bound peptide arrays: preparation and applications. *Biotechnol Genet Eng Rev* **24**, 31-106.
Hilton, J. L., Kearney, P. C. and Ames, B. N. (1965). Mode of action of the herbicide, 3-amino-1,2,4-triazole(amitrole): inhibition of an enzyme of histidine biosynthesis. *Arch Biochem Biophys* **112**, 544-7.
Hong, M., Fitzgerald, M. X., Harper, S., Luo, C., Speicher, D. W. and Marmorstein, R. (2008). Structural basis for dimerization in DNA recognition by Gal4. *Structure* **16**, 1019-26.
Hwang, J. S., Schilf, R., Drach, J. C., Townsend, L. B. and Bogner, E. (2009). Susceptibilities of human cytomegalovirus clinical isolates and other herpesviruses to new acetylated, tetrahalogenated benzimidazole D-ribonucleosides. *Antimicrob Agents Chemother* **53**, 5095-101.
Hyman, R. W., Ecker, J. R. and Tenser, R. B. (1983). Varicella-zoster virus RNA in human trigeminal ganglia. *Lancet* **2**, 814-6.
Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* **98**, 4569-74.

Ivanov, A. S., Gnedenko, O. V., Molnar, A. A., Mezentsev, Y. V., Lisitsa, A. V. and Archakov, A. I. (2007). Protein-protein interactions as new targets for drug design: virtual and experimental approaches. *J Bioinform Comput Biol* **5**, 579-92.
Jacobson, J. G., Yang, K., Baines, J. D. and Homa, F. L. (2006). Linker insertion mutations in the herpes simplex virus type 1 UL28 gene: effects on UL28 interaction with UL15 and UL33 and identification of a second-site mutation in the UL15 gene that suppresses a lethal UL28 mutation. *J Virol* **80**, 12312-23.
James, P. (2001). Yeast two-hybrid vectors and strains. *Methods Mol Biol* **177**, 41-84.
Jin, F., Avramova, L., Huang, J. and Hazbun, T. (2007). A yeast two-hybrid smart-pool-array system for protein-interaction mapping. *Nat Methods* **4**, 405-7.
Jin, F., Hazbun, T., Michaud, G. A., Salcius, M., Predki, P. F., Fields, S. and Huang, J. (2006). A pooling-deconvolution strategy for biological network elucidation. *Nat Methods* **3**, 183-9.
Johnsson, N. and Varshavsky, A. (1994). Ubiquitin-assisted dissection of protein transport across membranes. *Embo* **13**, 2686-98.
Jones, S. and Thornton, J. M. (1996). Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* **93**, 13-20.
Junker, B. H., Koschutzki, D. and Schreiber, F. (2006). Exploration of biological network centralities with CentiBiN. *BMC Bioinformatics* **7**, 219.
Kaleeba, J. A. and Berger, E. A. (2006). Kaposi's sarcoma-associated herpesvirus fusion-entry receptor: cystine transporter xCT. *Science* **311**, 1921-4.
Kibbe, W. A. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* **35**, W43-6.
Kleinschmidt-DeMasters, B. K., Amlie-Lefond, C. and Gilden, D. H. (1996). The patterns of varicella zoster virus encephalitis. *Hum Pathol* **27**, 927-38.
Koegl, M. and Uetz, P. (2007). Improving yeast two-hybrid screening systems. *Brief Funct Genomic Proteomic* **6**, 302-12.
Kopp, M., Granzow, H., Fuchs, W., Klupp, B. and Mettenleiter, T. C. (2004). Simultaneous deletion of pseudorabies virus tegument protein UL11 and glycoprotein M severely impairs secondary envelopment. *J Virol* **78**, 3024-34.
Kopp, M., Granzow, H., Fuchs, W., Klupp, B. G., Mundt, E., Karger, A. and Mettenleiter, T. C. (2003). The pseudorabies virus UL11 protein is a virion component involved in secondary envelopment in the cytoplasm. *J Virol* **77**, 5339-51.
LaCount, D. J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J. R., Schoenfeld, L. W., Ota, I., Sahasrabudhe, S., Kurschner, C. et al. (2005). A protein interaction network of the malaria parasite Plasmodium falciparum. *Nature* **438**, 103-7.
LaGuardia, J. J., Cohrs, R. J. and Gilden, D. H. (1999). Prevalence of varicella-zoster virus DNA in dissociated human trigeminal ganglion neurons and nonneuronal cells. *J Virol* **73**, 8571-7.
Lagunoff, M. and Ganem, D. (1997). The structure and coding organization of the genomic termini of Kaposi's sarcoma-associated herpesvirus. *Virology* **236**, 147-54.
Lake, C. M. and Hutt-Fletcher, L. M. (2004). The Epstein-Barr virus BFRF1 and BFLF2 proteins interact and coexpression alters their cellular localization. *Virology* **320**, 99-106.
Landgraf, C., Panni, S., Montecchi-Palazzi, L., Castagnoli, L., Schneider-Mergener, J., Volkmer-Engert, R. and Cesareni, G. (2004). Protein interaction networks by proteome peptide scanning. *PLoS Biol* **2**, E14.

Landy, A. (1989). Dynamic, Structural, and Regulatory Aspects of Lambda Site-specific Recombination. *Ann. Rev. Biochem.* **58**, 913-949.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-8.
Legrain, P., Dokhelar, M. C. and Transy, C. (1994). Detection of protein-protein interactions using different vectors in the two-hybrid system. *Nucleic Acids Res* **22**, 3241-2.
Leung, D. W., Prins, K. C., Borek, D. M., Farahbakhsh, M., Tufariello, J. M., Ramanan, P., Nix, J. C., Helgeson, L. A., Otwinowski, Z., Honzatko, R. B. et al. (2010). Structural basis for dsRNA recognition and interferon antagonism by Ebola VP35. *Nat Struct Mol Biol* **17**, 165-72.
Li, Q., Ali, M. A. and Cohen, J. I. (2006). Insulin degrading enzyme is a cellular receptor mediating varicella-zoster virus infection and cell-to-cell spread. *Cell* **127**, 305-16.
Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., Hao, T. et al. (2004). A map of the interactome network of the metazoan C. elegans. *Science* **303**, 540-3.
Lievens, S., Lemmens, I. and Tavernier, J. (2009). Mammalian two-hybrids come of age. *Trends Biochem Sci* **34**, 579-88.
Lim, J., Hao, T., Shaw, C., Patel, A. J., Szabo, G., Rual, J. F., Fisk, C. J., Li, N., Smolyar, A., Hill, D. E. et al. (2006). A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* **125**, 801-14.
Lischka, P., Hewlett, G., Wunberg, T., Baumeister, J., Paulsen, D., Goldner, T., Ruebsamen-Schaeff, H. and Zimmermann, H. (2010). In vitro and in vivo activities of the novel anticytomegalovirus compound AIC246. *Antimicrob Agents Chemother* **54**, 1290-7.
Lischka, P. and Zimmermann, H. (2008). Antiviral strategies to combat cytomegalovirus infections in transplant recipients. *Curr Opin Pharmacol* **8**, 541-8.
Lynn, D. J., Winsor, G. L., Chan, C., Richard, N., Laird, M. R., Barsky, A., Gardy, J. L., Roche, F. M., Chan, T. H., Shah, N. et al. (2008). InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Mol Syst Biol* **4**, 218.
Marchler-Bauer, A., Anderson, J. B., Derbyshire, M. K., DeWeese-Scott, C., Gonzales, N. R., Gwadz, M., Hao, L., He, S., Hurwitz, D. I., Jackson, J. D. et al. (2007). CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res* **35**, D237-40.
McGeoch, D. J. and Gatherer, D. (2005). Integrating reptilian herpesviruses into the family herpesviridae. *J Virol* **79**, 725-31.
McGeoch, D. J., Rixon, F. J. and Davison, A. J. (2006). Topics in herpesvirus genomics and evolution. *Virus Res* **117**, 90-104.
Merrifield, R. B. (1963). Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem Soc* **85**, 2149-2154.
Mettenleiter, T. C. (2002). Herpesvirus assembly and egress. *J Virol* **76**, 1537-47.
Millson, S. H., Truman, A. W. and Piper, P. W. (2003). Vectors for N- or C-terminal positioning of the yeast Gal4p DNA binding or activator domains. *Biotechniques* **35**, 60-4.
Montgomery, R. I., Warner, M. S., Lum, B. J. and Spear, P. G. (1996). Herpes simplex virus-1 entry into cells mediated by a novel member of the TNF/NGF receptor family. *Cell* **87**, 427-36.

Moriuchi, H., Moriuchi, M., Straus, S. E. and Cohen, J. I. (1993). Varicella-zoster virus open reading frame 10 protein, the herpes simplex virus VP16 homolog, transactivates herpesvirus immediate-early gene promoters. *J Virol* **67**, 2739-46.
Morse, L. S., Buchman, T. G., Roizman, B. and Schaffer, P. A. (1977). Anatomy of herpes simplex virus DNA. IX. Apparent exclusion of some parental DNA arrangements in the generation of intertypic (HSV-1 X HSV-2) recombinants. *J Virol* **24**, 231-48.
Muranyi, W., Haas, J., Wagner, M., Krohne, G. and Koszinowski, U. H. (2002). Cytomegalovirus recruitment of cellular kinases to dissolve the nuclear lamina. *Science* **297**, 854-7.
Naranatt, P. P., Krishnan, H. H., Smith, M. S. and Chandran, B. (2005). Kaposi's sarcoma-associated herpesvirus modulates microtubule dynamics via RhoA-GTP-diaphanous 2 signaling and utilizes the dynein motors to deliver its DNA to the nucleus. *J Virol* **79**, 1191-206.
Nyfeler, B., Michnick, S. W. and Hauri, H. P. (2005). Capturing protein interactions in the secretory pathway of living cells. *Proc Natl Acad Sci U S A* **102**, 6350-5.
Orchard, S., Kerrien, S., Jones, P., Ceol, A., Chatr-Aryamontri, A., Salwinski, L., Nerothin, J. and Hermjakob, H. (2007). Submit your interaction data the IMEx way: a step by step guide to trouble-free deposition. *Proteomics* **7 Suppl 1**, 28-34.
Osterman, A., Stellberger, T., Vizoso Pinto, M. G., Von Brunn, A., Haas, J. and Baiker, A. (2010). Manuscript in preparation.
Osterman, A., Stellberger, T., Vizoso Pinto, M. G., Von Brunn, A., Haas, J. and Baiker, A. (Manuscript in preparation).
Osterrieder, N., Kamil, J. P., Schumacher, D., Tischer, B. K. and Trapp, S. (2006). Marek's disease virus: from miasma to model. *Nat Rev Microbiol* **4**, 283-94.
Otte, L., Wiedemann, U., Schlegel, B., Pires, J. R., Beyermann, M., Schmieder, P., Krause, G., Volkmer-Engert, R., Schneider-Mergener, J. and Oschkinat, H. (2003). WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains. *Protein Sci* **12**, 491-500.
Pagel, P., Kovac, S., Oesterheld, M., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Mark, P., Stumpflen, V., Mewes, H. W. et al. (2005). The MIPS mammalian protein-protein interaction database. *Bioinformatics* **21**, 832-4.
Parrish, J. R., Yu, J., Liu, G., Hines, J. A., Chan, J. E., Mangiola, B. A., Zhang, H., Pacifico, S., Fotouhi, F., DiRita, V. J. et al. (2007). A proteome-wide protein interaction map for Campylobacter jejuni. *Genome Biol* **8**, R130.
Perdue, M. L., Cohen, J. C., Randall, C. C. and O'Callaghan, D. J. (1976). Biochemical studies of the maturation of herpesvirus nucleocapsid species. *Virology* **74**, UNKNOWN.
Pevenstein, S. R., Williams, R. K., McChesney, D., Mont, E. K., Smialek, J. E. and Straus, S. E. (1999). Quantitation of latent varicella-zoster virus and herpes simplex virus genomes in human trigeminal ganglia. *J Virol* **73**, 10514-8.
Pujol, A., Mosca, R., Farres, J. and Aloy, P. (2010). Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol Sci* **31**, 115-23.
R-Development-Core-Team. (2004). R: A language and environment for statistical computing, (ed. Vienna, Austria: R Foundation for Statistical Computing.
Rahaus, M., Desloges, N., Wolff, M.H. (2006). Molecular Biology of Varicella–Zoster Virus. In *Monographs in Virology*, vol. 26 (ed. G. Gross, Doerr, H.W.), pp. 1-8. Basel: Karger.

Rain, J. C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V. et al. (2001). The protein-protein interaction map of Helicobacter pylori. *Nature* **409**, 211-5.
Rajagopala, S. V., Blasche, S. and Uetz, P. (work in progress).
Rajagopala, S. V., Hughes, K. T. and Uetz, P. (2009). Benchmarking yeast two-hybrid systems using the interactions of bacterial motility proteins. *Proteomics* **9**, 5296-302.
Rajagopala, S. V., Titz, B., Goll, J., Parrish, J. R., Wohlbold, K., McKevitt, M. T., Palzkill, T., Mori, H., Finley, R. L., Jr. and Uetz, P. (2007). The protein network of bacterial motility. *Mol Syst Biol* **3**, 128.
Rajagopala, S. V. and Uetz, P. (2009). Analysis of protein-protein interactions using array-based yeast two-hybrid screens. *Methods Mol Biol* **548**, 223-45.
Rao, V. B. and Feiss, M. (2008). The bacteriophage DNA packaging motor. *Annu Rev Genet* **42**, 647-81.
Rappocciolo, G., Jenkins, F. J., Hensler, H. R., Piazza, P., Jais, M., Borowski, L., Watkins, S. C. and Rinaldo, C. R., Jr. (2006). DC-SIGN is a receptor for human herpesvirus 8 on dendritic cells and macrophages. *J Immunol* **176**, 1741-9.
Reefschlaeger, J., Bender, W., Hallenberger, S., Weber, O., Eckenberg, P., Goldmann, S., Haerter, M., Buerger, I., Trappe, J., Herrington, J. A. et al. (2001). Novel non-nucleoside inhibitors of cytomegaloviruses (BAY 38-4766): in vitro and in vivo antiviral activity and mechanism of action. *J Antimicrob Chemother* **48**, 757-67.
Reineke, U., Kramer, A. and Schneider-Mergener, J. (1999). Antigen sequence- and library-based mapping of linear and discontinuous protein-protein-interaction sites by spot synthesis. *Curr Top Microbiol Immunol* **243**, 23-36.
Reynolds, A. E., Ryckman, B. J., Baines, J. D., Zhou, Y., Liang, L. and Roller, R. J. (2001). U(L)31 and U(L)34 proteins of herpes simplex virus type 1 form a complex that accumulates at the nuclear rim and is required for envelopment of nucleocapsids. *J Virol* **75**, 8803-17.
Reynolds, A. E., Wills, E. G., Roller, R. J., Ryckman, B. J. and Baines, J. D. (2002). Ultrastructural localization of the herpes simplex virus type 1 UL31, UL34, and US3 proteins suggests specific roles in primary envelopment and egress of nucleocapsids. *J Virol* **76**, 8939-52.
Rogers, S., Wells, R. and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-8.
Roizman, B. (1996). Herpesviridae. Philadelphia, New York: Lippincott-Raven.
Roizman, B., Gu, H. and Mandel, G. (2005). The first 30 minutes in the life of a virus: unREST in the nucleus. *Cell Cycle* **4**, 1019-21.
Rozen, R., Sathish, N., Li, Y. and Yuan, Y. (2008). Virion-wide protein interactions of Kaposi's sarcoma-associated herpesvirus. *J Virol* **82**, 4742-50.
Rual, J. F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P. O. et al. (2004). Human ORFeome version 1.1: a platform for reverse proteomics. *Genome Res* **14**, 2128-35.
Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N. et al. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature* **437**, 1173-8.
Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U. and Eisenberg, D. (2004). The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res* **32**, D449-51.

Sampathkumar, P., Drage, L. A. and Martin, D. P. (2009). Herpes zoster (shingles) and postherpetic neuralgia. *Mayo Clin Proc* **84**, 274-80.
Sander, G., Konrad, A., Thurau, M., Wies, E., Leubert, R., Kremmer, E., Dinkel, H., Schulz, T., Neipel, F. and Sturzl, M. (2008). Intracellular localization map of human herpesvirus 8 proteins. *J Virol* **82**, 1908-22.
Sauerbrei, A., Deinhardt, S., Zell, R. and Wutzler, P. (2010). Phenotypic and genotypic characterization of acyclovir-resistant clinical isolates of herpes simplex virus. *Antiviral Res* **86**, 246-52.
Savva, C. G., Holzenburg, A. and Bogner, E. (2004). Insights into the structure of human cytomegalovirus large terminase subunit pUL56. *FEBS Lett* **563**, 135-40.
Scheffczik, H., Savva, C. G., Holzenburg, A., Kolesnikova, L. and Bogner, E. (2002). The terminase subunits pUL56 and pUL89 of human cytomegalovirus are DNA-metabolizing proteins with toroidal structure. *Nucleic Acids Res* **30**, 1695-703.
Seifert, M. H. (2005). ProPose: steered virtual screening by simultaneous protein-ligand docking and ligand-ligand alignment. *J Chem Inf Model* **45**, 449-60.
Serebriiskii, I., Estojak, J., Berman, M. and Golemis, E. A. (2000). Approaches to detecting false positives in yeast two-hybrid systems. *Biotechniques* **28**, 328-30, 332-6.
Serebriiskii, I. G. and Golemis, E. A. (2001). Two-hybrid system and false positives. Approaches to detection and elimination. *Methods Mol Biol* **177**, 123-34.
Seyboldt, C., Granzow, H. and Osterrieder, N. (2000). Equine herpesvirus 1 (EHV-1) glycoprotein M: effect of deletions of transmembrane domains. *Virology* **278**, 477-89.
Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504.
Sharan, R., Ulitsky, I. and Shamir, R. (2007). Network-based prediction of protein function. *Mol Syst Biol* **3**, 88.
Simonis, N., Rual, J. F., Carvunis, A. R., Tasan, M., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Sahalie, J. M., Venkatesan, K., Gebreab, F. et al. (2009). Empirically controlled mapping of the Caenorhabditis elegans protein-protein interactome network. *Nat Methods* **6**, 47-54.
Singer, G. P., Newcomb, W. W., Thomsen, D. R., Homa, F. L. and Brown, J. C. (2005). Identification of a region in the herpes simplex virus scaffolding protein required for interaction with the portal. *J Virol* **79**, 132-9.
Skepper, J. N., Whiteley, A., Browne, H. and Minson, A. (2001). Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment --> deenvelopment --> reenvelopment pathway. *J Virol* **75**, 5697-702.
Sourvinos, G. and Everett, R. D. (2002). Visualization of parental HSV-1 genomes and replication compartments in association with ND10 in live infected cells. *Embo J* **21**, 4989-97.
Steiner, I., Kennedy, P. G. and Pachner, A. R. (2007). The neurotropic herpes viruses: herpes simplex and varicella-zoster. *Lancet Neurol* **6**, 1015-28.
Stellberger, T., Hauser, R., Baiker, A., Pothineni, V. R., Haas, J. and Uetz, P. (2010). Improving the yeast two-hybrid system with permutated fusions proteins: the Varicella Zoster Virus interactome. *Proteome Sci* **8**, 8.
Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S. et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122**, 957-68.

Takahashi, M., Otsuka, T., Okuno, Y., Asano, Y. and Yazaki, T. (1974). Live vaccine used to prevent the spread of varicella in children in hospital. *Lancet* **2**, 1288-90.

Taylor, T. J., McNamee, E. E., Day, C. and Knipe, D. M. (2003). Herpes simplex virus replication compartments can form by coalescence of smaller compartments. *Virology* **309**, 232-47.

Titz, B., Rajagopala, S. V., Goll, J., Hauser, R., McKevitt, M. T., Palzkill, T. and Uetz, P. (2008). The binary protein interactome of Treponema pallidum--the syphilis spirochete. *PLoS One* **3**, e2292.

Tornow, J. and Santangelo, G. M. (1990). Efficient expression of the Saccharomyces cerevisiae glycolytic gene ADH1 is dependent upon a cis-acting regulatory element (UASRPG) found initially in genes encoding ribosomal proteins. *Gene* **90**, 79-85.

Trus, B. L., Cheng, N., Newcomb, W. W., Homa, F. L., Brown, J. C. and Steven, A. C. (2004). Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J Virol* **78**, 12668-71.

Trus, B. L., Newcomb, W. W., Cheng, N., Cardone, G., Marekov, L., Homa, F. L., Brown, J. C. and Steven, A. C. (2007). Allosteric signaling and a nuclear exit strategy: binding of UL25/UL17 heterodimers to DNA-Filled HSV-1 capsids. *Mol Cell* **26**, 479-89.

Uetz, P., Dong, Y. A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S. V., Roupelieva, M., Rose, D., Fossum, E. et al. (2006). Herpesviral protein networks and their interaction with the human proteome. *Science* **311**, 239-42.

Uetz, P. and Finley, R. L., Jr. (2005). From protein networks to biological systems. *FEBS Lett* **579**, 1821-7.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P. et al. (2000a). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**, 623-7.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P. et al. (2000b). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**, 623-627.

Underwood, M. R., Harvey, R. J., Stanat, S. C., Hemphill, M. L., Miller, T., Drach, J. C., Townsend, L. B. and Biron, K. K. (1998). Inhibition of human cytomegalovirus DNA maturation by a benzimidazole ribonucleoside is mediated through the UL89 gene product. *J Virol* **72**, 717-25.

UniProt-Consortium. (2010). The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res* **38**, D142-8.

Van Criekinge, W. and Beyaert, R. (1999). Yeast Two-Hybrid: State of the Art. *Biol Proced Online* **2**, 1-38.

van Zeijl, M., Fairhurst, J., Jones, T. R., Vernon, S. K., Morin, J., LaRocque, J., Feld, B., O'Hara, B., Bloom, J. D. and Johann, S. V. (2000). Novel class of thiourea compounds that inhibit herpes simplex virus type 1 DNA cleavage and encapsidation: resistance maps to the UL6 gene. *J Virol* **74**, 9054-61.

Venkatesan, K., Rual, J. F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K. I. et al. (2009). An empirical framework for binary interactome mapping. *Nat Methods* **6**, 83-90.

Vidal, M., Brachmann, R. K., Fattaey, A., Harlow, E. and Boeke, J. D. (1996). Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc Natl Acad Sci U S A* **93**, 10315-20.

Vidalain, P. O., Boxem, M., Ge, H., Li, S. and Vidal, M. (2004). Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods* **32**, 363-70.

Visalli, R. J., Fairhurst, J., Srinivas, S., Hu, W., Feld, B., DiGrandi, M., Curran, K., Ross, A., Bloom, J. D., van Zeijl, M. et al. (2003). Identification of small molecule compounds that selectively inhibit varicella-zoster virus replication. *J Virol* **77**, 2349-58.

Visalli, R. J., Knepper, J., Goshorn, B., Vanover, K., Burnside, D. M., Irven, K., McGauley, R. and Visalli, M. (2009). Characterization of the Varicella-zoster virus ORF25 gene product: pORF25 interacts with multiple DNA encapsidation proteins. *Virus Res* **144**, 58-64.

Visalli, R. J., Nicolosi, D. M., Irven, K. L., Goshorn, B., Khan, T. and Visalli, M. A. (2007). The Varicella-zoster virus DNA encapsidation genes: Identification and characterization of the putative terminase subunits. *Virus Res* **129**, 200-11.

Vizoso Pinto, M. G., Stellberger, T., Haase, R., Haas, J., Sommer, M., Arvin, A. and Baiker, A. (2010). Varicella Zoster Virus ORF25 gene product: a promiscuous but essential
protein (Manuscript in preparation).

Vizoso Pinto, M. G., Stellberger, T., Haase, R., Haas, J., Sommer, M., Arvin, A. and Baiker, A. (Manuscript in preparation). Varicella Zoster Virus ORF25 gene product: a promiscuous but essential
protein (Manuscript in preparation).

Vizoso Pinto, M. G., Villegas, J. M., Peter, J., Haase, R., Haas, J., Lotz, A. S., Muntau, A. C. and Baiker, A. (2009). LuMPIS--a modified luminescence-based mammalian interactome mapping pull-down assay for the investigation of protein-protein interactions encoded by GC-low ORFs. *Proteomics* **9**, 5303-8.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S. and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**, 399-403.

Walhout, A. J. and Vidal, M. (2001). Protein interaction maps for model organisms. *Nat Rev Mol Cell Biol* **2**, 55-62.

Wang, Q. Y., Zhou, C., Johnson, K. E., Colgrove, R. C., Coen, D. M. and Knipe, D. M. (2005). Herpesviral latency-associated transcript gene promotes assembly of heterochromatin on viral lytic-gene promoters in latent infection. *Proc Natl Acad Sci U S A* **102**, 16055-9.

Wang, Y., Addess, K. J., Chen, J., Geer, L. Y., He, J., He, S., Lu, S., Madej, T., Marchler-Bauer, A., Thiessen, P. A. et al. (2007). MMDB: annotating protein sequences with Entrez's 3D-structure database. *Nucleic Acids Res* **35**, D298-300.

Wang, Y., Geer, L. Y., Chappey, C., Kans, J. A. and Bryant, S. H. (2000). Cn3D: sequence and structure views for Entrez. *Trends Biochem Sci* **25**, 300-2.

Warner, M. S., Geraghty, R. J., Martinez, W. M., Montgomery, R. I., Whitbeck, J. C., Xu, R., Eisenberg, R. J., Cohen, G. H. and Spear, P. G. (1998). A cell surface protein with herpesvirus entry activity (HveB) confers susceptibility to infection by mutants of herpes simplex virus type 1, herpes simplex virus type 2, and pseudorabies virus. *Virology* **246**, 179-89.

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. and Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* **393**, 440-2.
Weigelt, J. (2010). Structural genomics-impact on biomedicine and drug discovery. *Exp Cell Res* **316**, 1332-8.
White, C. A., Stow, N. D., Patel, A. H., Hughes, M. and Preston, V. G. (2003). Herpes simplex virus type 1 portal protein UL6 interacts with the putative terminase subunits UL15 and UL28. *J Virol* **77**, 6351-8.
Xie, L., Li, J., Xie, L. and Bourne, P. E. (2009). Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput Biol* **5**, e1000387.
Xin, X., Rual, J. F., Hirozane-Kishikawa, T., Hill, D. E., Vidal, M., Boone, C. and Thierry-Mieg, N. (2009). Shifted Transversal Design smart-pooling for high coverage interactome mapping. *Genome Res* **19**, 1262-9.
Yamauchi, Y., Wada, K., Goshima, F., Takakuwa, H., Daikoku, T., Yamada, M. and Nishiyama, Y. (2001). The UL14 protein of herpes simplex virus type 2 translocates the minor capsid protein VP26 and the DNA cleavage and packaging UL33 protein into the nucleus of coexpressing cells. *J Gen Virol* **82**, 321-30.
Yang, K. and Baines, J. D. (2006). The putative terminase subunit of herpes simplex virus 1 encoded by UL28 is necessary and sufficient to mediate interaction between pUL15 and pUL33. *J Virol* **80**, 5733-9.
Yang, K., Homa, F. and Baines, J. D. (2007). Putative terminase subunits of herpes simplex virus 1 form a complex in the cytoplasm and interact with portal protein in the nucleus. *J Virol* **81**, 6419-33.
Yefenof, E., Klein, G., Jondal, M. and Oldstone, M. B. (1976). Surface markers on human B and T-lymphocytes. IX. Two-color immunofluorescence studies on the association between ebv receptors and complement receptors on the surface of lymphoid cell lines. *Int J Cancer* **17**, 693-700.
Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N. et al. (2008). High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104-10.
Zhang, Z., Selariu, A., Warden, C., Huang, G., Huang, Y., Zaccheus, O., Cheng, T., Xia, N. and Zhu, H. (2010a). Genome-wide mutagenesis reveals that ORF7 is a novel VZV skin-tropic factor. *PLoS Pathog* **6**, e1000971.
Zhang, Z. T., Zhou, Y., Li, Y., Shao, S. Q., Li, B. Y., Shi, H. Y. and Li, X. B. (2010b). Interactome analysis of the six cotton 14-3-3s that are preferentially expressed in fibres and involved in cell elongation. *J Exp Bot* **61**, 3331-44.
Zhong, J., Zhang, H., Stanyon, C. A., Tromp, G. and Finley, R. L., Jr. (2003). A strategy for constructing large protein interaction maps using the yeast two-hybrid system: regulated expression arrays and two-phase mating. *Genome Res* **13**, 2691-9.

6 Appendix

6.1 Supplementary Data

6.2 Curriculum vitae

- Pre-degree in biology at the University of Karlsruhe, in 2004.
  Concentrations: Zoology, Phytology, Chemistry, Physics
- Diploma in Biology at the Universität Karlsruhe in 2006
  Concentrations: Microbiology, Zoology
  Thesis: "Investigations of the Interaction Network of the Small GTPase Rho1p in *Saccharomyces cerevisiae*."
  The diploma thesis was performed in the research group of Prof. Dr. Nils Johnsson at the University of Münster.
- Publications:

1) Stellberger, T., Hauser, R., Baiker, A., Pothineni, V. R., Haas, J. and Uetz, P. (2010). Improving the yeast two-hybrid system with permutated fusions proteins: the Varicella Zoster Virus interactome. **Proteome Sci.** 8, 8.

2) Fossum, E., Friedel, C. C., Rajagopala, S. V., Titz, B., Baiker, A., Schmidt, T., Kraus, T., Stellberger, T., Rutenberg, C., Suthram, S. et al. (2009). Evolutionarily conserved herpesviral protein interaction networks. **PLoS Pathog.** 5, e1000570.

3) Chen, Y., Rajagopala, S. V., Stellberger, T. and Uetz, P. (2010). Exhaustive benchmarking of the yeast two-hybrid system. **Nat Methods** 7, 667 - 668.

Appendix

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
1	ORF24N	ORF1N	CN	1	1	ORF1-ORF24						
2	ORF1N	ORF25	NN CN CC	3	1	ORF1-ORF25						
3	ORF1N	ORF27	NN	1	1	ORF1-ORF27						
4	ORF3	ORF1N	CN	1	1	ORF1-ORF3						
5	ORF43	ORF1N	NC	1	1	ORF1-ORF43						
6	ORF1	ORF60C	CN	1	2	ORF1-ORF60					redundant construct	red
7	ORF1N	ORF60C	NN	1	red 2	ORF1-ORF60						perm-red
8	ORF1N	ORF62	NN	1	1	ORF1-ORF62						
9	ORF10	ORF32	CC	1	1	ORF10-ORF32						
10	ORF10	ORF56C	CN	1	1	ORF10-ORF56						
11	ORF57	ORF10	NN NC CN CC	4	1	ORF10-ORF57						
12	ORF11	ORF13	CN	1	1	ORF11-ORF13						
13	ORF11	ORF27	NN	1	1	ORF11-ORF27						
14	ORF11	ORF38	NN	1	1	ORF11-ORF38					HSV1_UL47-HSV1_UL21	Y2H
15	ORF12	ORF12	CC	1	1	ORF12-ORF12						
16	ORF12	ORF16	CN	1	1	ORF12-ORF16						
17	ORF12	ORF24N	NC	1	1	ORF12-ORF24						
18	ORF12	ORF25	NN CN CC	3	2	ORF12-ORF25					HSV1_UL46-HSV1_UL33, redundant construct	Y2H red
19	ORF12C	ORF25	CN CC	2	red 2	ORF12-ORF25						perm-red
20	ORF12	ORF27	CN	1	1	ORF12-ORF27						
21	ORF33	ORF12C	NN	1	1	ORF12-ORF33						
22	ORF33.5	ORF12C	NN	1	1	ORF12-ORF33.5						
23	ORF12C	ORF37N	CC	1	1	ORF12-ORF37						
24	ORF12	ORF38	CN	1	1	ORF12-ORF38	HSV1_UL46-HSV1_UL21	18602131				LC
25	ORF12	ORF39	NC	1	1	ORF12-ORF39						
26	ORF12	ORF42	NC	1	2	ORF12-ORF42					redundant construct	red
27	ORF12C	ORF42	CC	1	red 2	ORF12-ORF42						perm-red
28	ORF12	ORF60C	NN CN	2	1	ORF12-ORF60						
29	ORF65	ORF12C	NN	1	1	ORF12-ORF65						
30	ORF12	ORF9	NC	1	1	ORF12-ORF9	HSV1_UL46-HSV1_UL49	18602131				LC
31	ORF6	ORF13	CC	1	1	ORF13-ORF6						
32	ORF14	ORF35	NC	1	1	ORF14-ORF35						
33	ORF64	ORF14N	CC	1	1	ORF14-ORF64						
34	ORF24N	ORF15	CN	1	1	ORF15-ORF24						
35	ORF15F	ORF25	NN	1	2	ORF15-ORF25					HSV1_UL43-HSV1_UL33, redundant construct	Y2H red
36	ORF15N	ORF25	NN NC	2	red 2	ORF15-ORF25						perm-red
37	ORF15	ORF42	CC	1	1	ORF15-ORF42						
38	ORF15F	ORF60C	NN	1	1	ORF15-ORF60						
39	ORF68F	ORF15	CN	1	1	ORF15-ORF68						
40	ORF9	ORF15	CN	1	1	ORF15-ORF9						
41	ORF16	ORF16	NN CN CC	3	1	ORF16-ORF16	EBV_BMRF1-EBV_BMRF1, KSHV_59-KSHV_59	15286084	15075322			LC
42	ORF19	ORF16	NN CN CC	3	1	ORF16-ORF19						
43	ORF24N	ORF16	NN CN CC	3	1	ORF16-ORF24						
44	ORF3	ORF16	NN	1	1	ORF16-ORF3						
45	ORF33	ORF16	NN	1	1	ORF16-ORF33						
46	ORF37N	ORF16	CN	1	1	ORF16-ORF37						

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
47	ORF16	ORF38	CN	1	2	ORF16-ORF38					redundant direction	red
48	ORF38	ORF16	CN	1	red 2	ORF16-ORF38						perm-red
49	ORF42	ORF16	NN	1	1	ORF16-ORF42						
50	ORF58	ORF16	NN	1	1	ORF16-ORF58						
51	ORF16	ORF62	NN CN	2	1	ORF16-ORF62						
52	ORF64	ORF16	NN CN	2	1	ORF16-ORF64						
53	ORF8	ORF16	CN CC	2	1	ORF16-ORF8						
54	ORF39N	ORF17	CN	1	1	ORF17-ORF39						
55	ORF18C	ORF18C	CN	1	2	ORF18-ORF18	EBV_BaRF1-EBV_BaRF1	17446270			HSV1_UL40-HSV1_UL40, KSHV_60-KSHV_60, redundant construct	LC Y2H red
56	ORF18C	ORF18N	CN CC	2	red 2	ORF18-ORF18						perm-red
57	ORF18C	ORF19	NN CN	2	2	ORF18-ORF19	HSV1_UL39-HSV1_UL40	1322407	3012359		KSHV_61-KSHV_60, redundant construct	LC Y2H red
58	ORF19	ORF18C	NN CC	2	red 2	ORF18-ORF19						perm-red
59	ORF18C	ORF23	CN CC	2	3	ORF18-ORF23					redundant construct	red
60	ORF18N	ORF23	NN CN CC	3	red 3	ORF18-ORF23						perm-red
61	ORF23	ORF18N	NC	1	red 3	ORF18-ORF23						perm-red
62	ORF18C	ORF24N	CC	1	1	ORF18-ORF24						
63	ORF18	ORF25	NN CN CC	3	3	ORF18-ORF25					HSV1_UL40-HSV1_UL33, KSHV_67.5-KSHV_60, redundant construct	Y2H red
64	ORF18C	ORF25	NN CN CC	3	red 3	ORF18-ORF25						perm-red
65	ORF18N	ORF25	CN CC	2	red 3	ORF18-ORF25						perm-red
66	ORF18C	ORF26	CC	1	3	ORF18-ORF26					KSHV_68-KSHV_60, redundant construct	Y2H red
67	ORF26	ORF18C	CN	1	red 3	ORF18-ORF26						perm-red
68	ORF26	ORF18N	CN CC	2	red 3	ORF18-ORF26						perm-red
69	ORF28	ORF18C	CC	1	2	ORF18-ORF28					HSV1_UL30-HSV1_UL40, redundant construct	Y2H red
70	ORF28	ORF18N	CC	1	red 2	ORF18-ORF28						perm-red
71	ORF29	ORF18C	CN	1	2	ORF18-ORF29					redundant construct	red
72	ORF29	ORF18N	CN CC	2	red 2	ORF18-ORF29						perm-red
73	ORF33	ORF18N	NC CC	2	1	ORF18-ORF33						
74	ORF33.5	ORF18C	NN	1	2	ORF18-ORF33.5					redundant construct	red
75	ORF33.5	ORF18N	NN NC	2	red 2	ORF18-ORF33.5						perm-red
76	ORF18C	ORF37N	CC	1	1	ORF18-ORF37						
77	ORF18C	ORF41	CN CC	2	2	ORF18-ORF41					redundant direction	red
78	ORF41	ORF18C	NN	1	red 2	ORF18-ORF41						perm-red
79	ORF43	ORF18N	CC	1	1	ORF18-ORF43						
80	ORF18C	ORF53	CC	1	1	ORF18-ORF53						
81	ORF18C	ORF56C	CC	1	1	ORF18-ORF56						
82	ORF57	ORF18N	CN	1	1	ORF18-ORF57						
83	ORF18C	ORF58	CC	1	2	ORF18-ORF58					redundant construct	red
84	ORF58	ORF18N	NN CN	2	red 2	ORF18-ORF58						perm-red
85	ORF18C	ORF60C	CN CC	2	1	ORF18-ORF60						
86	ORF62	ORF18C	CN	1	1	ORF18-ORF62						
87	ORF65	ORF18C	NN	1	2	ORF18-ORF65					redundant construct	red
88	ORF65	ORF18N	NN	1	red 2	ORF18-ORF65						perm-red
89	ORF18C	ORF68	CC	1	2	ORF18-ORF68					redundant construct	red
90	ORF18C	ORF68F	CC	1	red 2	ORF18-ORF68						perm-red
91	ORF18C	ORF9a	CC	1	1	ORF18-ORF9a	EBV_BaRF1-EBV_BLRF1	17446270				LC

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
92	ORF18C	S/L C	CC	1	1	ORF18-S/L						
93	ORF19	ORF19	NN CN	2	1	ORF19-ORF19					MCMV_M45-MCMV_M45, KSHV_61-KSHV_61	Y2H
94	ORF19	ORF24N	CC	1	2	ORF19-ORF24					redundant direction	red
95	ORF24N	ORF19	CN	1	red 2	ORF19-ORF24						perm-red
96	ORF19	ORF25	NN NC CN CC	4	1	ORF19-ORF25					MCMV_M45-MCMV_M51	Y2H
97	ORF19	ORF26	NC CC	2	2	ORF19-ORF26					redundant direction	red
98	ORF26	ORF19	CN	1	red 2	ORF19-ORF26						perm-red
99	ORF19	ORF27	NN	1	1	ORF19-ORF27						
100	ORF19	ORF38	NN CN	2	1	ORF19-ORF38						
101	ORF19	ORF39N	CN	1	1	ORF19-ORF39						
102	ORF19	ORF41	NC	1	1	ORF19-ORF41						
103	ORF19	ORF42	CN	1	1	ORF19-ORF42						
104	ORF43	ORF19	NN	1	1	ORF19-ORF43						
105	ORF19	ORF48	NC	1	1	ORF19-ORF48						
106	ORF19	ORF53	NC CC	2	1	ORF19-ORF53						
107	ORF19	ORF57	CN	1	1	ORF19-ORF57						
108	ORF19	ORF58	NC	1	1	ORF19-ORF58						
109	ORF19	ORF60C	CN	1	1	ORF19-ORF60						
110	ORF19	ORF62	NN CN	2	1	ORF19-ORF62						
111	ORF19	ORF68F	NC	1	1	ORF19-ORF68						
112	ORF8	ORF19	NN	1	1	ORF19-ORF8					MCMV_M45-MCMV_M72	Y2H
113	ORF19	ORF9a	CC	1	1	ORF19-ORF9a						
114	ORF2	ORF25	NN NC	2	1	ORF2-ORF25						
115	ORF2	ORF56C	CN	1	1	ORF2-ORF56						
116	ORF4	ORF20	NC	1	1	ORF20-ORF4						
117	ORF21	ORF22N	NN	1	1	ORF21-ORF22	HSV1_UL36-HSV1_UL37, KSHV_63-KSHV_64	16014918	18602131	18321973	EBV_BOLF1-EBV_BPLF1	LC Y2H
118	ORF21	ORF23	NN	1	1	ORF21-ORF23	HSV1_UL35-HSV1_UL37	18602131			KSHV_65-KSHV_63	LC Y2H
119	ORF24N	ORF21	NC	1	1	ORF21-ORF24						
120	ORF21	ORF27	NN	1	2	ORF21-ORF27					redundant direction	red
121	ORF27	ORF21	NN NC	2	red 2	ORF21-ORF27						perm-red
122	ORF3	ORF21	NC	1	1	ORF21-ORF3						
123	ORF33	ORF21	NN NC	2	1	ORF21-ORF33						
124	ORF33.5	ORF21	NN NC	2	1	ORF21-ORF33.5						
125	ORF21	ORF34	NN	1	1	ORF21-ORF34						
126	ORF21	ORF42	CN	1	2	ORF21-ORF42					redundant direction	red
127	ORF42	ORF21	NN NC	2	red 2	ORF21-ORF42						perm-red
128	ORF43	ORF21	NC	1	1	ORF21-ORF43						
129	ORF21	ORF60C	NN	1	1	ORF21-ORF60	KSHV_63-KSHV_47	18321973				LC
130	ORF64	ORF21	NC	1	1	ORF21-ORF64						
131	ORF65	ORF21	NN	1	1	ORF21-ORF65						
132	ORF24N	ORF22N	NN	1	1	ORF22-ORF24						
133	ORF33	ORF22N	NN	1	1	ORF22-ORF33						
134	ORF33.5	ORF22N	NN	1	1	ORF22-ORF33.5					EBV_BPLF1-EBV_BDRF1	Y2H
135	ORF38	ORF22N	NN	1	1	ORF22-ORF38					EBV_BPLF1-EBV_BTRF1	Y2H
136	ORF41	ORF22N	NN	1	1	ORF22-ORF41	KSHV_64-KSHV_26	18321973				LC

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
137	ORF43	ORF22N	NN	1	1	ORF22-ORF43						
138	ORF65	ORF22N	NN	1	1	ORF22-ORF65						
139	ORF68	ORF22N	NN	1	1	ORF22-ORF68						
140	ORF9a	ORF22N	NN	1	1	ORF22-ORF9a						
141	ORF24N	ORF23	CN CC	2	1	ORF23-ORF24						
142	ORF23	ORF25	NC	1	1	ORF23-ORF25					MCMV_M48.2-MCMV_M51, EBV_BFRF3-EBV_BFRF4	Y2H
143	ORF26	ORF23	CN CC	2	1	ORF23-ORF26						
144	ORF23	ORF27	NN	1	1	ORF23-ORF27						
145	ORF3	ORF23	NC	1	1	ORF23-ORF3						
146	ORF33	ORF23	NN NC	2	1	ORF23-ORF33						
147	ORF23	ORF39N	NN	1	2	ORF23-ORF39					redundant direction	red
148	ORF39N	ORF23	NC	1	red 2	ORF23-ORF39						perm-red
149	ORF23	ORF60C	NN	1	2	ORF23-ORF60					redundant direction	red
150	ORF60C	ORF23	NC CN CC	3	red 2	ORF23-ORF60						perm-red
151	ORF65	ORF23	NN	1	1	ORF23-ORF65						
152	ORF68	ORF23	NN	1	1	ORF23-ORF68						
153	ORF7	ORF23	NN	1	1	ORF23-ORF7	HSV1_UL35-HSV1_UL51	18602131				LC
154	ORF23	ORF9	NC	1	2	ORF23-ORF9					redundant direction	red
155	ORF9	ORF23	CN CC	2	red 2	ORF23-ORF9						perm-red
156	ORF9a	ORF23	NN	1	1	ORF23-ORF9a						
157	ORF24N	ORF24N	CN CC	2	1	ORF24-ORF24					EBV_BFRF1-EBV_BFRF1	Y2H
158	ORF24N	ORF25	NN NC CN CC	4	1	ORF24-ORF25					EBV_BFRF1-EBV_BFRF4	Y2H
159	ORF26	ORF24N	CC	1	1	ORF24-ORF26						
160	ORF24	ORF27	NN NC CN CC	4	4	ORF24-ORF27	HSV1_UL34-HSV1_UL31	10627546	11507225	15731273	HSV1_UL31-HSV1_UL34, redundant direction construct	LC Y2H red
161	ORF24N	ORF27	NN NC CN CC	4	red 4	ORF24-ORF27	EBV_BFRF1-EBV_BFLF2	15003866	15731265	17446270	MCMV_M53-MCMV_M50	perm-red
162	ORF27	ORF24	NN	1	red 4	ORF24-ORF27	MCMV_M53-MCMV_M50	17005637			EBV_BFRF1-EBV_BFLF2	perm-red
163	ORF27	ORF24N	NC CN CC	3	red 4	ORF24-ORF27						perm-red
164	ORF24N	ORF34	NN	1	1	ORF24-ORF34						
165	ORF24N	ORF38	CN	1	1	ORF24-ORF38						
166	ORF24	ORF39N	NN	1	2	ORF24-ORF39					redundant construct	red
167	ORF24N	ORF39N	NN CN	2	red 2	ORF24-ORF39						perm-red
168	ORF24N	ORF41	NC CN CC	3	1	ORF24-ORF41					MCMV_M85-MCMV_M50	Y2H
169	ORF24N	ORF42	CN	1	2	ORF24-ORF42					redundant direction	red
170	ORF42	ORF24N	CC	1	red 2	ORF24-ORF42						perm-red
171	ORF24	ORF46	CN	1	1	ORF24-ORF46					HSV1_UL34-HSV1_UL14	Y2H
172	ORF24N	ORF52	NN	1	1	ORF24-ORF52						
173	ORF24N	ORF53	CC	1	1	ORF24-ORF53						
174	ORF57	ORF24N	CN	1	1	ORF24-ORF57						
175	ORF24N	ORF58	CN	1	1	ORF24-ORF58						
176	ORF60C	ORF24N	CC	1	1	ORF24-ORF60					EBV_BKRF2-EBV_BFRF1	Y2H
177	ORF24N	ORF62	NN CN	2	1	ORF24-ORF62						
178	ORF24N	ORF64	CN	1	1	ORF24-ORF64						
179	ORF24N	ORF66	CN	1	1	ORF24-ORF66	HSV1_UL34-HSV1_US3	10627546				LC
180	ORF67C	ORF24N	CC	1	1	ORF24-ORF67						
181	ORF24N	ORF9a	CC	1	1	ORF24-ORF9a						

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
182	ORF25	ORF25	NN NC CC	3	1	ORF25-ORF25					HSV1_UL33-HSV1_UL33, MCMV_M51-MCMV_M51, EBV_BFRF4-EBV_BFRF4	Y2H
183	ORF26	ORF25	NC CN CC	3	1	ORF25-ORF26						
184	ORF27	ORF25	NC CN CC	3	1	ORF25-ORF27					MCMV_M53-MCMV_M51, EBV_BFLF2-EBV_BFRF4, KSHV_67.5-KSHV_69	Y2H
185	ORF28	ORF25	CN CC	2	1	ORF25-ORF28					HSV1_UL30-HSV1_UL33, MCMV_M51-MCMV_M54, EBV_BALF5-EBV_BFRF4, KSHV_67.5-KSHV_9	Y2H
186	ORF29	ORF25	CN CC	2	1	ORF25-ORF29					EBV_BALF2-EBV_BFRF4	Y2H
187	ORF3	ORF25	NN NC CN CC	4	1	ORF25-ORF3					HSV1_UL55-HSV1_UL33	Y2H
188	ORF30	ORF25	NN	1	1	ORF25-ORF30	HSV1_UL28-HSV1_UL33	17035316			HSV1_UL28-HSV1_UL33, EBV_BFRF4-EBV_BALF3	LC Y2H
189	ORF25	ORF31N	CN	1	2	ORF25-ORF31					EBV_BALF4-EBV_BFRF4, redundant construct	Y2H red
190	ORF31C	ORF25	CC	1	red 2	ORF25-ORF31						perm-red
191	ORF32	ORF25	NN NC CN CC	4	1	ORF25-ORF32						
192	ORF33	ORF25	NN NC CC	3	1	ORF25-ORF33						
193	ORF33.5	ORF25	NN NC	2	1	ORF25-ORF33.5					EBV_BdRF1-EBV_BFRF4, EBV_BFRF4-EBV_BDRF1	Y2H
194	ORF36	ORF25	NN NC	2	1	ORF25-ORF36						
195	ORF25	ORF38	NN	1	1	ORF25-ORF38					EBV_BFRF4-EBV_BTRF1, KSHV_67.5-KSHV_23	Y2H
196	ORF25	ORF39N	NN	1	2	ORF25-ORF39					redundant direction	red
197	ORF39N	ORF25	NN NC CN CC	4	red 2	ORF25-ORF39						perm-red
198	ORF41	ORF25	NN NC	2	1	ORF25-ORF41						
199	ORF42	ORF25	NN NC CN CC	4	1	ORF25-ORF42						
200	ORF43	ORF25	CN CC	2	1	ORF25-ORF43					HSV1_UL17-HSV1_UL33, MCMV_M93-MCMV_M51, EBV_BGLF1-EBV_BFRF4	Y2H
201	ORF44	ORF25	NN	1	1	ORF25-ORF44					HSV1_UL16-HSV1_UL33	Y2H
202	ORF48	ORF25	NC	1	1	ORF25-ORF48						
203	ORF49	ORF25	NN NC CN	3	1	ORF25-ORF49						
204	ORF50	ORF25	NN NC CN CC	4	2	ORF25-ORF50					MCMV_M100-MCMV_M51, EBV_BFRF4-EBV_BBRF3, redundant construct	Y2H red
205	ORF50C	ORF25	CN CC	2	red 2	ORF25-ORF50						perm-red
206	ORF51	ORF25	NN	1	1	ORF25-ORF51					HSV1_UL9-HSV1_UL33	Y2H
207	ORF52	ORF25	NN NC	2	1	ORF25-ORF52					EBV_BBLF2-EBV_BFRF4	Y2H
208	ORF55	ORF25	NN NC	2	1	ORF25-ORF55					EBV_BBLF4-EBV_BFRF4	Y2H
209	ORF57	ORF25	NN NC CN CC	4	1	ORF25-ORF57						
210	ORF58	ORF25	NN CN CC	3	1	ORF25-ORF58						
211	ORF59	ORF25	NC	1	1	ORF25-ORF59					HSV1_UL2-HSV1_UL33	Y2H
212	ORF6	ORF25	CN CC	2	1	ORF25-ORF6						
213	ORF60C	ORF25	NC CN CC	3	1	ORF25-ORF60					EBV_BKRF2-EBV_BFRF4	Y2H
214	ORF62	ORF25	CN CC	2	1	ORF25-ORF62						
215	ORF64	ORF25	NN NC CN CC	4	1	ORF25-ORF64					HSV1_US10-HSV1_UL33	Y2H
216	ORF65	ORF25	NN NC	2	2	ORF25-ORF65					redundant construct	red
217	ORF65N	ORF25	NN NC CN CC	4	red 2	ORF25-ORF65						perm-red
218	ORF67	ORF25	NC	1	2	ORF25-ORF67					redundant construct	red
219	ORF67C	ORF25	NC CN CC	3	red 2	ORF25-ORF67						perm-red
220	ORF68C	ORF25	NN NC CN CC	4	2	ORF25-ORF68					redundant construct	red
221	ORF68F	ORF25	NN NC	2	red 2	ORF25-ORF68						perm-red

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
222	ORF8	ORF25	NN NC CN CC	4	1	ORF25-ORF8					MCMV_M72-MCMV_M51	Y2H
223	ORF9	ORF25	NC CN CC	3	1	ORF25-ORF9						
224	ORF9a	ORF25	NN CN CC	3	2	ORF25-ORF9a					HSV1_UL49A-HSV1_UL33, redundant construct	Y2H red
225	ORF9aN	ORF25	NN NC	2	red 2	ORF25-ORF9a						perm-red
226	S/L	ORF25	NN	1	1	ORF25-S/L						
227	ORF26	ORF26	NN CC	2	1	ORF26-ORF26					EBV_BFLF1-EBV_BFLF1	Y2H
228	ORF28	ORF26	CC	1	1	ORF26-ORF28					KSHV_68-KSHV_9	Y2H
229	ORF29	ORF26	CC	1	1	ORF26-ORF29						
230	ORF33	ORF26	CC	1	1	ORF26-ORF33						
231	ORF26	ORF38	NN	1	1	ORF26-ORF38						
232	ORF39N	ORF26	CN	1	1	ORF26-ORF39						
233	ORF42	ORF26	CC	1	1	ORF26-ORF42						
234	ORF43	ORF26	CC	1	1	ORF26-ORF43						
235	ORF26	ORF53	CC	1	1	ORF26-ORF53						
236	ORF26	ORF56C	CC	1	1	ORF26-ORF56						
237	ORF26	ORF60C	CN CC	2	2	ORF26-ORF60					redundant direction	red
238	ORF60C	ORF26	CC	1	red 2	ORF26-ORF60						perm-red
239	ORF62	ORF26	CC	1	1	ORF26-ORF62						
240	ORF26	ORF68	CC	1	3	ORF26-ORF68					redundant construct direction	red
241	ORF26	ORF68F	CC	1	red 3	ORF26-ORF68						perm-red
242	ORF68F	ORF26	NN	1	red 3	ORF26-ORF68						perm-red
243	ORF9	ORF26	CN	1	1	ORF26-ORF9						
244	ORF26	ORF9a	CC	1	1	ORF26-ORF9a						
245	ORF26	S/L C	CC	1	1	ORF26-S/L						
246	ORF27	ORF27	NN	1	1	ORF27-ORF27						
247	ORF33	ORF27	NN CN	2	1	ORF27-ORF33						
248	ORF33.5	ORF27	NN	1	1	ORF27-ORF33.5					EBV_BFLF2-EBV_BDRF1	Y2H
249	ORF38	ORF27	NN	1	1	ORF27-ORF38						
250	ORF4	ORF27	NC	1	1	ORF27-ORF4						
251	ORF42	ORF27	NN	1	1	ORF27-ORF42						
252	ORF50C	ORF27	NN NC	2	1	ORF27-ORF50						
253	ORF27	ORF60C	NN	1	1	ORF27-ORF60						
254	ORF64	ORF27	NN	1	1	ORF27-ORF64						
255	ORF65	ORF27	NN	1	2	ORF27-ORF65					redundant construct	red
256	ORF65N	ORF27	NN	1	red 2	ORF27-ORF65						perm-red
257	ORF68	ORF27	NN	1	1	ORF27-ORF68						
258	ORF9a	ORF27	NN	1	1	ORF27-ORF9a						
259	ORF28	ORF3	NC	1	2	ORF28-ORF3					redundant direction	red
260	ORF3	ORF28	CN	1	red 2	ORF28-ORF3						perm-red
261	ORF28	ORF38	CN	1	1	ORF28-ORF38						
262	ORF4	ORF28	NC	1	1	ORF28-ORF4						
263	ORF28	ORF56C	CC	1	1	ORF28-ORF56						
264	ORF28	ORF60C	NN	1	1	ORF28-ORF60					KSHV_47-KSHV_9	Y2H
265	ORF28	ORF68C	NC	1	1	ORF28-ORF68						
266	ORF28	ORF9a	CC	1	1	ORF28-ORF9a						
267	ORF29	ORF56C	CC	1	1	ORF29-ORF56						
268	ORF29	ORF60C	CN CC	2	1	ORF29-ORF60						

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
269	ORF29	ORF65N	CC	1	1	ORF29-ORF65						
270	ORF29	ORF68F	CC	1	1	ORF29-ORF68						
271	ORF29	ORF9a	CC	1	1	ORF29-ORF9a						
272	ORF3	ORF3	NC	1	1	ORF3-ORF3						
273	ORF33	ORF3	NC	1	1	ORF3-ORF33						
274	ORF3	ORF34	NN	1	1	ORF3-ORF34						
275	ORF3	ORF38	CN	1	1	ORF3-ORF38						
276	ORF3	ORF39N	NN	1	1	ORF3-ORF39						
277	ORF3	ORF41	NC	1	1	ORF3-ORF41						
278	ORF3	ORF46	NN CN	2	1	ORF3-ORF46						
279	ORF3	ORF53	CC	1	1	ORF3-ORF53						
280	ORF3	ORF60C	NN CN	2	1	ORF3-ORF60						
281	ORF3	ORF62	NN	1	1	ORF3-ORF62						
282	ORF31	ORF45	NN	1	1	ORF31-ORF45						
283	ORF31	ORF61	CC	1	1	ORF31-ORF61						
284	ORF32	ORF38	CN	1	1	ORF32-ORF38						
285	ORF32	ORF39N	NN	1	1	ORF32-ORF39						
286	ORF9	ORF32	CN	1	1	ORF32-ORF9						
287	ORF33.5	ORF41	NC	1	1	ORF33.5-ORF41						
288	ORF50C	ORF33.5	CN	1	1	ORF33.5-ORF50					EBV_BdRF1-EBV_BBRF3	Y2H
289	ORF33.5	ORF56	NN	1	1	ORF33.5-ORF56						
290	ORF33.5	ORF60C	NN	1	1	ORF33.5-ORF60						
291	ORF33.5	ORF66	NN	1	1	ORF33.5-ORF66						
292	ORF33	ORF33	NN CN CC	3	1	ORF33-ORF33					MCMV_M80-MCMV_M80	Y2H
293	ORF33	ORF33.5	CN	1	2	ORF33-ORF33.5	HSV1_UL26.5-HSV1_UL26	8661404			redundant direction	LC red
294	ORF33.5	ORF33	NN	1	red 2	ORF33-ORF33.5						perm-red
295	ORF33	ORF41	NC	1	1	ORF33-ORF41						
296	ORF33	ORF53	CC	1	1	ORF33-ORF53						
297	ORF33	ORF56	NN	1	2	ORF33-ORF56					redundant construct	red
298	ORF33	ORF56C	CC	1	red 2	ORF33-ORF56						perm-red
299	ORF33	ORF60C	NN CN CC	3	1	ORF33-ORF60						
300	ORF33	ORF66	NN CN	2	1	ORF33-ORF66						
301	ORF33	ORF68F	CC	1	1	ORF33-ORF68						
302	ORF42	ORF34	NN	1	1	ORF34-ORF42						
303	ORF34	ORF60C	NN	1	1	ORF34-ORF60						
304	ORF8	ORF34	NN	1	1	ORF34-ORF8						
305	ORF35	ORF41	CC	1	1	ORF35-ORF41						
306	ORF65	ORF35	CC	1	1	ORF35-ORF65						
307	ORF36	ORF36	NN	1	1	ORF36-ORF36	KSHV_21-KSHV_21	18321973				LC
308	ORF36	S/L C	NN	1	2	ORF36-S/L					redundant construct	red
309	S/L	ORF36	NN	1	red 2	ORF36-S/L						perm-red
310	ORF39N	ORF37N	CC	1	1	ORF37-ORF39						
311	ORF9	ORF37N	CC	1	1	ORF37-ORF9						
312	ORF39N	ORF38	CN	1	1	ORF38-ORF39					HSV1_UL21-HSV1_UL20	Y2H

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
313	ORF4	ORF38	NN CN	2	1	ORF38-ORF4					KSHV_23-KSHV_57	Y2H
314	ORF42	ORF38	CN	1	1	ORF38-ORF42						
315	ORF38	ORF44	NN NC	2	2	ORF38-ORF44					HSV1_UL16-HSV1_UL21, redundant direction	Y2H red
316	ORF44	ORF38	NN	1	red 2	ORF38-ORF44						perm-red
317	ORF46	ORF38	NN	1	1	ORF38-ORF46						
318	ORF49	ORF38	CN	1	1	ORF38-ORF49					EBV_BBLF1-EBV_BTRF1	Y2H
319	ORF57	ORF38	CN	1	1	ORF38-ORF57						
320	ORF58	ORF38	CN	1	1	ORF38-ORF58						
321	ORF38	ORF61	CC	1	1	ORF38-ORF61						
322	ORF65N	ORF38	CN	1	1	ORF38-ORF65						
323	ORF67C	ORF38	CN	1	1	ORF38-ORF67						
324	ORF9	ORF38	CN	1	1	ORF38-ORF9						
325	ORF39	ORF39	NN	1	1	ORF39-ORF39						
326	ORF41	ORF39N	NN	1	1	ORF39-ORF41						
327	ORF42	ORF39N	NN	1	1	ORF39-ORF42						
328	ORF39N	ORF50	CN	1	1	ORF39-ORF50						
329	ORF39N	ORF57	NC	1	2	ORF39-ORF57					redundant direction	red
330	ORF57	ORF39N	CN	1	red 2	ORF39-ORF57						perm-red
331	ORF58	ORF39N	NN CN	2	1	ORF39-ORF58						
332	ORF65	ORF39N	NN	1	2	ORF39-ORF65					redundant construct	red
333	ORF65N	ORF39N	NN	1	red 2	ORF39-ORF65						perm-red
334	ORF39	ORF68C	NC	1	2	ORF39-ORF68					redundant construct	red
335	ORF68F	ORF39N	NN	1	red 2	ORF39-ORF68						perm-red
336	ORF8	ORF39N	NN	1	1	ORF39-ORF8						
337	ORF9a	ORF39N	CN	1	1	ORF39-ORF9a						
338	ORF4	ORF4	CN	1	1	ORF4-ORF4	HSV1_UL54-HSV1_UL54, KSHV_57-KSHV_57	10329545	15269354		HSV1_UL54-HSV1_UL54, MCMV_M69-MCMV_M69, KSHV_57-KSHV_57	LC Y2H
339	ORF4	ORF41	NC CN CC	3	1	ORF4-ORF41						
340	ORF4	ORF62	NN	1	1	ORF4-ORF62	HSV1_UL54-HSV1_RS1, VZV_ORF4-VZV_ORF62	8995681	10873781			LC
341	ORF4	ORF9aN	CC	1	1	ORF4-ORF9a						
342	ORF42	ORF41	NC	1	1	ORF41-ORF42						
343	ORF41	ORF56	NN	1	1	ORF41-ORF56						
344	ORF41	ORF60C	NN	1	1	ORF41-ORF60						
345	ORF42	ORF42	CN	1	1	ORF42-ORF42	HSV1_UL15-HSV1_UL15	11086131				LC
346	ORF42	ORF53	CC	1	1	ORF42-ORF53					HSV1_UL7-HSV1_UL15, MCMV_M103-MCMV_M89	Y2H
347	ORF57	ORF42	CN	1	1	ORF42-ORF57						
348	ORF42	ORF60C	NN CN CC	3	1	ORF42-ORF60						
349	ORF42	ORF68F	CC	1	1	ORF42-ORF68						
350	ORF9	ORF42	CN CC	2	1	ORF42-ORF9						
351	ORF42	ORF9a	CC	1	2	ORF42-ORF9a					HSV1_UL49A-HSV1_UL15, redundant construct	Y2H red
352	ORF9aN	ORF42	CC	1	red 2	ORF42-ORF9a						perm-red
353	ORF43	ORF56C	CC	1	1	ORF43-ORF56						
354	ORF43	ORF60C	NN	1	1	ORF43-ORF60						
355	ORF43	ORF65N	NN NC CN CC	4	2	ORF43-ORF65					redundant direction	red
356	ORF65N	ORF43	NN CN	2	red 2	ORF43-ORF65						perm-red
357	ORF43	ORF9	NC	1	1	ORF43-ORF9						
358	ORF43	ORF9a	CC	1	1	ORF43-ORF9a						

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
359	ORF44	ORF49	NC	1	2	ORF44-ORF49	HSV1_UL11-HSV1_UL16	16014918	18602131		HSV1_UL11-HSV1_UL16, MCMV_M94-MCMV_M99, EBV_BBLF1-EBV_BGLF2, redundant direction	LC Y2H red
360	ORF49	ORF44	CN	1	red 2	ORF44-ORF49						perm-red
361	ORF44	ORF53	NC	1	1	ORF44-ORF53					HSV1_UL7-HSV1_UL16	Y2H
362	ORF44	ORF61	NN	1	1	ORF44-ORF61						
363	ORF44	ORF62	NN NC	2	1	ORF44-ORF62						
364	ORF57	ORF46	CN	1	1	ORF46-ORF57						
365	ORF50C	ORF50C	NN	1	1	ORF50-ORF50						
366	ORF50C	ORF61	NN	1	1	ORF50-ORF61						
367	ORF50C	ORF62	NN	1	1	ORF50-ORF62						
368	ORF50C	ORF66	CN	1	1	ORF50-ORF66						
369	ORF50	ORF9a	CC	1	1	ORF50-ORF9a	EBV_BLRF1-EBV_BBRF3	11070013			HSV1_UL10-HSV1_UL49A	LC Y2H
370	ORF60C	ORF53	CC	1	1	ORF53-ORF60						
371	ORF53	ORF7	NN	1	1	ORF53-ORF7					EBV_BSRF1-EBV_BBRF2	Y2H
372	ORF55	ORF61	NN	1	1	ORF55-ORF61						
373	ORF60C	ORF56C	CC	1	1	ORF56-ORF60						
374	ORF62	ORF56C	CC	1	1	ORF56-ORF62						
375	ORF65	ORF56	NN	1	1	ORF56-ORF65						
376	ORF9a	ORF56	NN	1	1	ORF56-ORF9a						
377	ORF57	ORF62	CN	1	1	ORF57-ORF62						
378	ORF57	ORF9a	CN	1	1	ORF57-ORF9a						
379	ORF58	ORF60C	NN	1	1	ORF58-ORF60						
380	ORF60C	ORF60	NC	1	2	ORF60-ORF60					redundant construct	red
381	ORF60C	ORF60C	CN CC	2	red 2	ORF60-ORF60						perm-red
382	ORF62	ORF60C	CN CC	2	1	ORF60-ORF62						
383	ORF65	ORF60C	NN	1	2	ORF60-ORF65					redundant construct	red
384	ORF65N	ORF60C	NN	1	red 2	ORF60-ORF65						perm-red
385	ORF68	ORF60C	NN	1	1	ORF60-ORF68						
386	ORF9a	ORF60C	NN	1	1	ORF60-ORF9a						
387	ORF61	ORF61	NN	1	1	ORF61-ORF61	HSV1_RL2-HSV1_RL2	8151788	7966607			LC
388	ORF62	ORF68F	CC	1	1	ORF62-ORF68						
389	ORF9	ORF62	CN	1	1	ORF62-ORF9	VZV_ORF9-VZV_ORF62	17079304				LC
390	ORF62	ORF9a	CC	1	2	ORF62-ORF9a					redundant construct	red
391	ORF62	ORF9aN	CN	1	red 2	ORF62-ORF9a						perm-red
392	S/L C	ORF62	NN	1	1	ORF62-S/L						
393	ORF63	ORF63	CC	1	1	ORF63-ORF63						
394	ORF64	ORF64	NC	1	1	ORF64-ORF64						
395	ORF65	ORF64	CC	1	1	ORF64-ORF65						
396	ORF68C	ORF64	CN	1	1	ORF64-ORF68						
397	ORF9	ORF64	CN	1	1	ORF64-ORF9						
398	ORF67C	ORF67C	CN	1	1	ORF67-ORF67						
399	ORF67C	ORF68C	CC	1	1	ORF67-ORF68	HSV1_US7-HSV1_US8, VZV_ORF67-VZV_ORF68	7995945	18945783			LC
400	ORF8	ORF67C	NN	1	1	ORF67-ORF8						
401	ORF67C	ORF9a	CC	1	1	ORF67-ORF9a						
402	ORF7	ORF7	NN	1	1	ORF7-ORF7						
403	ORF8	ORF8	CN	1	1	ORF8-ORF8						
404	ORF9	ORF9aN	CN CC	2	1	ORF9-ORF9a						

Interaction #	VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID			Y2H_interolog	verification
405	ORF9a	ORF9a	CC	1	1	ORF9a-ORF9a					HSV1_UL49A-HSV1_UL49A	Y2H
		26	LC	(incl. 11 interologs)								
		62	interologs									
column	explanation											
VZV_bait	bait protein											
VZV_prey	prey protein											
permutation	configuration in which this interaction was found (either NN, NC, CN, CC, or subsets thereof)											
perm_count	number of permutations this protein pair was found in, e.g. 2 when found as NN and CC											
PPIs	number of interactions a certain protein pair was found in. For instance, ORF1 and ORF60 are found in 2 different construct pairs, ORF1-ORF60C and ORF1N-ORF60C, shown on 2 different lines)											
VZV_pairs	all pairs of interacting proteins listed so that the smaller ORF number is always shown first, in order to allow searches for redundant interactions. For example, see rows 7 and 8.											
LC_interologs	Interologs of published interactions											
PMID	PubMed IDs of LC_interologs											
Y2H_interolog	Interologs from large-scale Y2H screens conducted by our lab (Uetz et al. 2006 and Fossum et al. 2009)											
	This column also indicates verified interactions from this study that have been found in additional constructs or directions (such as inverted bait/preys)											
verification	Source of verification data: LC = literature-curated data from Fossum et al., Y2H = Y2H data from Fossum et al., red = redundant data from this study (additional constructs/directions etc.)											
	Interactions that are found by a different permutation are labelled "perm-red".											
VZV_bait	bait protein											

Supplementary Table S1: Y2H data - all interactions detected in combinatorial Y2H-screens.

Note that this table is redundant in certain ways. The list contains a total of 405 interactions which are all different in terms of interacting proteins. For example, the two interactions between ORF1 and ORF60C (Interactions # 6 and 7) are listed separately because they involve different constructs (full-length ORF1 in # 6 and an N-terminal fragment of ORF1 in # 7). Such redundancies are indicated by "red" in the column PPIs. All interacting protein pairs irrespective of constructs are listed in VZV_pairs. This column is redundant because the same protein pairs may occur multiple times (as in interactions 6 and 7). Y2H signals that were clearly contaminations or clearly unspecific (i.e. appearing with many unrelated baits) were considered as false-positives and left out from this table. The protein interactions from this study can also be found in the IntAct database of the European Bioinformatics Institute. The interaction data is provided by the IMEx-consortium (http://imex.sf.net) through IntAct (http://www.ebi.ac.uk/intact/main.xhtml?conversationContext=1) and was given the Identifier IM-117118.

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
VZV	ORF1N	ORF25	Uetz et al.; this study	HSV-1	UL5	UL52	10075707;10501495	10075707;10501495	LC Fossum et al.
VZV	ORF1N	ORF27	Uetz et al.	HSV-1	UL30	UL42	10998337;11333878;15157875;2173776;7964622;8380085;8380091;8382792;8396660	10998337;11333878;15157875;2173776;7964622;8380085;8380091;8382792;8396660	LC Fossum et al.
VZV	ORF1	ORF60C	this study	HSV-1	UL9	UL29	12758170;8397405	12758170;8397405	LC Fossum et al.
VZV	ORF1N	ORF60C	Uetz et al.	HSV-1	UL19	UL26	10497121;8661404	10497121;De Sai et al. (LS)	LC Fossum et al.
VZV	ORF1N	ORF62	Uetz et al.	HSV-1	UL19	UL26.5	10497121;8523566;8661404;8995652;9188587	10497121;8523566;De Sai et al. (LS);8995652;9188587	LC Fossum et al.
VZV	ORF10	ORF32	this study	HSV-1	UL52	UL8	10501495;7931156;9344911	10501495;7931156;9344911	LC Fossum et al.
VZV	ORF10	ORF56C	this study	HSV-1	UL34	UL31	10627546;11507225;15731273	Ye et al. (SS);Reynolds et al. (SS);Liang et al. (SS)	LC Fossum et al.
VZV	ORF11	ORF13	this study	HSV-1	UL34	UL19	10627546	Ye et al. (SS)	LC Fossum et al.
VZV	ORF11	ORF27	Uetz et al.	HSV-1	UL15	UL28	11086131;17035316	Abbotts et al. (SS);Jacobson et al. (SS)	LC Fossum et al.
VZV	ORF11	ORF38	Uetz et al.	HSV-1	UL15	UL15	11086131	Abbotts et al. (SS)	LC Fossum et al.
VZV	ORF12	ORF12	this study	HSV-1	UL28	UL33	17035316	Jacobson et al. (SS)	LC Fossum et al.
VZV	ORF12	ORF16	this study	HSV-1	UL25	UL19	11152516	11152516	LC Fossum et al.
VZV	ORF12	ORF24N	this study	HSV-1	UL25	UL38	11152516	11152516	LC Fossum et al.
VZV	ORF12C	ORF25	this study	HSV-1	US7	US8	7995945	Basu et al. (SS)	LC Fossum et al.
VZV	ORF12	ORF25	Uetz et al.; this study	HSV-1	UL26.5	UL26.5	8661404;9188587	De Sai et al. (LS);9188587	LC Fossum et al.
VZV	ORF12	ORF27	this study	HSV-1	UL26.5	UL26	8661404	De Sai et al. (LS)	LC Fossum et al.
VZV	ORF12C	ORF37N	this study	HSV-1	UL41	UL48	8139019;8642633;12584313;15564467	8139019;8642633;12584313;15564467	LC Fossum et al.
VZV	ORF12	ORF38	this study	HSV-1	UL6	UL15	12743292	12743292	LC Fossum et al.
VZV	ORF12	ORF39	this study	HSV-1	UL6	UL28	12743292	12743292	LC Fossum et al.
VZV	ORF12C	ORF42	this study	HSV-1	UL39	UL40	1322407;3012359	Filatov et al. (SS);Dutia et al. (SS)	LC Fossum et al.
VZV	ORF12	ORF42	this study	HSV-1	UL9	UL42	9454723	9454723	LC Fossum et al.
VZV	ORF12	ORF60C	Uetz et al.; this study	HSV-1	UL9	UL8	7931156	7931156	LC Fossum et al.
VZV	ORF12	ORF9	this study	HSV-1	UL9	UL29	8397405	8397405	LC Fossum et al.
VZV	ORF14	ORF35	this study	HSV-1	UL8	UL52	7931156;9344911	7931156;9344911	LC Fossum et al.
VZV	ORF15F	ORF25	Uetz et al.	HSV-1	UL48	UL22	14675620;15857998	14675620;15857998	LC Fossum et al.
VZV	ORF15N	ORF25	Uetz et al.; this study	HSV-1	UL12	UL29	15078942	15078942	LC Fossum et al.
VZV	ORF15	ORF42	this study	HSV-1	UL12.5	UL29	15078942	15078942	LC Fossum et al.
VZV	ORF15F	ORF60C	Uetz et al.	HSV-1	UL54	UL29	15582656	15582656	LC Fossum et al.
VZV	ORF16	ORF16	Uetz et al.; this study	HSV-1	UL6	UL26.5	12941896	12941896	LC Fossum et al.
VZV	ORF16	ORF38	this study	HSV-1	US6	UL49	15659744	15659744	LC Fossum et al.
VZV	ORF16	ORF62	Uetz et al.; this study	HSV-1	UL11	UL16	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18C	ORF18C	this study	HSV-1	UL36	UL37	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18C	ORF18N	this study	HSV-1	UL36	UL48	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18C	ORF19	Uetz et al.; this study	HSV-1	UL46	UL48	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18C	ORF23	this study	HSV-1	UL47	UL48	16014918	Vittone et al. (LS)	LC Fossum et al.
VZV	ORF18N	ORF23	Uetz et al.; this study	HSV-1	UL49	UL48	16014918;16189010;18602131;16160145	Vittone et al. (LS);16189010;Lee et al. (LS);16160145	LC Fossum et al.
VZV	ORF18C	ORF24N	this study	HSV-1	US11	US11	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18N	ORF25	this study	HSV-1	UL37	UL37	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18	ORF25	Uetz et al.; this	HSV-1	UL49	UL49	16014918;18602131	Vittone et al. (LS);Lee et al. (LS)	LC Fossum et al.

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
			study						
VZV	ORF18C	ORF25	Uetz et al.; this study	HSV-1	UL5	UL8	7931156;9344911	7931156;9344911	LC Fossum et al.
VZV	ORF18C	ORF26	this study	HSV-1	UL27	UL27	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF18C	ORF37N	this study	HSV-1	UL27	UL48	7941326	7941326	LC Fossum et al.
VZV	ORF18C	ORF41	this study	HSV-1	RS1	RL2	7966607	Yao et al. (SS)	LC Fossum et al.
VZV	ORF18C	ORF53	this study	HSV-1	RL2	RL2	8151788;7966607	Ciufo et al. (SS);Yao et al. (SS)	LC Fossum et al.
VZV	ORF18C	ORF56C	this study	HSV-1	UL18	UL38	8661404;18602131	De Sai et al. (LS);Lee et al. (LS)	LC Fossum et al.
VZV	ORF18C	ORF58	this study	HSV-1	UL54	RS1	8995681	Panagiotidis et al. (SS)	LC Fossum et al.
VZV	ORF18C	ORF60C	this study	HSV-1	UL30	UL8	9261356	9261356	LC Fossum et al.
VZV	ORF18C	ORF68	this study	HSV-1	UL8	UL29	9278436	9278436	LC Fossum et al.
VZV	ORF18C	ORF68F	this study	HSV-1	UL22	UL1	10441558	10441558	LC Fossum et al.
VZV	ORF18C	ORF9a	this study	HSV-1	UL13	US8	9454715	9454715	LC Fossum et al.
VZV	ORF18C	S/L C	this study	HSV-1	UL18	UL18	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF16	Uetz et al.; this study	HSV-1	UL35	UL19	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF18C	Uetz et al.; this study	HSV-1	UL54	UL54	10329545	Zhi et al. (SS)	LC Fossum et al.
VZV	ORF19	ORF19	Uetz et al.; this study	HSV-1	UL34	US3	10627546	Ye et al. (SS)	LC Fossum et al.
VZV	ORF19	ORF24N	this study	HSV-1	UL35	UL18	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF25	Uetz et al.; this study	HSV-1	UL35	UL25	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF26	this study	HSV-1	UL35	UL38	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF27	Uetz et al.	HSV-1	UL19	UL16	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF38	Uetz et al.; this study	HSV-1	UL19	UL21	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF39N	this study	HSV-1	UL19	UL48	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF41	this study	HSV-1	UL35	UL11	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF42	this study	HSV-1	UL35	UL14	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF48	this study	HSV-1	UL35	UL16	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF53	this study	HSV-1	UL35	UL21	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF57	this study	HSV-1	UL35	UL37	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF58	this study	HSV-1	UL35	UL48	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF60C	this study	HSV-1	UL35	UL51	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF62	Uetz et al.; this study	HSV-1	UL35	US3	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF68F	this study	HSV-1	UL37	UL38	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF19	ORF9a	this study	HSV-1	UL46	UL18	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF2	ORF25	Uetz et al.; this study	HSV-1	UL46	UL19	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF2	ORF56C	this study	HSV-1	UL46	UL25	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF21	ORF22N	Uetz et al.	HSV-1	UL46	UL38	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF21	ORF23	Uetz et al.	HSV-1	UL4	UL56	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF21	ORF27	Uetz et al.	HSV-1	UL46	UL21	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF21	ORF34	Uetz et al.	HSV-1	UL46	UL37	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF21	ORF42	this study	HSV-1	UL46	UL49	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF21	ORF60C	Uetz et al.	HSV-1	UL46	US3	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF23	ORF18N	this study	HSV-1	UL46	US10	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF23	ORF25	this study	HSV-1	UL49	UL48	18602131	Lee et al. (LS)	LC Fossum et al.

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
VZV	ORF23	ORF27	Uetz et al.	HSV-1	UL46	UL10	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF23	ORF39N	Uetz et al.	HSV-1	US9	UL49	18602131	Lee et al. (LS)	LC Fossum et al.
VZV	ORF23	ORF60C	Uetz et al.	EBV	BSLF1	BBLF4	10580049	10580049	LC Fossum et al.
VZV	ORF23	ORF9	this study	EBV	BSLF1	BBLF2/BBLF3	10580049	10580049	LC Fossum et al.
VZV	ORF24N	ORF1N	this study	EBV	BBLF4	BBLF2/BBLF3	10580049	10580049	LC Fossum et al.
VZV	ORF24N	ORF15	this study	EBV	BLRF1	BBRF3	11070013	Lake et al. (SS)	LC Fossum et al.
VZV	ORF24N	ORF16	Uetz et al.; this study	EBV	BZLF1	BBLF4	11507224;9765394	11507224;9765394	LC Fossum et al.
VZV	ORF24N	ORF19	this study	EBV	BZLF1	BBLF2/BBLF3	9765394	9765394	LC Fossum et al.
VZV	ORF24N	ORF21	this study	EBV	BFRF1	BFLF2	15003866;15731265;17446270	Lake et al. (SS);Gonnella et al. (SS);Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF24N	ORF22N	Uetz et al.	EBV	BGLF5	BBLF2/BBLF3	15596820	15596820	LC Fossum et al.
VZV	ORF24N	ORF23	this study	EBV	BZLF1	BMRF1	8764021	8764021	LC Fossum et al.
VZV	ORF24N	ORF24N	this study	EBV	BSLF2	BMLF1	9765385	9765385	LC Fossum et al.
VZV	ORF24N	ORF25	Uetz et al.; this study	EBV	EBNA1	EBNA1	11684888;17446270	11684888;Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF24	ORF27	Uetz et al.; this study	EBV	BZLF1	BZLF1	10727769	10727769	LC Fossum et al.
VZV	ORF24N	ORF27	Uetz et al.; this study	EBV	BALF5	BBLF2/BBLF3	10684269	10684269	LC Fossum et al.
VZV	ORF24N	ORF34	Uetz et al.	EBV	BALF5	BSLF1	10684269	10684269	LC Fossum et al.
VZV	ORF24N	ORF38	this study	EBV	BALF5	BBLF4	10684269	10684269	LC Fossum et al.
VZV	ORF24	ORF39N	Uetz et al.	EBV	BXLF2	BZLF2	7539502	7539502	LC Fossum et al.
VZV	ORF24N	ORF39N	Uetz et al.; this study	EBV	BCRF1	BCRF1	9159483	9159483	LC Fossum et al.
VZV	ORF24N	ORF41	this study	EBV	EBNA-LP	EBNA-LP	11773378;14732686	11773378;14732686	LC Fossum et al.
VZV	ORF24N	ORF42	this study	EBV	BMRF1	BMRF1	15286084	Makhov et al. (SS)	LC Fossum et al.
VZV	ORF24	ORF46	this study	EBV	BGLF5	BMRF1	8396819	8396819	LC Fossum et al.
VZV	ORF24N	ORF52	Uetz et al.	EBV	EBNA-LP	EBNA2	14732686	14732686	LC Fossum et al.
VZV	ORF24N	ORF53	this study	EBV	BRLF1	BRLF1	1645863	1645863	LC Fossum et al.
VZV	ORF24N	ORF58	this study	EBV	EBNA2	EBNA2	8207803	8207803	LC Fossum et al.
VZV	ORF24N	ORF62	Uetz et al.; this study	EBV	BaRF1	BaRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF24N	ORF64	this study	EBV	BcLF1	BFRF3	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF24N	ORF66	this study	EBV	BGLF2	BSLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF24N	ORF9a	this study	EBV	BGLF4	BXLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF25	ORF25	Uetz et al.; this study	EBV	BNRF1	BLRF2	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF25	ORF31N	this study	EBV	LF2	BRLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF25	ORF38	Uetz et al.	EBV	LF2	LF2	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF25	ORF39N	Uetz et al.	EBV	BNRF2	BLRF3	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF18C	this study	EBV	BPLF1	BaRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF18N	this study	EBV	BPLF1	BPLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF19	this study	EBV	EBNA2	BDLF2	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF23	this study	EBV	EBNA2	EBNA3A	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF24N	this study	EBV	BZLF1	EBNA2	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF25	this study	EBV	EBNA3A	EBNA3A	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF26	Uetz et al.; this	EBV	BORF2	BcLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
			study						
VZV	ORF26	ORF38	Uetz et al.	EBV	BRRF2	LMP2A	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF53	this study	EBV	EBNA3A	EBNA3C	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF56C	this study	EBV	EBNA3C	BXLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF60C	this study	EBV	BNLF2a	BNLF2a	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF68	this study	EBV	BORF2	BALF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF68F	this study	EBV	BXLF1	BILF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	ORF9a	this study	EBV	BXLF1	BRLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF26	S/L C	this study	EBV	BBRF1	BSRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF27	ORF21	Uetz et al.; this study	EBV	BdRF1	BALF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF27	ORF24N	this study	EBV	BNRF1	BaRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF27	ORF24	Uetz et al.	EBV	BZLF2	BTRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF27	ORF25	this study	EBV	LMP1	BKRF2	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF27	ORF27	Uetz et al.	EBV	BILF1	EBNA1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF27	ORF60C	Uetz et al.	EBV	BLLF1	BBLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF18C	this study	EBV	BLLF2	EBNA3B	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF18N	this study	EBV	LMP1	BLLF2	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF25	this study	EBV	LMP1	BLRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF26	this study	EBV	BaRF1	BLRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF3	this study	EBV	BILF1	BaRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF38	this study	EBV	BILF1	BXRF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF56C	this study	EBV	BMRF1	EBNA3B	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF60C	Uetz et al.	EBV	BNLF2b	BDLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF68C	this study	EBV	EBNA3B	BALF4	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF28	ORF9a	this study	EBV	EBNA3B	BDLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF29	ORF18C	this study	EBV	EBNA3C	EBNA1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF29	ORF18N	this study	EBV	LMP1	BcLF1	17446270	Calderwood et al. (LS)	LC Fossum et al.
VZV	ORF29	ORF25	this study	MCMV	M123	M112/113	15596821	15596821	LC Fossum et al.
VZV	ORF29	ORF26	this study	MCMV	m139	m141	12719548	12719548	LC Fossum et al.
VZV	ORF29	ORF56C	this study	MCMV	m142	m143	17005694	17005694	LC Fossum et al.
VZV	ORF29	ORF60C	this study	MCMV	M53	M50	17005637	Schnee et al. (SS)	LC Fossum et al.
VZV	ORF29	ORF65N	this study	MCMB	M45	M48	19244336	19244336	LC Whitehurst
VZV	ORF29	ORF68F	this study	VZV	VZV_ORF63	VZV_ORF62	11483768;14722273;12429517	11483768;14722273;12429517	LC Fossum et al.
VZV	ORF29	ORF9a	this study	VZV	VZV_ORF37	VZV_ORF60	7618278	7618278	LC Fossum et al.
VZV	ORF3	ORF1N	this study	VZV	VZV_ORF62	VZV_ORF66	16439528	16439528	LC Fossum et al.
VZV	ORF3	ORF16	Uetz et al.	VZV	VZV_ORF67	VZV_ORF68	18945783	Berarducci et al. (SS)	LC Fossum et al.
VZV	ORF3	ORF21	this study	VZV	VZV_ORF68	VZV_ORF68	18945783	Berarducci et al. (SS)	LC Fossum et al.
VZV	ORF3	ORF23	this study	VZV	VZV_ORF4	VZV_ORF62	10873781	Spengler et al. (SS)	LC Fossum et al.
VZV	ORF3	ORF25	Uetz et al.; this study	VZV	VZV_ORF9	VZV_ORF62	17079304	Cilloniz et al. (SS)	LC Fossum et al.
VZV	ORF3	ORF28	this study	VZV	VZV_ORF37	VZV_ORF68	14990707	14990707	LC Fossum et al.
VZV	ORF3	ORF3	this study	VZV	VZV_ORF42	VZV_ORF30	17868947	17868947	LC Fossum et al.

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
VZV	ORF3	ORF34	Uetz et al.	KSHV	59	9	11069986	11069986	LC Fossum et al.
VZV	ORF3	ORF38	this study	KSHV	K8	50	12502859	12502859	LC Fossum et al.
VZV	ORF3	ORF39N	Uetz et al.	KSHV	25	65	12634386	12634386	LC Fossum et al.
VZV	ORF3	ORF41	this study	KSHV	73	50	15163750	15163750	LC Fossum et al.
VZV	ORF3	ORF46	Uetz et al.; this study	KSHV	57	50	15269354	Malik et al. (SS)	LC Fossum et al.
VZV	ORF3	ORF53	this study	KSHV	59	59	15075322	Chen et al. (SS)	LC Fossum et al.
VZV	ORF3	ORF60C	Uetz et al.; this study	KSHV	50	50	12915555	12915555	LC Fossum et al.
VZV	ORF3	ORF62	Uetz et al.	KSHV	K12	K12	11336706	11336706	LC Fossum et al.
VZV	ORF30	ORF25	Uetz et al.	KSHV	57	57	15269354	Malik et al. (SS)	LC Fossum et al.
VZV	ORF31C	ORF25	this study	KSHV	K8	K8	15919946	15919946	LC Fossum et al.
VZV	ORF31	ORF45	Uetz et al.	KSHV	11	21	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF31	ORF61	this study	KSHV	11	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF32	ORF25	Uetz et al.; this study	KSHV	11	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF32	ORF38	this study	KSHV	21	25	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF32	ORF39N	Uetz et al.	KSHV	21	62	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF12C	Uetz et al.	KSHV	21	21	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF16	Uetz et al.	KSHV	21	64	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF18N	this study	KSHV	21	75	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF21	Uetz et al.; this study	KSHV	21	22	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF22N	Uetz et al.	KSHV	21	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF23	Uetz et al.; this study	KSHV	21	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF25	Uetz et al.; this study	KSHV	27	39	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF26	this study	KSHV	27	22	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF27	Uetz et al.; this study	KSHV	27	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF3	this study	KSHV	27	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF33	Uetz et al.; this study	KSHV	33	52	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF33.5	this study	KSHV	33	28	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF41	this study	KSHV	33	39	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF53	this study	KSHV	33	22	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF56C	this study	KSHV	33	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF56	Uetz et al.	KSHV	45	26	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF60C	Uetz et al.; this study	KSHV	45	62	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF66	Uetz et al.; this study	KSHV	45	65	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33	ORF68F	this study	KSHV	45	11	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF12C	Uetz et al.	KSHV	45	21	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF18C	Uetz et al.	KSHV	45	27	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF18N	Uetz et al.; this study	KSHV	45	33	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF21	Uetz et al.; this study	KSHV	45	63	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF22N	Uetz et al.	KSHV	45	64	18321973	Rozen et al. (LS)	LC Fossum et al.

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
VZV	ORF33.5	ORF25	Uetz et al.; this study	KSHV	45	75	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF27	Uetz et al.	KSHV	45	39	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF33	Uetz et al.	KSHV	45	22	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF41	this study	KSHV	45	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF56	Uetz et al.	KSHV	45	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF60C	Uetz et al.	KSHV	52	26	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF33.5	ORF66	Uetz et al.	KSHV	52	45	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF34	ORF60C	Uetz et al.	KSHV	52	64	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF35	ORF41	this study	KSHV	52	75	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF36	ORF25	Uetz et al.; this study	KSHV	52	39	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF36	ORF36	Uetz et al.	KSHV	52	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF36	S/L C	Uetz et al.	KSHV	52	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF37N	ORF16	this study	KSHV	63	21	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF38	ORF16	this study	KSHV	63	45	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF38	ORF22N	Uetz et al.	KSHV	63	64	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF38	ORF27	Uetz et al.	KSHV	63	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF38	ORF44	Uetz et al.; this study	KSHV	63	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF38	ORF61	this study	KSHV	64	25	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF17	this study	KSHV	64	26	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF23	this study	KSHV	64	62	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF25	Uetz et al.; this study	KSHV	64	11	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF26	this study	KSHV	64	33	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF37N	this study	KSHV	64	64	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF38	this study	KSHV	64	75	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39	ORF39	Uetz et al.	KSHV	64	28	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF50	this study	KSHV	64	39	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39N	ORF57	this study	KSHV	64	22	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF39	ORF68C	this study	KSHV	64	47	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF4	ORF20	this study	KSHV	75	62	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF4	ORF27	this study	KSHV	75	39	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF4	ORF28	this study	KSHV	75	53	18321973	Rozen et al. (LS)	LC Fossum et al.
VZV	ORF4	ORF38	Uetz et al.; this study						
VZV	ORF4	ORF4	this study						
VZV	ORF4	ORF41	this study						
VZV	ORF4	ORF62	Uetz et al.						
VZV	ORF4	ORF9aN	this study						
VZV	ORF41	ORF18C	Uetz et al.						
VZV	ORF41	ORF22N	Uetz et al.						
VZV	ORF41	ORF25	Uetz et al.; this study						
VZV	ORF41	ORF39N	Uetz et al.						
VZV	ORF41	ORF56	Uetz et al.						
VZV	ORF41	ORF60C	Uetz et al.						
VZV	ORF42	ORF16	Uetz et al.						

Y2H Interaction Set				Literature-curated Interaction Set					
Species	Bait_ID	Prey_ID	Reference	Species	ORFA_ID	ORFB_ID	LC Reference (PMID)	Reference (SS: Small Scale Study; LS: Large Scale Study)	Reference Set
VZV	ORF42	ORF21	Uetz et al.; this study						
VZV	ORF42	ORF24N	this study						
VZV	ORF42	ORF25	Uetz et al.; this study						
VZV	ORF42	ORF26	this study						
VZV	ORF42	ORF27	Uetz et al.						
VZV	ORF42	ORF34	Uetz et al.						
VZV	ORF42	ORF38	this study						
VZV	ORF42	ORF39N	Uetz et al.						
VZV	ORF42	ORF41	this study						
VZV	ORF42	ORF42	this study						
VZV	ORF42	ORF53	this study						
VZV	ORF42	ORF60C	Uetz et al.; this study						
VZV	ORF42	ORF68F	this study						
VZV	ORF42	ORF9a	this study						
VZV	ORF43	ORF1N	this study						
VZV	ORF43	ORF18N	this study						
VZV	ORF43	ORF19	Uetz et al.						
VZV	ORF43	ORF21	this study						
VZV	ORF43	ORF22N	Uetz et al.						
VZV	ORF43	ORF25	this study						
VZV	ORF43	ORF26	this study						
VZV	ORF43	ORF56C	this study						
VZV	ORF43	ORF60C	Uetz et al.						
VZV	ORF43	ORF65N	Uetz et al.; this study						
VZV	ORF43	ORF9	this study						
VZV	ORF43	ORF9a	this study						
VZV	ORF44	ORF25	Uetz et al.						
VZV	ORF44	ORF38	Uetz et al.						
VZV	ORF44	ORF49	this study						

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
VZV	ORF44	ORF53	this study	VZV	ORF62	ORF9a	this study	VZV	ORF9	ORF25	this study
VZV	ORF44	ORF61	Uetz et al.	VZV	ORF62	ORF9aN	this study	VZV	ORF9	ORF26	this study
VZV	ORF44	ORF62	Uetz et al.; this study	VZV	ORF63	ORF63	this study	VZV	ORF9	ORF32	this study
VZV	ORF46	ORF38	Uetz et al.	VZV	ORF64	ORF14N	this study	VZV	ORF9	ORF37N	this study
VZV	ORF48	ORF25	this study	VZV	ORF64	ORF16	Uetz et al.; this study	VZV	ORF9	ORF38	this study
VZV	ORF49	ORF25	Uetz et al.; this study	VZV	ORF64	ORF21	this study	VZV	ORF9	ORF42	this study
VZV	ORF49	ORF38	this study	VZV	ORF64	ORF25	Uetz et al.; this study	VZV	ORF9	ORF62	this study
VZV	ORF49	ORF44	this study	VZV	ORF64	ORF27	Uetz et al.	VZV	ORF9	ORF64	this study
VZV	ORF50C	ORF25	this study	VZV	ORF64	ORF64	this study	VZV	ORF9	ORF9aN	this study
VZV	ORF50	ORF25	Uetz et al.; this study	VZV	ORF65	ORF12C	Uetz et al.	VZV	ORF9a	ORF22N	Uetz et al.
VZV	ORF50C	ORF27	Uetz et al.; this study	VZV	ORF65	ORF18C	Uetz et al.	VZV	ORF9a	ORF23	Uetz et al.
VZV	ORF50C	ORF33.5	this study	VZV	ORF65	ORF18N	Uetz et al.	VZV	ORF9a	ORF25	Uetz et al.; this study
VZV	ORF50C	ORF50C	Uetz et al.	VZV	ORF65	ORF21	Uetz et al.	VZV	ORF9aN	ORF25	Uetz et al.; this study
VZV	ORF50C	ORF61	Uetz et al.	VZV	ORF65	ORF22N	Uetz et al.	VZV	ORF9a	ORF27	Uetz et al.
VZV	ORF50C	ORF62	Uetz et al.	VZV	ORF65	ORF23	Uetz et al.	VZV	ORF9a	ORF39N	this study
VZV	ORF50C	ORF66	this study	VZV	ORF65	ORF25	Uetz et al.; this study	VZV	ORF9aN	ORF42	this study
VZV	ORF50	ORF9a	this study	VZV	ORF65N	ORF25	Uetz et al.; this study	VZV	ORF9a	ORF56	Uetz et al.
VZV	ORF51	ORF25	Uetz et al.	VZV	ORF65	ORF27	Uetz et al.	VZV	ORF9a	ORF60C	Uetz et al.
VZV	ORF52	ORF25	Uetz et al.; this study	VZV	ORF65N	ORF27	Uetz et al.	VZV	ORF9a	ORF9a	this study
VZV	ORF53	ORF7	Uetz et al.	VZV	ORF65	ORF35	this study	VZV	S/L	ORF25	Uetz et al.
VZV	ORF55	ORF25	Uetz et al.; this study	VZV	ORF65N	ORF38	this study	VZV	S/L	ORF36	Uetz et al.
VZV	ORF55	ORF61	Uetz et al.	VZV	ORF65	ORF39N	Uetz et al.	VZV	S/L C	ORF62	Uetz et al.
VZV	ORF57	ORF10	Uetz et al.; this study	VZV	ORF65N	ORF39N	Uetz et al.	VZV	ORF9	ORF25	this study
VZV	ORF57	ORF18N	this study	VZV	ORF65N	ORF43	Uetz et al.; this study	VZV	ORF9	ORF26	this study
VZV	ORF57	ORF24N	this study	VZV	ORF65	ORF56	Uetz et al.	VZV	ORF9	ORF32	this study
VZV	ORF57	ORF25	Uetz et al.; this study	VZV	ORF65	ORF60C	Uetz et al.	VZV	ORF9	ORF37N	this study
VZV	ORF57	ORF38	this study	VZV	ORF65N	ORF60C	Uetz et al.	VZV	ORF9	ORF38	this study
VZV	ORF57	ORF39N	this study	VZV	ORF65	ORF64	this study	VZV	ORF9	ORF42	this study
VZV	ORF57	ORF42	this study	VZV	ORF67C	ORF24N	this study	VZV	ORF9	ORF62	this study
VZV	ORF57	ORF46	this study	VZV	ORF67	ORF25	this study	VZV	ORF9	ORF64	this study
VZV	ORF57	ORF62	this study	VZV	ORF67C	ORF25	this study	VZV	ORF9	ORF9aN	this study
VZV	ORF57	ORF9a	this study	VZV	ORF67C	ORF38	this study	VZV	ORF9a	ORF22N	Uetz et al.
VZV	ORF58	ORF16	Uetz et al.	VZV	ORF67C	ORF67C	this study	VZV	ORF9a	ORF23	Uetz et al.
VZV	ORF58	ORF18N	Uetz et al.; this study	VZV	ORF67C	ORF68C	this study	VZV	ORF9a	ORF25	Uetz et al.; this study
VZV	ORF58	ORF25	Uetz et al.; this study	VZV	ORF67C	ORF9a	this study	VZV	ORF9aN	ORF25	Uetz et al.; this study
VZV	ORF58	ORF38	this study	VZV	ORF68F	ORF15	this study	VZV	ORF9a	ORF27	Uetz et al.
VZV	ORF58	ORF39N	Uetz et al.; this study	VZV	ORF68	ORF22N	Uetz et al.	VZV	ORF9a	ORF39N	this study
VZV	ORF58	ORF60C	Uetz et al.	VZV	ORF68	ORF23	Uetz et al.	VZV	ORF9aN	ORF42	this study
VZV	ORF59	ORF25	this study	VZV	ORF68F	ORF25	Uetz et al.; this study	VZV	ORF9a	ORF56	Uetz et al.
VZV	ORF6	ORF13	this study	VZV	ORF68C	ORF25	Uetz et al.; this study	VZV	ORF9a	ORF60C	Uetz et al.
VZV	ORF6	ORF25	this study	VZV	ORF68F	ORF26	Uetz et al.	VZV	ORF9a	ORF9a	this study
VZV	ORF60C	ORF23	this study	VZV	ORF68	ORF27	Uetz et al.	VZV	S/L	ORF25	Uetz et al.
VZV	ORF60C	ORF24N	this study	VZV	ORF68F	ORF39N	Uetz et al.	VZV	S/L	ORF36	Uetz et al.
VZV	ORF60C	ORF25	this study	VZV	ORF68	ORF60C	Uetz et al.	VZV	S/L C	ORF62	Uetz et al.
VZV	ORF60C	ORF26	this study	VZV	ORF68C	ORF64	this study	VZV	ORF9	ORF25	this study
VZV	ORF60C	ORF53	this study	VZV	ORF7	ORF23	**Uetz et al.**	**VZV**	**ORF9**	**ORF26**	**this study**
VZV	**ORF60C**	**ORF56C**	**this study**	**VZV**	**ORF7**	**ORF7**	**Uetz et al.**	**VZV**	**ORF9**	**ORF32**	**this study**

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
VZV	ORF60C	ORF60C	this study	VZV	ORF8	ORF16	this study	VZV	ORF9	ORF37N	this study
VZV	ORF60C	ORF60	this study	VZV	ORF8	ORF19	Uetz et al.	VZV	ORF9	ORF38	this study
VZV	ORF61	ORF61	Uetz et al.	VZV	ORF8	ORF25	Uetz et al.; this study	VZV	ORF9	ORF42	this study
VZV	ORF62	ORF18C	this study	VZV	ORF8	ORF34	Uetz et al.	VZV	ORF9	ORF62	this study
VZV	ORF62	ORF25	this study	VZV	ORF8	ORF39N	Uetz et al.	VZV	ORF9	ORF64	this study
VZV	ORF62	ORF26	this study	VZV	ORF8	ORF67C	Uetz et al.	VZV	ORF9	ORF9aN	this study
VZV	ORF62	ORF56C	this study	VZV	ORF8	ORF8	this study	VZV	ORF9a	ORF22N	Uetz et al.
VZV	ORF62	ORF60C	this study	VZV	ORF9	ORF15	this study	VZV	ORF9a	ORF23	Uetz et al.
VZV	ORF62	ORF68F	this study	VZV	ORF9	ORF23	this study	VZV	ORF9a	ORF25	Uetz et al.; this study
HSV-1	RL2	UL49	Fossum et al.	HSV-1	UL40	UL37	Fossum et al.	MCMV	M1	M88	Fossum et al.
HSV-1	UL2	UL12	Fossum et al.	HSV-1	UL40	UL40	Fossum et al.	MCMV	M3	M9	Fossum et al.
HSV-1	UL2	UL14	Fossum et al.	HSV-1	UL40	UL53	Fossum et al.	MCMV	M3	M10	Fossum et al.
HSV-1	UL2	UL33	Fossum et al.	HSV-1	UL43	UL33	Fossum et al.	MCMV	M3	M50	Fossum et al.
HSV-1	UL2	UL37	Fossum et al.	HSV-1	UL43	UL40	Fossum et al.	MCMV	M3	M55	Fossum et al.
HSV-1	UL2	UL40	Fossum et al.	HSV-1	UL43	UL45	Fossum et al.	MCMV	M3	M72	Fossum et al.
HSV-1	UL2	UL53	Fossum et al.	HSV-1	UL43	UL49A	Fossum et al.	MCMV	M3	M80	Fossum et al.
HSV-1	UL3	UL4	Fossum et al.	HSV-1	UL43	UL53	Fossum et al.	MCMV	M3	M106	Fossum et al.
HSV-1	UL7	UL7	Fossum et al.	HSV-1	UL43	US2	Fossum et al.	MCMV	M3	M119.2	Fossum et al.
HSV-1	UL7	UL14	Fossum et al.	HSV-1	UL44	UL17	Fossum et al.	MCMV	M3	M119.3	Fossum et al.
HSV-1	UL7	UL15	Fossum et al.	HSV-1	UL44	UL21	Fossum et al.	MCMV	M7	M125	Fossum et al.
HSV-1	UL7	UL16	Fossum et al.	HSV-1	UL44	UL33	Fossum et al.	MCMV	M7	M163	Fossum et al.
HSV-1	UL7	UL33	Fossum et al.	HSV-1	UL44	UL40	Fossum et al.	MCMV	M7	M164	Fossum et al.
HSV-1	UL7	UL45	Fossum et al.	HSV-1	UL45	UL45	Fossum et al.	MCMV	M7	M169	Fossum et al.
HSV-1	UL7	UL53	Fossum et al.	HSV-1	UL45	UL53	Fossum et al.	MCMV	M8	M50	Fossum et al.
HSV-1	UL7	US2	Fossum et al.	HSV-1	UL46	UL33	Fossum et al.	MCMV	M10	M95	Fossum et al.
HSV-1	UL9	UL15	Fossum et al.	HSV-1	UL46	UL45	Fossum et al.	MCMV	M11	M26	Fossum et al.
HSV-1	UL9	UL33	Fossum et al.	HSV-1	UL46	UL53	Fossum et al.	MCMV	M11	M72	Fossum et al.
HSV-1	UL9	UL45	Fossum et al.	HSV-1	UL47	UL14	Fossum et al.	MCMV	M11	M87	Fossum et al.
HSV-1	UL10	UL49A	Fossum et al.	HSV-1	UL47	UL15	Fossum et al.	MCMV	M11	M93	Fossum et al.
HSV-1	UL10	US8A	Fossum et al.	HSV-1	UL47	UL17	Fossum et al.	MCMV	M11	M125	Fossum et al.
HSV-1	UL11	UL16	Fossum et al.	HSV-1	UL47	UL21	Fossum et al.	MCMV	M11	M126	Fossum et al.
HSV-1	UL14	UL14	Fossum et al.	HSV-1	UL47	UL40	Fossum et al.	MCMV	M11	M162	Fossum et al.
HSV-1	UL16	UL14	Fossum et al.	HSV-1	UL47	UL48	Fossum et al.	MCMV	M11	M168	Fossum et al.
HSV-1	UL16	UL21	Fossum et al.	HSV-1	UL47	UL49	Fossum et al.	MCMV	M14	M119.3	Fossum et al.
HSV-1	UL16	UL33	Fossum et al.	HSV-1	UL47	US1	Fossum et al.	MCMV	M17	M3	Fossum et al.
HSV-1	UL17	UL33	Fossum et al.	HSV-1	UL47	US11	Fossum et al.	MCMV	M17	M9	Fossum et al.
HSV-1	UL17	UL45	Fossum et al.	HSV-1	UL49	UL49	Fossum et al.	MCMV	M17	M10	Fossum et al.
HSV-1	UL21	UL20	Fossum et al.	HSV-1	UL49	US8	Fossum et al.	MCMV	M17	M13	Fossum et al.
HSV-1	UL23	UL40	Fossum et al.	HSV-1	UL49A	UL15	Fossum et al.	MCMV	M17	M14	Fossum et al.
HSV-1	UL23	UL45	Fossum et al.	HSV-1	UL49A	UL33	Fossum et al.	MCMV	M17	M48.2	Fossum et al.
HSV-1	UL26	UL45	Fossum et al.	HSV-1	UL49A	UL49A	Fossum et al.	MCMV	M17	M55	Fossum et al.
HSV-1	UL26	UL53	Fossum et al.	HSV-1	UL53	UL33	Fossum et al.	MCMV	M17	M72	Fossum et al.
HSV-1	UL28	UL33	Fossum et al.	HSV-1	UL53	UL53	Fossum et al.	MCMV	M17	M106	Fossum et al.
HSV-1	UL28	UL40	Fossum et al.	HSV-1	UL54	UL54	Fossum et al.	MCMV	M17	M107	Fossum et al.
HSV-1	UL28	UL45	Fossum et al.	HSV-1	UL55	UL33	Fossum et al.	MCMV	M17	M119.2	Fossum et al.
HSV-1	UL28	UL53	Fossum et al.	HSV-1	UL55	UL45	Fossum et al.	MCMV	M17	M119.3	Fossum et al.
HSV-1	UL28	US2	Fossum et al.	HSV-1	UL55	UL53	Fossum et al.	MCMV	M17	M126	Fossum et al.

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
HSV-1	UL30	UL14	Fossum et al.	HSV-1	UL56	UL21	Fossum et al.	MCMV	M17	M163	Fossum et al.
HSV-1	UL30	UL33	Fossum et al.	HSV-1	UL56	UL49	Fossum et al.	MCMV	M17	M164	Fossum et al.
HSV-1	UL30	UL40	Fossum et al.	HSV-1	UL56	US1	Fossum et al.	MCMV	M17	M169	Fossum et al.
HSV-1	UL30	UL45	Fossum et al.	HSV-1	UL56	US11	Fossum et al.	MCMV	M18	M119.2	Fossum et al.
HSV-1	UL30	UL53	Fossum et al.	HSV-1	US1	US7	Fossum et al.	MCMV	M18	M163	Fossum et al.
HSV-1	UL31	UL34	Fossum et al.	HSV-1	US2	UL33	Fossum et al.	MCMV	M20	M40	Fossum et al.
HSV-1	UL31	UL45	Fossum et al.	HSV-1	US2	UL45	Fossum et al.	MCMV	M20	M119.2	Fossum et al.
HSV-1	UL33	UL33	Fossum et al.	HSV-1	US2	UL53	Fossum et al.	MCMV	M22	M48.2	Fossum et al.
HSV-1	UL34	UL14	Fossum et al.	HSV-1	US2	US2	Fossum et al.	MCMV	M22	M72	Fossum et al.
HSV-1	UL36	UL48	Fossum et al.	HSV-1	US4	UL53	Fossum et al.	MCMV	M22	M87	Fossum et al.
HSV-1	UL38	UL14	Fossum et al.	HSV-1	US5	UL43	Fossum et al.	MCMV	M22	M88	Fossum et al.
HSV-1	UL38	UL18	Fossum et al.	HSV-1	US8	UL53	Fossum et al.	MCMV	M22	M93	Fossum et al.
HSV-1	UL38	UL33	Fossum et al.	HSV-1	US10	UL14	Fossum et al.	MCMV	M22	M107	Fossum et al.
HSV-1	UL38	UL45	Fossum et al.	HSV-1	US10	UL23	Fossum et al.	MCMV	M22	M125	Fossum et al.
HSV-1	UL38	UL48	Fossum et al.	HSV-1	US10	UL33	Fossum et al.	MCMV	M22	M126	Fossum et al.
HSV-1	UL40	UL14	Fossum et al.	HSV-1	US10	UL45	Fossum et al.	MCMV	M22	M128	Fossum et al.
HSV-1	UL40	UL15	Fossum et al.	HSV-1	US11	US11	Fossum et al.	MCMV	M22	M134	Fossum et al.
HSV-1	UL40	UL33	Fossum et al.	MCMV	M43	M103	Fossum et al.	MCMV	M22	M142	Fossum et al.
MCMV	M22	M144	Fossum et al.	MCMV	M43	M144	Fossum et al.	MCMV	M84	M33	Fossum et al.
MCMV	M22	M162	Fossum et al.	MCMV	M43	M168	Fossum et al.	MCMV	M85	M50	Fossum et al.
MCMV	M22	M168	Fossum et al.	MCMV	M45	M45	Fossum et al.	MCMV	M85	M77	Fossum et al.
MCMV	M22	M169	Fossum et al.	MCMV	M45	M48	Fossum et al.	MCMV	M85	M85	Fossum et al.
MCMV	M23.1	M29.1	Fossum et al.	MCMV	M45	M51	Fossum et al.	MCMV	M86	M48.1	Fossum et al.
MCMV	M23.1	M51	Fossum et al.	MCMV	M45	M72	Fossum et al.	MCMV	M87	M163	Fossum et al.
MCMV	M23.1	M87	Fossum et al.	MCMV	M45	M168	Fossum et al.	MCMV	M87	M164	Fossum et al.
MCMV	M23.1	M88	Fossum et al.	MCMV	M48	M10	Fossum et al.	MCMV	M89	M93	Fossum et al.
MCMV	M25.1a	M168	Fossum et al.	MCMV	M48	M94	Fossum et al.	MCMV	M90	M35	Fossum et al.
MCMV	M25.2	M69	Fossum et al.	MCMV	M48	M137	Fossum et al.	MCMV	M90	M51	Fossum et al.
MCMV	M26	M126	Fossum et al.	MCMV	M48.1	M51	Fossum et al.	MCMV	M90	M72	Fossum et al.
MCMV	M26	M168	Fossum et al.	MCMV	M48.2	M48.2	Fossum et al.	MCMV	M90	M88	Fossum et al.
MCMV	M28	M106	Fossum et al.	MCMV	M48.2	M51	Fossum et al.	MCMV	M90	M93	Fossum et al.
MCMV	M28	M119.3	Fossum et al.	MCMV	M48.2	M86	Fossum et al.	MCMV	M90	M103	Fossum et al.
MCMV	M28	M169	Fossum et al.	MCMV	M48.2	M87	Fossum et al.	MCMV	M90	M107	Fossum et al.
MCMV	M29.1	M29.1	Fossum et al.	MCMV	M48.2	M91	Fossum et al.	MCMV	M90	M119.2	Fossum et al.
MCMV	M29.1	M30	Fossum et al.	MCMV	M48.2	M119.2	Fossum et al.	MCMV	M90	M144	Fossum et al.
MCMV	M29.1	M51	Fossum et al.	MCMV	M50	M10	Fossum et al.	MCMV	M90	M168	Fossum et al.
MCMV	M29.1	M72	Fossum et al.	MCMV	M50	M14	Fossum et al.	MCMV	M90	M169	Fossum et al.
MCMV	M29.1	M97	Fossum et al.	MCMV	M50	M55	Fossum et al.	MCMV	M90	M29.1	Fossum et al.
MCMV	M29.1	M106	Fossum et al.	MCMV	M50	M119.2	Fossum et al.	MCMV	M93	M50	Fossum et al.
MCMV	M29.1	M107	Fossum et al.	MCMV	M50	M119.3	Fossum et al.	MCMV	M93	M51	Fossum et al.
MCMV	M29.1	M127	Fossum et al.	MCMV	M50	M126	Fossum et al.	MCMV	M93	M72	Fossum et al.
MCMV	M29.1	M144	Fossum et al.	MCMV	M50	M168	Fossum et al.	MCMV	M93	M77	Fossum et al.
MCMV	M29.1	M162	Fossum et al.	MCMV	M50	M169	Fossum et al.	MCMV	M93	M85	Fossum et al.
MCMV	M29.1	M169	Fossum et al.	MCMV	M51	M48	Fossum et al.	MCMV	M93	M87	Fossum et al.
MCMV	M30	M106	Fossum et al.	MCMV	M51	M51	Fossum et al.	MCMV	M93	M88	Fossum et al.
MCMV	M30	M144	Fossum et al.	MCMV	M51	M54	Fossum et al.	MCMV	M93	M93	Fossum et al.
MCMV	M32	M168	Fossum et al.	MCMV	M51	M97	Fossum et al.	MCMV	M93	M106	Fossum et al.

Appendix

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
MCMV	M35	M106	Fossum et al.	MCMV	M51	M126	Fossum et al.	MCMV	M93	M107	Fossum et al.
MCMV	M35	M144	Fossum et al.	MCMV	M52	M75	Fossum et al.	MCMV	M93	M25.1a	Fossum et al.
MCMV	M36	M25.1a	Fossum et al.	MCMV	M52	M97	Fossum et al.	MCMV	M94	M99	Fossum et al.
MCMV	M36	M126	Fossum et al.	MCMV	M53	M50	Fossum et al.	MCMV	M95	M51	Fossum et al.
MCMV	M36	M144	Fossum et al.	MCMV	M53	M51	Fossum et al.	MCMV	M97	M54	Fossum et al.
MCMV	M36	M163	Fossum et al.	MCMV	M53	M169	Fossum et al.	MCMV	M97	M95	Fossum et al.
MCMV	M36	M164	Fossum et al.	MCMV	M55	M125	Fossum et al.	MCMV	M98	M97	Fossum et al.
MCMV	M36	M168	Fossum et al.	MCMV	M55	M163	Fossum et al.	MCMV	M98	M106	Fossum et al.
MCMV	M36	M169	Fossum et al.	MCMV	M55	M164	Fossum et al.	MCMV	M98	M114	Fossum et al.
MCMV	M37	M10	Fossum et al.	MCMV	M56	M1	Fossum et al.	MCMV	M99	M10	Fossum et al.
MCMV	M37	M55	Fossum et al.	MCMV	M56	M14	Fossum et al.	MCMV	M99	M35	Fossum et al.
MCMV	M37	M119.2	Fossum et al.	MCMV	M69	M69	Fossum et al.	MCMV	M99	M72	Fossum et al.
MCMV	M37	M144	Fossum et al.	MCMV	M69	M72	Fossum et al.	MCMV	M99	M90	Fossum et al.
MCMV	M37	M163	Fossum et al.	MCMV	M69	M89	Fossum et al.	MCMV	M99	M106	Fossum et al.
MCMV	M37	M169	Fossum et al.	MCMV	M72	M51	Fossum et al.	MCMV	M99	M119.2	Fossum et al.
MCMV	M40	M106	Fossum et al.	MCMV	M73	M55	Fossum et al.	MCMV	M100	M26	Fossum et al.
MCMV	M40	M119.3	Fossum et al.	MCMV	M73	M119.3	Fossum et al.	MCMV	M100	M48.2	Fossum et al.
MCMV	M40	M146	Fossum et al.	MCMV	M73	M168	Fossum et al.	MCMV	M100	M51	Fossum et al.
MCMV	M40	M168	Fossum et al.	MCMV	M73	M169	Fossum et al.	MCMV	M100	M72	Fossum et al.
MCMV	M41	M9	Fossum et al.	MCMV	M77	M48	Fossum et al.	MCMV	M100	M73.5	Fossum et al.
MCMV	M41	M50	Fossum et al.	MCMV	M77	M51	Fossum et al.	MCMV	M100	M87	Fossum et al.
MCMV	M41	M72	Fossum et al.	MCMV	M77	M88	Fossum et al.	MCMV	M100	M168	Fossum et al.
MCMV	M41	M119.2	Fossum et al.	MCMV	M77	M103	Fossum et al.	MCMV	M103	M51	Fossum et al.
MCMV	M41	M147	Fossum et al.	MCMV	M77	M126	Fossum et al.	MCMV	M103	M53	Fossum et al.
MCMV	M43	M29.1	Fossum et al.	MCMV	M77	M168	Fossum et al.	MCMV	M103	M89	Fossum et al.
MCMV	M43	M87	Fossum et al.	MCMV	M80	M80	Fossum et al.	MCMV	M103	M97	Fossum et al.
MCMV	M43	M88	Fossum et al.	MCMV	M84	M23.1	Fossum et al.	MCMV	M108	M25.1a	Fossum et al.
MCMV	M43	M93	Fossum et al.	MCMV	M120	M144	Fossum et al.	MCMV	M108	M26	Fossum et al.
MCMV	M108	M48.2	Fossum et al.	MCMV	M120	M168	Fossum et al.	MCMV	M140	M93	Fossum et al.
MCMV	M108	M51	Fossum et al.	MCMV	M120	M169	Fossum et al.	MCMV	M140	M107	Fossum et al.
MCMV	M108	M72	Fossum et al.	MCMV	M124	M10	Fossum et al.	MCMV	M140	M124.1	Fossum et al.
MCMV	M108	M87	Fossum et al.	MCMV	M124	M106	Fossum et al.	MCMV	M140	M126	Fossum et al.
MCMV	M108	M88	Fossum et al.	MCMV	M124	M119.3	Fossum et al.	MCMV	M140	M128	Fossum et al.
MCMV	M108	M93	Fossum et al.	MCMV	M124	M169	Fossum et al.	MCMV	M140	M136	Fossum et al.
MCMV	M108	M103	Fossum et al.	MCMV	M124.1	M50	Fossum et al.	MCMV	M140	M141	Fossum et al.
MCMV	M117.1	M3	Fossum et al.	MCMV	M125	M72	Fossum et al.	MCMV	M140	M168	Fossum et al.
MCMV	M117.1	M9	Fossum et al.	MCMV	M125	M126	Fossum et al.	MCMV	M142	M126	Fossum et al.
MCMV	M117.1	M10	Fossum et al.	MCMV	M125	M163	Fossum et al.	MCMV	M142	M128	Fossum et al.
MCMV	M117.1	M13	Fossum et al.	MCMV	M125	M168	Fossum et al.	MCMV	M142	M143	Fossum et al.
MCMV	M117.1	M14	Fossum et al.	MCMV	M125	M169	Fossum et al.	MCMV	M142	M161	Fossum et al.
MCMV	M117.1	M55	Fossum et al.	MCMV	M126	M72	Fossum et al.	MCMV	M142	M168	Fossum et al.
MCMV	M117.1	M91	Fossum et al.	MCMV	M126	M87	Fossum et al.	MCMV	M144	M144	Fossum et al.
MCMV	M117.1	M106	Fossum et al.	MCMV	M126	M88	Fossum et al.	MCMV	M146	M48.2	Fossum et al.
MCMV	M117.1	M119.2	Fossum et al.	MCMV	M126	M126	Fossum et al.	MCMV	M147	M72	Fossum et al.
MCMV	M117.1	M119.3	Fossum et al.	MCMV	M126	M168	Fossum et al.	MCMV	M150	M40	Fossum et al.
MCMV	M117.1	M163	Fossum et al.	MCMV	M127	M78	Fossum et al.	MCMV	M150	M136	Fossum et al.
MCMV	M117.1	M164	Fossum et al.	MCMV	M127	M88	Fossum et al.	MCMV	M150	M169	Fossum et al.

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
MCMV	M117.1	M169	Fossum et al.	MCMV	M127	M93	Fossum et al.	MCMV	M151	M9	Fossum et al.
MCMV	M119.1	M26	Fossum et al.	MCMV	M127	M103	Fossum et al.	MCMV	M151	M10	Fossum et al.
MCMV	M119.1	M48.2	Fossum et al.	MCMV	M130	M71	Fossum et al.	MCMV	M151	M13	Fossum et al.
MCMV	M119.1	M51	Fossum et al.	MCMV	M131	M26	Fossum et al.	MCMV	M151	M14	Fossum et al.
MCMV	M119.1	M72	Fossum et al.	MCMV	M131	M126	Fossum et al.	MCMV	M151	M48.2	Fossum et al.
MCMV	M119.1	M87	Fossum et al.	MCMV	M131	M134	Fossum et al.	MCMV	M151	M51	Fossum et al.
MCMV	M119.1	M126	Fossum et al.	MCMV	M131	M48.2	Fossum et al.	MCMV	M151	M55	Fossum et al.
MCMV	M119.1	M152	Fossum et al.	MCMV	M131	M87	Fossum et al.	MCMV	M151	M72	Fossum et al.
MCMV	M119.1	M155	Fossum et al.	MCMV	M134	M9	Fossum et al.	MCMV	M151	M87	Fossum et al.
MCMV	M119.1	M162	Fossum et al.	MCMV	M134	M48.2	Fossum et al.	MCMV	M151	M119.2	Fossum et al.
MCMV	M119.1	M163	Fossum et al.	MCMV	M134	M126	Fossum et al.	MCMV	M151	M119.3	Fossum et al.
MCMV	M119.1	M168	Fossum et al.	MCMV	M134	M168	Fossum et al.	MCMV	M151	M162	Fossum et al.
MCMV	M119.2	M10	Fossum et al.	MCMV	M135	M119.3	Fossum et al.	MCMV	M154	M119.2	Fossum et al.
MCMV	M119.2	M12	Fossum et al.	MCMV	M138	M2	Fossum et al.	MCMV	M154	M169	Fossum et al.
MCMV	M119.2	M14	Fossum et al.	MCMV	M138	M3	Fossum et al.	MCMV	M155	M119.2	Fossum et al.
MCMV	M119.2	M55	Fossum et al.	MCMV	M138	M9	Fossum et al.	MCMV	M159	M3	Fossum et al.
MCMV	M119.2	M71	Fossum et al.	MCMV	M138	M10	Fossum et al.	MCMV	M159	M10	Fossum et al.
MCMV	M119.2	M72	Fossum et al.	MCMV	M138	M12	Fossum et al.	MCMV	M159	M13	Fossum et al.
MCMV	M119.2	M91	Fossum et al.	MCMV	M138	M13	Fossum et al.	MCMV	M159	M14	Fossum et al.
MCMV	M119.2	M106	Fossum et al.	MCMV	M138	M14	Fossum et al.	MCMV	M159	M55	Fossum et al.
MCMV	M119.2	M119.2	Fossum et al.	MCMV	M138	M20	Fossum et al.	MCMV	M159	M91	Fossum et al.
MCMV	M119.2	M125	Fossum et al.	MCMV	M138	M29.1	Fossum et al.	MCMV	M159	M106	Fossum et al.
MCMV	M119.2	M128	Fossum et al.	MCMV	M138	M30	Fossum et al.	MCMV	M159	M119.3	Fossum et al.
MCMV	M119.2	M136	Fossum et al.	MCMV	M138	M40	Fossum et al.	MCMV	M159	M163	Fossum et al.
MCMV	M119.2	M164	Fossum et al.	MCMV	M138	M48.2	Fossum et al.	MCMV	M159	M164	Fossum et al.
MCMV	M119.2	M169	Fossum et al.	MCMV	M138	M55	Fossum et al.	MCMV	M159	M169	Fossum et al.
MCMV	M119.3	M119.2	Fossum et al.	MCMV	M138	M72	Fossum et al.	MCMV	M161	M10	Fossum et al.
MCMV	M119.4	M29.1	Fossum et al.	MCMV	M138	M107	Fossum et al.	MCMV	M161	M13	Fossum et al.
MCMV	M119.4	M126	Fossum et al.	MCMV	M138	M119.2	Fossum et al.	MCMV	M161	M55	Fossum et al.
MCMV	M119.4	M152	Fossum et al.	MCMV	M138	M119.3	Fossum et al.	MCMV	M161	M106	Fossum et al.
MCMV	M119.4	M155	Fossum et al.	MCMV	M138	M119.5	Fossum et al.	MCMV	M161	M119.2	Fossum et al.
MCMV	M119.5	M87	Fossum et al.	MCMV	M138	M164	Fossum et al.	MCMV	M162	M26	Fossum et al.
MCMV	M119.5	M88	Fossum et al.	MCMV	M138	M169	Fossum et al.	MCMV	M162	M48.2	Fossum et al.
MCMV	M119.5	M162	Fossum et al.	MCMV	M139	M26	Fossum et al.	MCMV	M162	M72	Fossum et al.
MCMV	M119.5	M168	Fossum et al.	MCMV	M140	M50	Fossum et al.	MCMV	M162	M88	Fossum et al.
MCMV	M120	M125	Fossum et al.	MCMV	M140	M72	Fossum et al.	MCMV	M162	M125	Fossum et al.
MCMV	M120	M126	Fossum et al.	MCMV	M140	M88	Fossum et al.	MCMV	M162	M126	Fossum et al.
MCMV	M120	M127	Fossum et al.	EBV	BARF1	EBNA3A	Fossum et al.	MCMV	M162	M134	Fossum et al.
MCMV	M162	M162	Fossum et al.	EBV	BBLF1	BGLF2	Fossum et al.	EBV	BFLF1	BDRF1	Fossum et al.
MCMV	M162	M168	Fossum et al.	EBV	BBLF1	BTRF1	Fossum et al.	EBV	BFLF1	BFLF1	Fossum et al.
MCMV	M163	M119.2	Fossum et al.	EBV	BBLF2	BALF3	Fossum et al.	EBV	BFLF1	BFLF2	Fossum et al.
MCMV	M163	M119.3	Fossum et al.	EBV	BBLF2	BARF0	Fossum et al.	EBV	BFLF1	BGLF2	Fossum et al.
MCMV	M163	M163	Fossum et al.	EBV	BBLF2	BDRF1	Fossum et al.	EBV	BFLF1	BORF1	Fossum et al.
MCMV	M163	M169	Fossum et al.	EBV	BBLF2	BFLF2	Fossum et al.	EBV	BFLF2	BALF3	Fossum et al.
MCMV	M168	M119.2	Fossum et al.	EBV	BBLF2	BFRF4	Fossum et al.	EBV	BFLF2	BCRF1	Fossum et al.
MCMV	M170	M78	Fossum et al.	EBV	BBLF2	BORF1	Fossum et al.	EBV	BFLF2	BDRF1	Fossum et al.
EBV	A73	BALF3	Fossum et al.	EBV	BBLF3	EBNA3A	Fossum et al.	EBV	BFLF2	BFRF4	Fossum et al.

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
EBV	A73	BARF0	Fossum et al.	EBV	BBLF4	BALF3	Fossum et al.	EBV	BFLF2	BGLF2	Fossum et al.
EBV	A73	BBLF2	Fossum et al.	EBV	BBLF4	BFRF4	Fossum et al.	EBV	BFRF1	BDLF2	Fossum et al.
EBV	A73	BDRF1	Fossum et al.	EBV	BBRF1	BMLF1	Fossum et al.	EBV	BFRF1	BFLF2	Fossum et al.
EBV	A73	BFRF4	Fossum et al.	EBV	BBRF2	BTRF1	Fossum et al.	EBV	BFRF1	BFRF1	Fossum et al.
EBV	A73	BGLF2	Fossum et al.	EBV	BBRF3	BDLF2	Fossum et al.	EBV	BFRF1	BFRF4	Fossum et al.
EBV	A73	BGLF3	Fossum et al.	EBV	BBRF3	BFRF1	Fossum et al.	EBV	BFRF3	BFRF4	Fossum et al.
EBV	A73	BGLF5	Fossum et al.	EBV	BBRF3	BNLF2a	Fossum et al.	EBV	BFRF4	BALF3	Fossum et al.
EBV	A73	BLLF2	Fossum et al.	EBV	BcLF1	BNLF2a	Fossum et al.	EBV	BFRF4	BBRF3	Fossum et al.
EBV	A73	BTRF1	Fossum et al.	EBV	BCRF1	EBNA3A	Fossum et al.	EBV	BFRF4	BCRF1	Fossum et al.
EBV	BALF1	BALF3	Fossum et al.	EBV	BDLF2	BALF3	Fossum et al.	EBV	BFRF4	BDRF1	Fossum et al.
EBV	BALF1	BFRF4	Fossum et al.	EBV	BDLF2	BARF0	Fossum et al.	EBV	BFRF4	BFRF4	Fossum et al.
EBV	BALF1	BGLF5	Fossum et al.	EBV	BDLF2	BDLF2	Fossum et al.	EBV	BFRF4	BTRF1	Fossum et al.
EBV	BALF1	BSRF1	Fossum et al.	EBV	BDLF2	BDRF1	Fossum et al.	EBV	BGLF1	BARF0	Fossum et al.
EBV	BALF1	EBNA3A	Fossum et al.	EBV	BDLF2	BFRF4	Fossum et al.	EBV	BGLF1	BDRF1	Fossum et al.
EBV	BALF2	BALF3	Fossum et al.	EBV	BDLF2	BGLF3	Fossum et al.	EBV	BGLF1	BFRF1	Fossum et al.
EBV	BALF2	BFRF4	Fossum et al.	EBV	BDLF2	BSRF1	Fossum et al.	EBV	BGLF1	BFRF4	Fossum et al.
EBV	BALF3	BGLF5	Fossum et al.	EBV	BDLF3	EBNA3A	Fossum et al.	EBV	BGLF1	BLLF2	Fossum et al.
EBV	BALF3	BORF1	Fossum et al.	EBV	BdRF1	BALF3	Fossum et al.	EBV	BGLF1	BNLF2a	Fossum et al.
EBV	BALF3	BTRF1	Fossum et al.	EBV	BdRF1	BARF0	Fossum et al.	EBV	BGLF1	BTRF1	Fossum et al.
EBV	BALF4	BALF3	Fossum et al.	EBV	BdRF1	BBLF2	Fossum et al.	EBV	BGLF1	EBNA3A	Fossum et al.
EBV	BALF4	BARF0	Fossum et al.	EBV	BdRF1	BBRF3	Fossum et al.	EBV	BHRF1	BALF3	Fossum et al.
EBV	BALF4	BBLF2	Fossum et al.	EBV	BdRF1	BCRF1	Fossum et al.	EBV	BHRF1	BDLF2	Fossum et al.
EBV	BALF4	BBRF3	Fossum et al.	EBV	BdRF1	BDRF1	Fossum et al.	EBV	BHRF1	BFRF1	Fossum et al.
EBV	BALF4	BCRF1	Fossum et al.	EBV	BdRF1	BFRF4	Fossum et al.	EBV	BHRF1	BFRF3	Fossum et al.
EBV	BALF4	BDLF4	Fossum et al.	EBV	BdRF1	BGLF2	Fossum et al.	EBV	BHRF1	BFRF4	Fossum et al.
EBV	BALF4	BDRF1	Fossum et al.	EBV	BdRF1	BGLF3	Fossum et al.	EBV	BHRF1	BNLF2a	Fossum et al.
EBV	BALF4	BFLF2	Fossum et al.	EBV	BdRF1	BGLF5	Fossum et al.	EBV	BHRF1	EBNA3A	Fossum et al.
EBV	BALF4	BFRF4	Fossum et al.	EBV	BdRF1	BLLF2	Fossum et al.	EBV	BKRF2	BDLF2	Fossum et al.
EBV	BALF4	BGLF2	Fossum et al.	EBV	BdRF1	BMLF1	Fossum et al.	EBV	BKRF2	BFRF1	Fossum et al.
EBV	BALF4	BGLF3	Fossum et al.	EBV	BdRF1	BTRF1	Fossum et al.	EBV	BKRF2	BFRF4	Fossum et al.
EBV	BALF4	BGLF5	Fossum et al.	EBV	BDRF1	BALF3	Fossum et al.	EBV	BKRF2	BNLF2a	Fossum et al.
EBV	BALF4	BKRF3	Fossum et al.	EBV	BDRF1	BDRF1	Fossum et al.	EBV	BLLF1	BLLF1	Fossum et al.
EBV	BALF4	BLLF2	Fossum et al.	EBV	BDRF1	BGLF2	Fossum et al.	EBV	BLLF1	BLLF2	Fossum et al.
EBV	BALF4	BLRF2	Fossum et al.	EBV	BDRF1	BGLF5	Fossum et al.	EBV	BLLF1	EBNA3A	Fossum et al.
EBV	BALF4	BOLF1	Fossum et al.	EBV	BDRF1	BTRF1	Fossum et al.	EBV	BLRF1	BDLF2	Fossum et al.
EBV	BALF4	BORF1	Fossum et al.	EBV	BERF3	BALF3	Fossum et al.	EBV	BLRF1	BNLF2a	Fossum et al.
EBV	BALF4	BPLF1	Fossum et al.	EBV	BERF3	BARF0	Fossum et al.	EBV	BLRF1	EBNA3A	Fossum et al.
EBV	BALF4	BSRF1	Fossum et al.	EBV	BERF3	BBLF2	Fossum et al.	EBV	BLRF2	BLRF2	Fossum et al.
EBV	BALF4	BTRF1	Fossum et al.	EBV	BERF3	BDRF1	Fossum et al.	EBV	BLRF2	EBNA-LP	Fossum et al.
EBV	BALF4	EBNA3A	Fossum et al.	EBV	BERF3	BFLF2	Fossum et al.	EBV	BMRF2	BDLF2	Fossum et al.
EBV	BALF4	EBNA-LP	Fossum et al.	EBV	BERF3	BFRF4	Fossum et al.	EBV	BMRF2	BFRF1	Fossum et al.
EBV	BALF5	BALF3	Fossum et al.	EBV	BERF3	BGLF3	Fossum et al.	EBV	BMRF2	BNLF2a	Fossum et al.
EBV	BALF5	BARF0	Fossum et al.	EBV	BERF3	BGLF5	Fossum et al.	EBV	BNLF2a	BDLF2	Fossum et al.
EBV	BALF5	BFRF4	Fossum et al.	EBV	BERF3	BMLF1	Fossum et al.	EBV	BNLF2a	BFRF1	Fossum et al.
EBV	BALF5	BGLF2	Fossum et al.	EBV	BERF3	BORF1	Fossum et al.	EBV	BNLF2a	BNLF2a	Fossum et al.
EBV	BALF5	BNRF1	Fossum et al.	EBV	BERF3	BTRF1	Fossum et al.	EBV	BNLF2b	EBNA3A	Fossum et al.
EBV	BALF5	EBNA3A	Fossum et al.	EBV	BERF3	EBNA3A	Fossum et al.	EBV	BNRF1	BBLF1	Fossum et al.

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
EBV	BARF0	BFRF4	Fossum et al.	KSHV	23	28	Uetz et al.	EBV	BNRF1	BFRF4	Fossum et al.
EBV	BNRF1	BLRF2	Fossum et al.	KSHV	23	29b	Uetz et al.	KSHV	65	63	Uetz et al.
EBV	BNRF1	EBNA-LP	Fossum et al.	KSHV	23	30	Uetz et al.	KSHV	67.5	9	Uetz et al.
EBV	BOLF1	BPLF1	Fossum et al.	KSHV	23	45	Uetz et al.	KSHV	67.5	23	Uetz et al.
EBV	BPLF1	BALF3	Fossum et al.	KSHV	23	57	Uetz et al.	KSHV	67.5	28	Uetz et al.
EBV	BPLF1	BARF0	Fossum et al.	KSHV	23	60	Uetz et al.	KSHV	67.5	29b	Uetz et al.
EBV	BPLF1	BBRF3	Fossum et al.	KSHV	23	63	Uetz et al.	KSHV	67.5	34	Uetz et al.
EBV	BPLF1	BCRF1	Fossum et al.	KSHV	23	K09	Uetz et al.	KSHV	67.5	59	Uetz et al.
EBV	BPLF1	BDRF1	Fossum et al.	KSHV	25	65	Uetz et al.	KSHV	67.5	60	Uetz et al.
EBV	BPLF1	BFLF2	Fossum et al.	KSHV	27	58	Uetz et al.	KSHV	67.5	63	Uetz et al.
EBV	BPLF1	BFRF4	Fossum et al.	KSHV	28	28	Uetz et al.	KSHV	67.5	69	Uetz et al.
EBV	BPLF1	BHRF1	Fossum et al.	KSHV	28	29b	Uetz et al.	KSHV	67.5	75	Uetz et al.
EBV	BPLF1	BPLF1	Fossum et al.	KSHV	28	30	Uetz et al.	KSHV	68	9	Uetz et al.
EBV	BPLF1	BTRF1	Fossum et al.	KSHV	28	K11	Uetz et al.	KSHV	68	29b	Uetz et al.
EBV	BPLF1	EBNA3A	Fossum et al.	KSHV	29b	50	Uetz et al.	KSHV	68	57	Uetz et al.
EBV	BRLF1	BDLF2	Fossum et al.	KSHV	29b	54	Uetz et al.	KSHV	68	59	Uetz et al.
EBV	BSRF1	BARF0	Fossum et al.	KSHV	29b	72	Uetz et al.	KSHV	68	60	Uetz et al.
EBV	BSRF1	BBRF2	Fossum et al.	KSHV	29b	K08.1	Uetz et al.	KSHV	68	75	Uetz et al.
EBV	BSRF1	BDRF1	Fossum et al.	KSHV	29b	K10.5	Uetz et al.	KSHV	72	37	Uetz et al.
EBV	BSRF1	BFLF2	Fossum et al.	KSHV	30	29b	Uetz et al.	KSHV	74	27	Uetz et al.
EBV	BSRF1	BFRF4	Fossum et al.	KSHV	31	30	Uetz et al.	KSHV	74	29b	Uetz et al.
EBV	BSRF1	BGLF5	Fossum et al.	KSHV	31	31	Uetz et al.	KSHV	K03	53	Uetz et al.
EBV	BSRF1	BTRF1	Fossum et al.	KSHV	31	41	Uetz et al.	KSHV	K03	60	Uetz et al.
EBV	BSRF1	EBNA3A	Fossum et al.	KSHV	31	67.5	Uetz et al.	KSHV	K03	K03	Uetz et al.
EBV	BTRF1	BLRF2	Fossum et al.	KSHV	31	68	Uetz et al.	KSHV	K03	K07	Uetz et al.
EBV	BTRF1	EBNA-LP	Fossum et al.	KSHV	31	K11	Uetz et al.	KSHV	K05	6	Uetz et al.
EBV	BXRF1	BMLF1	Fossum et al.	KSHV	36	48	Uetz et al.	KSHV	K05	28	Uetz et al.
EBV	BZLF1	BGLF2	Fossum et al.	KSHV	36	56	Uetz et al.	KSHV	K05	34	Uetz et al.
EBV	BZLF1	BSRF1	Fossum et al.	KSHV	36	61	Uetz et al.	KSHV	K05	53	Uetz et al.
EBV	BZLF2	BDLF2	Fossum et al.	KSHV	37	K08	Uetz et al.	KSHV	K05	59	Uetz et al.
EBV	BZLF2	BLLF1	Fossum et al.	KSHV	39	9	Uetz et al.	KSHV	K05	60	Uetz et al.
EBV	BZLF2	BNLF2a	Fossum et al.	KSHV	41	9	Uetz et al.	KSHV	K07	74	Uetz et al.
EBV	EBNA3C	BBLF2	Fossum et al.	KSHV	41	28	Uetz et al.	KSHV	K07	K05	Uetz et al.
EBV	EBNA3C	BFRF4	Fossum et al.	KSHV	41	29b	Uetz et al.	KSHV	K08	57	Uetz et al.
EBV	EBNA3C	BLLF2	Fossum et al.	KSHV	41	63	Uetz et al.	KSHV	K08	60	Uetz et al.
EBV	EBNA-LP	BDLF2	Fossum et al.	KSHV	45	36	Uetz et al.	KSHV	K08.1	75	Uetz et al.
EBV	EBNA-LP	BFRF1	Fossum et al.	KSHV	45	50	Uetz et al.	KSHV	K09	63	Uetz et al.
EBV	EBNA-LP	BNLF2a	Fossum et al.	KSHV	45	72	Uetz et al.	KSHV	K09	69	Uetz et al.
EBV	LF2	LF2	Fossum et al.	KSHV	47	9	Uetz et al.	KSHV	K10	2	Uetz et al.
EBV	LMP1	BcLF1	Fossum et al.	KSHV	50	57	Uetz et al.	KSHV	K10	9	Uetz et al.
EBV	LMP1	BDLF2	Fossum et al.	KSHV	50	75	Uetz et al.	KSHV	K10	28	Uetz et al.
EBV	LMP1	BFRF1	Fossum et al.	KSHV	52	34	Uetz et al.	KSHV	K10	29b	Uetz et al.
EBV	LMP1	BNLF2a	Fossum et al.	KSHV	52	49	Uetz et al.	KSHV	K10	31	Uetz et al.
EBV	LMP2A	BALF3	Fossum et al.	KSHV	52	52	Uetz et al.	KSHV	K10	37	Uetz et al.
EBV	LMP2A	BARF0	Fossum et al.	KSHV	52	57	Uetz et al.	KSHV	K10	39	Uetz et al.
EBV	LMP2A	BDRF1	Fossum et al.	KSHV	52	59	Uetz et al.	KSHV	K10	41	Uetz et al.
EBV	LMP2A	BFLF2	Fossum et al.	KSHV	52	60	Uetz et al.	KSHV	K10	47	Uetz et al.

Y2H Interaction Set				Y2H Interaction Set				Y2H Interaction Set			
Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference	Species	Bait_ID	Prey_ID	Reference
EBV	LMP2A	BFRF4	Fossum et al.	KSHV	52	69	Uetz et al.	KSHV	K10	49	Uetz et al.
EBV	LMP2A	BLLF2	Fossum et al.	KSHV	54	36	Uetz et al.	KSHV	K10	59	Uetz et al.
EBV	LMP2A	BTRF1	Fossum et al.	KSHV	54	62	Uetz et al.	KSHV	K10	60	Uetz et al.
EBV	LMP2A	EBNA3A	Fossum et al.	KSHV	56	60	Uetz et al.	KSHV	K10	61	Uetz et al.
EBV	LMP2B	BDLF2	Fossum et al.	KSHV	57	57	Uetz et al.	KSHV	K10	67.5	Uetz et al.
EBV	RPMS1	BBLF2	Fossum et al.	KSHV	60	60	Uetz et al.	KSHV	K10	68	Uetz et al.
EBV	RPMS1	BGLF3	Fossum et al.	KSHV	60	K01	Uetz et al.	KSHV	K10	K12	Uetz et al.
EBV	RPMS1	BLLF2	Fossum et al.	KSHV	61	57	Uetz et al.	KSHV	K10.5	56	Uetz et al.
EBV	RPMS1	BMLF1	Fossum et al.	KSHV	61	60	Uetz et al.	KSHV	K10.5	75	Uetz et al.
KSHV	6	52	Uetz et al.	KSHV	61	61	Uetz et al.	KSHV	K11	34	Uetz et al.
KSHV	6	K15	Uetz et al.					KSHV	K11	29b	Uetz et al.
KSHV	K11	59	Uetz et al.								
KSHV	K11	60	Uetz et al.								
KSHV	K11	61	Uetz et al.								
KSHV	K11	69	Uetz et al.								
KSHV	K12	29b	Uetz et al.								
KSHV	K12	60	Uetz et al.								
KSHV	K12	K12	Uetz et al.								

Supplementary Table S2: List of all published PPIs
All PPIs: Interactions of this study combined with all previously published interactions among herpesviral proteins which were used as reference set.

VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID	PMID	Y2H_interolog	verification
ORF18C	ORF26	CC	1	3	ORF18-ORF26				KSHV_68-KSHV_60, redundant construct	Y2H red
ORF18C	ORF9a	CC	1	1	ORF18-ORF9a	EBV_BaRF1-EBV_BLRF1	17446270			LC
ORF28	ORF18C	CC	1	2	ORF18-ORF28				HSV1_UL30-HSV1_UL40, redundant construct	Y2H red
ORF28	ORF26	CC	1	1	ORF26-ORF28				KSHV_68-KSHV_9	Y2H
ORF42	ORF53	CC	1	1	ORF42-ORF53				HSV1_UL7-HSV1_UL15, MCMV_M103-MCMV_M89	Y2H
ORF42	ORF9a	CC	1	2	ORF42-ORF9a				HSV1_UL49A-HSV1_UL15, redundant construct	Y2H red
ORF50	ORF9a	CC	1	1	ORF50-ORF9a	EBV_BLRF1-EBV_BBRF3	11070013		HSV1_UL10-HSV1_UL49A	LC Y2H
ORF60C	ORF24N	CC	1	1	ORF24-ORF60				EBV_BKRF2-EBV_BFRF1	Y2H
ORF67C	ORF68C	CC	1	1	ORF67-ORF68	HSV1_US7-HSV1_US8, VZV_ORF67-VZV_ORF68	7995945	18945783		LC
ORF9a	ORF9a	CC	1	1	ORF9a-ORF9a				HSV1_UL49A-HSV1_UL49A	Y2H
ORF12	ORF38	CN	1	1	ORF12-ORF38	HSV1_UL46-HSV1_UL21	18602131			LC
ORF18C	ORF18C	CN	1	2	ORF18-ORF18	EBV_BaRF1-EBV_BaRF1	17446270		HSV1_UL40-HSV1_UL40, KSHV_60-KSHV_60, redundant construct	LC Y2H red
ORF24	ORF46	CN	1	1	ORF24-ORF46				HSV1_UL34-HSV1_UL14	Y2H
ORF24N	ORF66	CN	1	1	ORF24-ORF66	HSV1_UL34-HSV1_US3	10627546			LC
ORF25	ORF31N	CN	1	2	ORF25-ORF31				EBV_BALF4-EBV_BFRF4, redundant construct	Y2H red
ORF33	ORF33.5	CN	1	2	ORF33-ORF33.5	HSV1_UL26.5-HSV1_UL26	8661404		redundant direction	LC red
ORF39N	ORF38	CN	1	1	ORF38-ORF39				HSV1_UL21-HSV1_UL20	Y2H
ORF4	ORF4	CN	1	1	ORF4-ORF4	HSV1_UL54-HSV1_UL54, KSHV_57-KSHV_57	10329545	15269354	HSV1_UL54-HSV1_UL54, MCMV_M69-MCMV_M69, KSHV_57-KSHV_57	LC Y2H
ORF42	ORF42	CN	1	1	ORF42-ORF42	HSV1_UL15-HSV1_UL15	11086131			LC
ORF49	ORF38	CN	1	1	ORF38-ORF49				EBV_BBLF1-EBV_BTRF1	Y2H
ORF50C	ORF33.5	CN	1	1	ORF33.5-ORF50				EBV_BdRF1-EBV_BBRF3	Y2H
ORF24N	ORF24N	CN CC	2	1	ORF24-ORF24				EBV_BFRF1-EBV_BFRF1	Y2H
ORF28	ORF25	CN CC	2	1	ORF25-ORF28				HSV1_UL30-HSV1_UL33, MCMV_M51-MCMV_M54, EBV_BALF5-EBV_BFRF4, KSHV_67.5-KSHV_9	Y2H
ORF29	ORF25	CN CC	2	1	ORF25-ORF29				EBV_BALF2-EBV_BFRF4	Y2H
ORF43	ORF25	CN CC	2	1	ORF25-ORF43				HSV1_UL17-HSV1_UL33, MCMV_M93-MCMV_M51, EBV_BGLF1-EBV_BFRF4	Y2H
ORF12	ORF9	NC	1	1	ORF12-ORF9	HSV1_UL46-HSV1_UL49	18602131			LC

VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID	PMID	Y2H_interolog	verification
ORF23	ORF25	NC	1	1	ORF23-ORF25				MCMV_M48.2-MCMV_M51, EBV_BFRF3-EBV_BFRF4	Y2H
ORF44	ORF49	NC	1	2	ORF44-ORF49	HSV1_UL11-HSV1_UL16	16014918	18602131	HSV1_UL11-HSV1_UL16, MCMV_M94-MCMV_M99, EBV_BBLF1-EBV_BGLF2, redundant direction	LC Y2H red
ORF44	ORF53	NC	1	1	ORF44-ORF53				HSV1_UL7-HSV1_UL16	Y2H
ORF59	ORF25	NC	1	1	ORF25-ORF59				HSV1_UL2-HSV1_UL33	Y2H
ORF24N	ORF41	NC CN CC	3	1	ORF24-ORF41				MCMV_M85-MCMV_M50	Y2H
ORF27	ORF25	NC CN CC	3	1	ORF25-ORF27				MCMV_M53-MCMV_M51, EBV_BFLF2-EBV_BFRF4, KSHV_67.5-KSHV_69	Y2H
ORF60C	ORF25	NC CN CC	3	1	ORF25-ORF60				EBV_BKRF2-EBV_BFRF4	Y2H
ORF11	ORF38	NN	1	1	ORF11-ORF38				HSV1_UL47-HSV1_UL21	Y2H
ORF15F	ORF25	NN	1	2	ORF15-ORF25				HSV1_UL43-HSV1_UL33, redundant construct	Y2H red
ORF21	ORF22N	NN	1	1	ORF21-ORF22	HSV1_UL36-HSV1_UL37, KSHV_63-KSHV_64	16014918	18602131; 18321973	EBV_BOLF1-EBV_BPLF1	LC Y2H
ORF21	ORF23	NN	1	1	ORF21-ORF23	HSV1_UL35-HSV1_UL37	18602131		KSHV_65-KSHV_63	LC Y2H
ORF21	ORF60C	NN	1	1	ORF21-ORF60	KSHV_63-KSHV_47	18321973			LC
ORF25	ORF38	NN	1	1	ORF25-ORF38				EBV_BFRF4-EBV_BTRF1, KSHV_67.5-KSHV_23	Y2H
ORF28	ORF60C	NN	1	1	ORF28-ORF60				KSHV_47-KSHV_9	Y2H
ORF30	ORF25	NN	1	1	ORF25-ORF30	HSV1_UL28-HSV1_UL33	17035316		HSV1_UL28-HSV1_UL33, EBV_BFRF4-EBV_BALF3	LC Y2H
ORF33.5	ORF22N	NN	1	1	ORF22-ORF33.5				EBV_BPLF1-EBV_BDRF1	Y2H
ORF33.5	ORF27	NN	1	1	ORF27-ORF33.5				EBV_BFLF2-EBV_BDRF1	Y2H
ORF36	ORF36	NN	1	1	ORF36-ORF36	KSHV_21-KSHV_21	18321973			LC
ORF38	ORF22N	NN	1	1	ORF22-ORF38				EBV_BPLF1-EBV_BTRF1	Y2H
ORF4	ORF62	NN	1	1	ORF4-ORF62	HSV1_UL54-HSV1_RS1, VZV_ORF4-VZV_ORF62	8995681	10873781		LC
ORF41	ORF22N	NN	1	1	ORF22-ORF41	KSHV_64-KSHV_26	18321973			LC
ORF44	ORF25	NN	1	1	ORF25-ORF44				HSV1_UL16-HSV1_UL33	Y2H
ORF51	ORF25	NN	1	1	ORF25-ORF51				HSV1_UL9-HSV1_UL33	Y2H
ORF53	ORF7	NN	1	1	ORF53-ORF7				EBV_BSRF1-EBV_BBRF2	Y2H
ORF61	ORF61	NN	1	1	ORF61-ORF61	HSV1_RL2-HSV1_RL2	8151788	7966607		LC
ORF7	ORF23	NN	1	1	ORF23-ORF7	HSV1_UL35-HSV1_UL51	18602131			LC
ORF8	ORF19	NN	1	1	ORF19-ORF8				MCMV_M45-MCMV_M72	Y2H
ORF26	ORF26	NN CC	2	1	ORF26-ORF26				EBV_BFLF1-EBV_BFLF1	Y2H
ORF18C	ORF19	NN CN	2	2	ORF18-ORF19	HSV1_UL39-HSV1_UL40	1322407	3012359	KSHV_61-KSHV_60, redundant construct	LC Y2H red

VZV_bait	VZV_prey	permutation	perm_count	PPIs	VZV_pairs	LC_interologs	PMID	PMID	Y2H_interolog	verification
ORF19	ORF19	NN CN	2	1	ORF19-ORF19				MCMV_M45-MCMV_M45, KSHV_61-KSHV_61	Y2H
ORF4	ORF38	NN CN	2	1	ORF38-ORF4				KSHV_23-KSHV_57	Y2H
ORF12	ORF25	NN CN CC	3	2	ORF12-ORF25				HSV1_UL46-HSV1_UL33, redundant construct	Y2H red
ORF16	ORF16	NN CN CC	3	1	ORF16-ORF16	EBV_BMRF1-EBV_BMRF1, KSHV_59-KSHV_59	15286084	15075322		LC
ORF18	ORF25	NN CN CC	3	3	ORF18-ORF25				HSV1_UL40-HSV1_UL33, KSHV_67.5-KSHV_60, redundant construct	Y2H red
ORF33	ORF33	NN CN CC	3	1	ORF33-ORF33				MCMV_M80-MCMV_M80	Y2H
ORF9a	ORF25	NN CN CC	3	2	ORF25-ORF9a				HSV1_UL49A-HSV1_UL33, redundant construct	Y2H red
ORF33.5	ORF25	NN NC	2	1	ORF25-ORF33.5				EBV_BdRF1-EBV_BFRF4, EBV_BFRF4-EBV_BDRF1	Y2H
ORF38	ORF44	NN NC	2	2	ORF38-ORF44				HSV1_UL16-HSV1_UL21, redundant direction	Y2H red
ORF52	ORF25	NN NC	2	1	ORF25-ORF52				EBV_BBLF2-EBV_BFRF4	Y2H
ORF55	ORF25	NN NC	2	1	ORF25-ORF55				EBV_BBLF4-EBV_BFRF4	Y2H
ORF25	ORF25	NN NC CC	3	1	ORF25-ORF25				HSV1_UL33-HSV1_UL33, MCMV_M51-MCMV_M51, EBV_BFRF4-EBV_BFRF4	Y2H
ORF19	ORF25	NN NC CN CC	4	1	ORF19-ORF25				MCMV_M45-MCMV_M51	Y2H
ORF24	ORF27	NN NC CN CC	4	4	ORF24-ORF27	HSV1_UL34-HSV1_UL31	10627546	11507225; 15731273	HSV1_UL31-HSV1_UL34, redundant direction construct	LC Y2H red
ORF24N	ORF25	NN NC CN CC	4	1	ORF24-ORF25				EBV_BFRF1-EBV_BFRF4	Y2H
ORF3	ORF25	NN NC CN CC	4	1	ORF25-ORF3				HSV1_UL55-HSV1_UL33	Y2H
ORF50	ORF25	NN NC CN CC	4	2	ORF25-ORF50				MCMV_M100-MCMV_M51, EBV_BFRF4-EBV_BBRF3, redundant construct	Y2H red
ORF64	ORF25	NN NC CN CC	4	1	ORF25-ORF64				HSV1_US10-HSV1_UL33	Y2H
ORF8	ORF25	NN NC CN CC	4	1	ORF25-ORF8				MCMV_M72-MCMV_M51	Y2H

Supplementary Table S3: Verification of Y2H data.
List of verified interaction data extracted from Supplementary Table S1.

ORF_ID	ROW	COLUMN	Array ID	c_{min} [3-AT]	Array ID	c_{min} [3-AT]
ORF1	a	1	VZV_GBKg	0	VZV_GBKCg	0
ORF1N	a	2	VZV_GBKg	1	VZV_GBKCg	0
ORF2	a	3	VZV_GBKg	1	VZV_GBKCg	0
ORF3	a	4	VZV_GBKg	1	VZV_GBKCg	**40**
ORF4	a	5	VZV_GBKg	**50**	VZV_GBKCg	**3**
ORF5	a	6	VZV_GBKg	0	VZV_GBKCg	0
ORF5F	a	7	VZV_GBKg	0	VZV_GBKCg	0
ORF6	a	8	VZV_GBKg	0	VZV_GBKCg	0
ORF7	a	9	VZV_GBKg	0	VZV_GBKCg	0
ORF8	a	10	VZV_GBKg	1	VZV_GBKCg	**10**
ORF9	a	11	VZV_GBKg	**10**	VZV_GBKCg	0
ORF9a	a	12	VZV_GBKg	**10**	VZV_GBKCg	**3**
ORF9aN	b	1	VZV_GBKg	1	VZV_GBKCg	0
ORF10	b	2	VZV_GBKg	1	VZV_GBKCg	0
ORF11	b	3	VZV_GBKg	**50**	VZV_GBKCg	0
ORF12	b	4	VZV_GBKg	**3**	VZV_GBKCg	1
ORF12N	b	5	VZV_GBKg	0	VZV_GBKCg	0
ORF12C	b	6	VZV_GBKg	0	VZV_GBKCg	0
ORF13	b	7	VZV_GBKg	0	VZV_GBKCg	0
ORF14	b	8	VZV_GBKg	0	VZV_GBKCg	0
ORF14N	b	9	VZV_GBKg	0	VZV_GBKCg	0
ORF15	b	10	VZV_GBKg	1	VZV_GBKCg	0
ORF15N	b	11	VZV_GBKg	1	VZV_GBKCg	0
ORF15F	b	12	VZV_GBKg	0	VZV_GBKCg	1
ORF16	c	1	VZV_GBKg	**3**	VZV_GBKCg	**25**
ORF17	c	2	VZV_GBKg	0	VZV_GBKCg	0
ORF18	c	3	VZV_GBKg	0	VZV_GBKCg	0
ORF18N	c	4	VZV_GBKg	1	VZV_GBKCg	0
ORF18C	c	5	VZV_GBKg	0	VZV_GBKCg	0

ORF_ID	ROW	COLUMN	Array ID	c_{min} [3-AT]	Array ID	c_{min} [3-AT]
ORF19	c	6	VZV_GBKg	**10**	VZV_GBKCg	**3**
ORF20	c	7	VZV_GBKg	0	VZV_GBKCg	0
ORF21	c	8	VZV_GBKg	0	VZV_GBKCg	0
ORF22N	c	9	VZV_GBKg	**40**	VZV_GBKCg	0
ORF23	c	10	VZV_GBKg	0	VZV_GBKCg	0
ORF24	c	11	VZV_GBKg	0	VZV_GBKCg	0
ORF24N	c	12	VZV_GBKg	1	VZV_GBKCg	**3**
ORF25	d	1	VZV_GBKg	**10**	VZV_GBKCg	**3**
ORF26	d	2	VZV_GBKg	**3**	VZV_GBKCg	0
ORF27	d	3	VZV_GBKg	**3**	VZV_GBKCg	0
ORF28	d	4	VZV_GBKg	0	VZV_GBKCg	**10**
ORF29	d	5	VZV_GBKg	0	VZV_GBKCg	0
ORF30	d	6	VZV_GBKg	0	VZV_GBKCg	0
ORF31	d	7	VZV_GBKg	0	VZV_GBKCg	0
ORF31N	d	8	VZV_GBKg	0	VZV_GBKCg	0
ORF31C	d	9	VZV_GBKg	**25**	VZV_GBKCg	1
ORF32	d	10	VZV_GBKg	0	VZV_GBKCg	**40**
ORF33	d	11	VZV_GBKg	0	VZV_GBKCg	0
ORF33.5	d	12	VZV_GBKg	0	VZV_GBKCg	0
ORF34	e	1	VZV_GBKg	0	VZV_GBKCg	0
ORF35	e	2	VZV_GBKg	0	VZV_GBKCg	0
ORF36	e	3	VZV_GBKg	1	VZV_GBKCg	0
ORF37	e	4	VZV_GBKg	0	VZV_GBKCg	0
ORF37N	e	5	VZV_GBKg	0	VZV_GBKCg	0
ORF38	e	6	VZV_GBKg	1	VZV_GBKCg	0
ORF39	e	7	VZV_GBKg	0	VZV_GBKCg	0
ORF39N	e	8	VZV_GBKg	1	VZV_GBKCg	**3**
ORF40	e	9	VZV_GBKg	0	VZV_GBKCg	0
ORF41	e	10	VZV_GBKg	0	VZV_GBKCg	0
ORF42	e	11	VZV_GBKg	0	VZV_GBKCg	1
ORF43	e	12	VZV_GBKg	0	VZV_GBKCg	**3**

ORF_ID	ROW	COLUMN	Array ID	c_{min} [3-AT]	Array ID	c_{min} [3-AT]
ORF43C	f	1	VZV_GBKg	**3**	VZV_GBKCg	0
ORF44	f	2	VZV_GBKg	**10**	VZV_GBKCg	0
ORF45	f	3	VZV_GBKg	0	VZV_GBKCg	0
ORF46	f	4	VZV_GBKg	**40**	VZV_GBKCg	**40**
ORF47	f	5	VZV_GBKg	0	VZV_GBKCg	0
ORF48	f	6	VZV_GBKg	1	VZV_GBKCg	0
ORF49	f	7	VZV_GBKg	**3**	VZV_GBKCg	**3**
ORF50	f	8	VZV_GBKg	1	VZV_GBKCg	0
ORF50C	f	9	VZV_GBKg	0	VZV_GBKCg	0
ORF51	f	10	VZV_GBKg	0	VZV_GBKCg	0
ORF52	f	11	VZV_GBKg	1	VZV_GBKCg	0
ORF53	f	12	VZV_GBKg	**3**	VZV_GBKCg	0
ORF54	g	1	VZV_GBKg	0	VZV_GBKCg	0
ORF55	g	2	VZV_GBKg	0	VZV_GBKCg	0
ORF56	g	3	VZV_GBKg	0	VZV_GBKCg	0
ORF56C	g	4	VZV_GBKg	0	VZV_GBKCg	0
ORF57	g	5	VZV_GBKg	0	VZV_GBKCg	**3**
ORF58	g	6	VZV_GBKg	0	VZV_GBKCg	**3**
ORF59	g	7	VZV_GBKg	0	VZV_GBKCg	0
ORF60	g	8	VZV_GBKg	0	VZV_GBKCg	0
ORF60C	g	9	VZV_GBKg	**3**	VZV_GBKCg	0
ORF61	g	10	VZV_GBKg	1	VZV_GBKCg	1
ORF62/71	g	11	VZV_GBKg	**3**	VZV_GBKCg	0
ORF63/70	g	12	VZV_GBKg	0	VZV_GBKCg	0
ORF64/69	h	1	VZV_GBKg	1	VZV_GBKCg	**3**
ORF65	h	2	VZV_GBKg	0	VZV_GBKCg	0
ORF65N	h	3	VZV_GBKg	1	VZV_GBKCg	**40**
ORF66	h	4	VZV_GBKg	**50**	VZV_GBKCg	0
ORF67	h	5	VZV_GBKg	0	VZV_GBKCg	0
ORF67N	h	6	VZV_GBKg	1	VZV_GBKCg	0
ORF67C	h	7	VZV_GBKg	0	VZV_GBKCg	**10**

ORF_ID	ROW	COLUMN	Array ID	c_{min} [3-AT]	Array ID	c_{min} [3-AT]
ORF68	h	8	VZV_GBKg	1	VZV_GBKCg	0
ORF68F	h	9	VZV_GBKg	1	VZV_GBKCg	**3**
ORF68C	h	10	VZV_GBKg	1	VZV_GBKCg	**3**
ORF0	h	11	VZV_GBKg	0	VZV_GBKCg	0
ORF0C	h	12	VZV_GBKg	0	VZV_GBKCg	0

Supplementary Table S4: Autoactivation properties differ between vector systems.
In 20 cases the N-terminal bait fusion autoactivated while 21 of the C-terminal baits did so (**bold numbers**). In 7 cases both fusions were autoactivators at 3 mM 3AT or higher (ORF names highlighted **bold underlined**) concentrations but only ORF46 required more than 25 mM in both cases and was thus not interpretable.

>ORF1 / ORF1N
MSRVSEYGVPEGVRESDSDTDSVFMYQHTELMQNNASPLVVQTRPPAVLIPLVDVPRPRSRRKASAQLKMQMDRLCNVLGVVLQMATLALVTYIAFVVHTRATSCKRE
>ORF2
MHVISETLAYGHVPAFIMGSTLVRPSLNATAEENPASETRCLLRVLAGRTVDLPGGGTLHITCTKTYVIIGKYSKPGERLSLARLIGRAMTPGGARTFIILAMKEKRSTTLGYECGT GLHLLAPSMGTFLRTHGLSNRDLCLWRGNIYDMHMQRLMFWENIAQNTTETPCITSTLTCNLTEDSGEAALTTSDRPTLPTLTAQGRPTVSNIRGILKGSPRQQPVCHRVRFAE PTEGVLM
>ORF3
MDTTGASESSQPIRVNLKPDPLASFTQVIPPLALETTWTCPANSHAPTPSPLYGVKRLCALRATCGRADDLHAFLIGLGRRDKPSESPMYVDLQPFCSLLNSQRLLPEMANYNT LCDAPFSAATQQMMLESGQLGVHLAAIGYHCHCKSPFSAECWTGASEAYDHVVCGGKARAAVGGL
>ORF4
MASASIPTDPDVSTICEDFMNLLPDEPSDDFALEVTDWANDEAIGSTPGEDSTTSRTVYVERTADTAYNPRYSKRRHGRRESYHHNRPKTLVVVLPDSNHHGGRDVETGYARI ERGHRRSSRSYNTQSSRKHRDRSLSNRRRRPTTPPAMTTGERNDQTHDESYRLRFSKRDARRERIRKEYDIPVDRITGRAIEVVSTAGASVTIDSVRHLDETIEKLVVRYATIQE GDSWASGGCFPGIKQNTSWPELMLYGHELYRTFESYKMDSRIARALRERVIRGESLIEALESADELLTWIKMLAAKNLPIYTNNPIVATSKSLLENLKLKLGPFVRCLLLNRDNDL GSRTLPELLRQQRFSDITCITTYMFVMIARIANIVVRGSKFVEYDDISCNVQVLQEYTPGSCLAGVLEALITHQRECGRVECTLSTWAGHLSDARPYGKYFKCSTFNC
>ORF5 / ORF5F
MQALGIKTEHFIIMCLLSGHAVFTLWYTARVKFEHECVYATTVINGGPVVWGSYNNSLIYVTFVNHSTFLDGLSGYDYSCRENLLSGDTMVKTAISTPLHDKIRIVLGTRNCHAYF WCVQLKMIFFAWFVYGMYLQFRRIRRMFGPFRSSCELISPTSYSLNYVTRVISNILLGYPYTKLARLLCDVSMRRDGMSKVFNADPISFLYMHKGVTLLMLLEVIAHISSGCIVLLT LGVAYTPCALLYPTYIRILAWVVVCTLAIVELISYVRPKPTKDNHLNHINTGGIRGICTTCCATVMSGLAIKCFYIVIFAIAVVIFMHYEQRVQVSLFGESENSQKH
>ORF6
MDKSSKPTIRLLFATKGCAISHSLLLLTGQISTEPLYVVSYTWTPDLDDVFVKNGREEITQVIPTKRPREVTENDEENQIMHLFCSRDVNVIFYLIGGFSTGDVRSRVWPIFFCCFK TQTDFKALYKALWYGAPLNPHIISDTLCISETFDIHSEVIQTLMVTTHHLNRKGLSDNGLCITEATLCKLVKKSVGRQELTSLYAHYERQVLAAYRRLYWGYGCSPFWYIVRFGPS EKTLVLATRYYLLQTDTSYNTLETPLYDLQAIKDLFLTYQVPALPNCSGYNISDLLSFDKLSMFCCSSTYTRGLTAKNALSYILQRIHTDTTEIHAVSEYITNDRKGLKVPDREFVDYI YLAHFECFNRKQIADHLQAVTYSDFVNKPVLLKSSNLGKRATANFFNHVRSRLNMRDYIKKNVICDVTELGPEIGHKYTITKTYTLSLTYAAKPSKFIGVCDLATTLTRRVENIEKQ FSPYGWSSTIPSNPPGFDELSNFEDSGVSAEALRAANFANDTPNQSGRTGFDTSPGITKLLLFFSAATGIATHDVSILSYKTPLEALIGHSEVTGPMPVYRVALPHGAQAFAVIAN DTWSSITNRYTLPHEARLIAEDLKQINPCNFVAASLRDMQLTLLLSTSVKNVSKISSNIPKDQLYINRNELFNTNLIITNLILDVDF
HIRKPIPLGILHAGMRAFRHGILTAMQLLFPKAVVNPNKDPCYFYKTACPEPTVEVLDDDNLLDITSHSDIDFYIENGELYTCVEENYTEDVWFFDTQTTSEVHTHADVSNNENLH ETLPCNCKEKIGFRVCVPIPNPYALVGSSTLKGFAQILQQAVLLEREFVEYIGPYLRDFSFIDTGVYSHGHSLRLPFFSKVTTTGTAVGQLLPFYVVPEQCIDILAFVTSHRNPANF HFHSRPQSNVPVQFILHNLGGEYAEFFERKVARNKQIFSSPQISLTKALKERGVTCLDAFTLEAFVDSTILESIVEHIAVHFPGRDREYTLTSSKCIAIKRDWVLFQLICGTKGFTCL RYPHRGGRTAPRTFVSLRVDHHNRLCISLAQQCFATKCDSNRMHTIFTLEVPNYPNLTSS
>ORF7
MQTVCASLCGYARIPTEEPSYEEVRVNTHPQGAALLRLQEALTAVNGLLPAPLTLEDVVASADNTRRLVRAQALARTYAACSRNIECLKQHHFTEDNPGLNAVVRSHMENSKRL ADMCLAAITHLYLSVGAVDVTTDDIVDQTLRMTAESEVVMSDVVLLEKTLGVVAKPQASFDVSHNHELSIAKGENVGLKTSPIKSEATQLSEIKPPLIEVSDNNTSNLTKKTYPTET

LQPVLTPKQTQDVQRTTPAIKKSHVMLV
>ORF8
MNEAVIDPILETAVNTGDMFCSQTIPNRCLKDTILIEVQPECADTLQCVLDDKVSRHQPLLLRNHKKLELPSEKSVTRGGFYMQQLELLVKSAPPNEYALLLIQCKDTALADEDNF FVANGVIDAGYRGVISALLYYRPGVTVILPGHLTIYLFPVKLRQSRLLPKNVLKHLDPIFKSIQVQPLSNSPSNYEKPVIPEFADISTVQQGQPLHRDSAEYHIDVPLTYKHIINPKRQ EDAGYDICVPYNLYLKRNEFIKIVLPIIRDWDLQHPSINAYIFGRSSKSRSGIIVCPTAWPAGEHCKFYVYNLTGDDIRIKTGDRLAQVLLIDHNTQIHLKHNVLSNIAFPYAIRGKCGI PGVQWYFTKTLDLIATPSERGTRGFGSTDKETNDVDFLLKH
>ORF9a / ORF9aN
MGSITASFILITMQILFFCEDSSGEPNFAERNFWHASCSARGVYIDGSMITTLFFYASLLGVCVALISLAYHACFRLFTRSVLRSTW
>ORF9
MASSDGDRLCRSNAVRRKTTPSYSGQYRTARRSVVVGPPDDSDDSLGYITTVGADSPSPVYADLYFEHKNTTPRVHQPNDSSGSEDDFEDIDEVVAAFREARLRHELVEDAV YENPLSVEKPSRSFTKNAAVKPKLEDSPKRAPPGAGAIASGRPISFSTAPKTATSSWCGPTPSYNKRVFCEAVRRVAAMQAQKAAEAAWNSNPPRNNAELDRLLTGAVIRITVH EGLNLIQAANEADLGEGASVSKRGHNRKTGDLQGGMGNEPMYAQVRKPKSRTDTQTTGRITNRSRARSASRTDTRK
>ORF10
MECNLGTEHPSTDTWNRSKTEQAVVDAFDESLFGDVASDIGFETSLYSHAVKTAPSPPWVASPKILYQQLIRDLDFSEGPRLLSCLETWNEDLFSCFPINEDLYSDMMVLSPDP DDVISTVSTKDHVEMFNLTTRGSVRLPSPPKQPTGLPAYVQEVQDSFTVELRAREEAYTKLLVTYCKSIIRYLQGTAKRTTIGLNIQNPDQKAYTQLRQSILLRYYREVASLARLLY LHLYLTVTREFSWRLYASQSAHPDVFAALKFTWTERRQFTCAFHPVLCNHGIVLLEGKPLTASALREINYRRRELGLPLVRCGLVEENKSPLVQQPSFSVHLPRSVGFLTHHIKR KLDAYAVKHPQEPRHVRADHPYAKVVENRNYGSSIEAMILAPPSPSEILPGDPPRPPTCGFLTR
>ORF11
MQSGHYNRRQSRRQRISSNTTDSPRHTHGTRYRSTNWYTHPPQILSNSETLVAVQELLNSEMDQDSSSDASDDFPGYALHHSTYNGSEQNTSTSRHENRIFKLTEREANEEI NINTDAIDDEGEAEEGEAEEDAIDDEGEAEEGEAEEDAIDDEGEAEEGEAEEDAIDDEGEAEEGEAEEGEAEEGEAEEDAIDDEGEAEEDAAEEDAIDDEGEAEEDYFSVSQVC SRDADEVYFTLDPEISYSTDLRIAKVMEPAVSKELNVSKRCVEPVTLTGSMLAHNGFDESWFAMRECTRREYITVQGLYDPIHLRYQFDTSRMTPPQILRTIPALPNMTLGELLLI FPIEFMAQPISIERILVEDVFLDRRASSKTHKYGPRWNSVYALPYNAGKMYVQHIPGFYDVSLRAVGQGTAIWHHMILSTAACAISNRISHGDGLGFLLDAAIRISANCIFLGRNDN FGVGDPCWLEDHLAGLPREAVPDVLQVTQLVLPNRGPTVAIMRGFFGALAYWPELRIAISEPSTSLVRYATGHMELAEWFLFSRTHSLKPQFTPTEREMLASFFTLYVTLGGGM LNWICRATAMYLAAPYHSRSAYIAVCESLPYYYIPVNSDLLCDLEVLLLGEVDLPTVCESYATIAHELTGYEAVRTAATNFMIEFADCYKESETDLMVSAYLGAVLLLQRVLGHANL LLLLLSGAALYGGCSIYIPRGILDAYNTLMLAASPLYAHQTLTSFWKDRDDAMQTLGIRPTTDVLPKEQDRIVQASPIEMNFRFVGLETIYPREQPIPSVDLAENLMQYRNEILGLD WKSVAMHLLRKY
>ORF12 / ORF12N / *ORF12C*
MFSRFARSFSSDDRTRKSYDGSYQSFNAGERDLPTPTRDWCSISQRITSERVRDGCLIPTPGEALETAVKALSEKTDSLTSPVLQSTERHSVLLGLHHNNVPESLVVSCMSNDV HDGFMQRYMETIQRCLDDLKLSGDGLWWVYENTYWQYLKYTTGAEVPVTSEKVNKKSKSTVLLFSSVVANKPISRHPFKSKVINSDYRGICQELREALGAVQKYMYFMRPDD PTNPSPDTRIRVQEIAAYTATGYGWMLWFLDVVDARVCRHLKLQFRRIRGPRASVIPDDLLRRHLKTGPAVSAGTGVAFILAATTASALTALLRISVLWRKEEWRDGLNGTAAAI VAAVELITLLHHHFQYLIN*MMLIGYACWGDGGLNDPYILKALRAQGRFLYFAGQLVRTMSTHSWVVLETSTHMWFSRAVAQSILAHGGKPTKYYAQVLAASKRYTPLHLRRISEP SSVSDQPYIRFNRLGSPIGTGIGNLECVCLTGNYLSDDVNASSHVINTEAPLNSIAPDTNRQRTSRVLVRPDTGLDVTVRKNHCLDIGHTDGSPVDPTYPDHYTRIKAEYEGPVR DESNTMFDQRSDLRHIETQASLNDHVYENIPPKEVGFNSSSDLDVDSLNGYTSGDMHTDDDLSPDFIPNDVPVRCKTTVTFRKNTPKSHH*
>ORF13

MGDLSCWTKVPGFTLTGELQYLKQVDDILRYGVRKRDRTGIGTLSLFGMQARYNLRNEFPLLTTKRVFWRAVVEELLWFIRGSTDSKELAAKDIHIWDIYGSSKFLNRNGFHKR
HTGDLGPIYGFQWRHFGAEYKDCQSNYLQQGIDQLQTVIDTIKTNPESRRMIISSWNPKDIPLMVLPPCHTLCQFYVANGELSCQVYQRSGDMGLGVPFNIAGYALLTYIVAHVT
GLKTGDLIHTMGDAHIYLNHIDALKVQLARSPKPFPCLKIIRNVTDINDFKWDDFQLDGYNPHPPLKMEMAL

>ORF14 / ORF14N

MKRIQINLILTIACIQLSTESQPTPVSITELYTSAATRKPDPAVAPTSAASRKPDPAVAPTSAASRKPDPAVAPTSAASRKPDPAVAPTSAATRKPDPAVAPTSAASRKPDPAVAPT
SAATRKPDPAVAPTSAASRKPDPAANTQHSQPPFLYENIQCVHGGIQSIPYFHTFIMPCYMRLTTGQQAAFKQQQKTYEQYSLDPEGSNITRWKSLIRPDLHIEVWFTRHLIDPH
RQLGNALIRMPDLPVMLYSNSADLNLINNPEIFTHAKENYVIPDVKTTSDFSVTILSMDATTEGTYIWRVVNTKTKNVISEHSITVTTYYRPNITVVGDPVLTGQTYAAYCNVSKYYP
PHSVRVRWTSRFGNIGKNFITDAIQEYANGLFSYVSAVRIPQQKQMDYPPPAIQCNVLWIRDGVSNMKYSAVVTPDVYPFPNVSIGIIDGHIVCTAKCVPRGVVHFVWWVNDSPI
NHENSEITGVCDQNKRFVNMQSSCPTSELDGPITYSCHLDGYPKKFPPFSAVYTYDASTYATTFSVVAVIIGVISILGTLGLIAVIATLCIRCCS

>ORF15 / ORF15N / *ORF15F*

MAVNGERAVHDENLGVLDRELIRAQSIQGCVGNPQECNSCAITSASRLFLVGLQASVITSGLILQYHVCEAAVNATIMGLIVVSGLWPTSVKFLRTLAKLGRCLQTVVVLGFAVL
WAVGCPISRDLPFVELLGISISAITGTVAAVHIHYYNFVTTFNGPHIYFYVMMLGTGLGGLLTVILYMYVSKYEVLIGLCISIVTLVSI*VDAATDLQDTCIYRKNRHKQLNTYTDLGFAV*
*VYTQNDRGRVCDHRESSRTLKRVFKGIR*IMSVIPPVLYIVTPLMWAISHIIKLNHFIKLTQVTLAVSIGGHIIAFGLQGFAVLYQEKKNLWVIVLYTTTSVTGIAVTFAGISWGAIIILTST
VAAGLTCIQMMRLSVKPIDCFMASHITKVYHVCVYIIINLCYLCGTYVS

>ORF16

MDLRSRTDDALDMELHAGFDAPEIARAVLTEKTLTGLISSISPLVNRLRDSILIFSDEGLIIHCSLETEQLYIPIPANMFDQYNWTGPRMVVLAATEGRSSLIDAFRHTKDPSTPTRL
YFKFTGQPPERSIIQTMVWQRPGDCGPDDQVQCYKQVVKRELACYTMMFPNLTPDISICLKRDQFTRLQRLLKTFGFTTCFILTATDMYIQTAGGGFISFNVSLDINGSKPTPYN
LIRSITNSKRILNNVVYGSGSMREFGVLLETHSGFRSAVQNLKLTRDETCYINFYLALTNSPMVGLYIQRSAPVHSFFYATFLSPKDLKEKLTSMQLFANMESVKDEPPLKKRRNL
LTKRNEKNTGNKMGGKLPETTWQEGIGIREYCVAPPVDPAGTLDYSELSRESDVICTVK

>ORF17

MGLFGLTRFIHEHKLVKPSIISTPPGVLTPVAVDVWNVMYTLLERLYPVGKRENLHGPSVTIHCLGVLLRLLTQRSYYPIFVLERCTDGPLSRGAKAIMSRAMNHDERGTSDLTRV
LLSSNTSCSIKYNKTSETYDSVFRNSSTSCIPSEENKSQDMFLDGCPRQTDKTICLRDQNVCSLTSTMPSRGHPNHRLYHKLCASLIRWMGYAYVEAVDIEADEACANLFHTRT
VALVYTTDTDLLFMGCDILLDAIPMFAPVVRCRDLLQYLGITYPEFLVAFVRCQTDLHTSDNLKSVQQVIQDTGLKVPHQMDTSTRSPTYDSWRHGEVFKSLTVATSGKTENGVS
VSKYASNRSEVTVDASWALNLLPPSSSPLDNLERAFVEHIIAVVTPLTRGRLKLMKRVNIMQNTADPYMVINTLYHNLKGEKMARQYARIFKQFIPTPLPLNTVLTKYWN

>ORF18 / ORF18N / *ORF18C*

MDQKDCSHFFYRPECPDINNLRALSISNRWLESDFIIEDDYQYLDCLTEDELIFYRFIFTFLSAADDLVNVNLGSLTQLFSQKDIHHYYIEQECIEVVHARVYSQIQLMLFRGDESLR
VQYVNVTINNPSIQQKVQWLEEKVRDNPSVAEKYILMILIEGIFFVSSFAAIAY*LRNNGLFVVTCQFNDLISRDEAIHTSASCCIYNNYVPEKPAITRIHQLFSEAVEIECAFLKSHAPK*
TRLVNVDAITQYVKFSADRLLSAINVPKLFNTPPPDSDFPLAFMIADKNTNFFERHSTSYAGTVINDL

>ORF19

MEFKRIFNTVHDIINRLCQHGYKEYIIPPESTTPVELMEYISTIVSKLKAVTRQDERVYRCCGELIHCRINLRSVSMETWLTSPILCLTPRVRQAIEGRRDEIRRAILEPFLKDQYPAL
ATLGLQSALKYEDFYLTKLEEGKLESLCQFFLRLAATVTTEIVNLPKIATLIPGINDGYTWTDVCRVFFTALACQKIVPATPVMMFLGRETGATASCYLMDPESITVGRAVRAITGD
VGTVLQSRGGVGISLQSLNLIPTENQTKGLLAVLKLLDCMVMAINSDCERPTGVCVYIEPWHVDLQTVLATRGMLVRDEIFRCDNIFCCLWTPDLFFERYLSYLKGASNVQWTLF
DNRADILRTLHGEAFTSTYLRLEREGLGVSSVPIQDIAFTIIRSAAVTGSPFLMFKDACNRNYHMNTQGNAITGSNLCTEIVQKADAHQHGVCNLASINLTTCLSKGPVSFNLNDL
QLTARTTVIFLNGVLAAGNFPCKKSCKGVKNNRSLGIGIQGLHTTCLRLGFDLTSQPARRLNVQIAELMLYETMKTSMEMCKIGGLAPFKGFTESKYAKGWLHQDGFSTISYLDL
PWCTLRDDICAYGLYNSQFLALMPTVSSAQVTECSEGFSPIYNNMFSKVTTSGELLRPNLDLMDELRDMYSCEEKRLEVINILEKNQWSVIRSFGCLSNSHPLLKYKTAFEYEQE

DLVDMCAERAPFIDQSQSMTLFIEERPDGTIPASKIMNLLIRAYKAGLKTGMYYCKIRKATNSGLFAGGELTCTSCAL

>ORF20

MGSQPTNSHFTLNEQTLCGTNISLLGNNRFIQIGNGLHMTYAPGFFGNWSRDLTIGPRFGGLNKQPIHVPPKRTETASIQVTPRSIVINRMNNIQINPTSIGNPQVTIRLPLNNFKS
TTQLIQQVSLTDFFRPDIEHAGSIVLILRHPSDMIGEANTLTQAGRDPDVLLEGLRNLFNACTAPWTVGEGGGLRAYVTSLSFIAACRAEEYTDKQAADANRTAIVSAYGCSRMET
RLIRFSECLRAMVQCHVFPHRFISFFGSLLEYTIQDNLCNITAVAKGPQEAARTDKTSTRRVTANIPACVFWDVDKDLHLSADGLKHVFLVFVYTQRRQREGVRLHLALSQLNEQ
CFGRGIGFLLGRIRAENAAWGTEGVANTHQPYNTRALPLVQLSNDPTSPRCSIGEITGVNWNLARQRLYQWTGDFRGLPTQLSCMYAAYTLIGTIPSESVRYTRRMERFGGYN
VPTIWLEGVVWGGTNTWNECYY

>ORF21

MEEPICYDTQKLLDDLSNLKVQEADNERPWSPEKTEIARVKVVKFLRSTQKIPAKHFIQIWEPLHSNICFVYSNTFLAEAAFTAENLPGLLFWRLDLDWTIEEPGNSLKILTQLSSV
VQDSETLHRLSANKLRTSSKFGPVSIHFIITDWINMYEVALKDATTAIESPFTHARIGMLESAIAALTQHKFAIIYDMPFVQEGIRVLTQYAGWLLPFNVMWNQIQNSSLTPLTRALFI
ICMIDEYLTETPVHSISELFADTVNLIKDEAFVSIEEAVTNPRTVHESRISSALAYRDPYVFETSPGMLARRLRLDNGIWESNLLSLSTPGIHIEALLHLLNSDPEAETTSGSNVAEHT
RGIWEKVQASTSPSMLISTLAESGFTRFSCKLLRRFIAHHTLAGFIHGSVVADEHITDFQQTLGCLALVGGLAYQLVETYAPTTEYVLTYTRTVNETEKRYETLLPALGLPPGGLG
QIMRRCFAPRPLIESIQATRVILLNEISHAEARETTYFKQTHNQSSGALLPQAGQSAVREAVLTWFDLRMDSRWGITPPVDVGMTPPICVDPPATGLEAVMITEALKIAYPTEYNR
SSVFVEPSFVPYIIATSTLDALSATIALSFDTRGIQQALSILQWARDYGSGTVPNADGYRTKLSALITILEPFTRTHPPVLLPSHVSTIDSLICELHRTVGIAVDLLPQHVRPLVPDRPS
ITNSVFLATLYYDELYGRWTRLDKTSQALVENFTSNALVVSRYMLMLQKFFACRFYPTPDLQAVGICNPKVERDEQFGVWRLNDLADAVGHIVGTIQGIRTQMRVGISSLRTIMA
DASSALRECENLMTKTSTSAIGPLFSTMASRYARFTQDQMDILMRVDKLTTGENIPGLANVEIFLNRWERIATACRHATAVPSAESIATVCNELRRGLKNIQEDRVNAPTSYMSH
ARNLEDHKAAVSFVMDSRQQFIVDSGPQMGAVLTSQCNIGTWENVNATFLHDNVKITTTVRDVISEAPTLIIGQRWLRPDEILSNVDLRLGVPGNTSGSDP

>ORF22 / ORF22N

MDIIPPIAVTVAGVGSRNQFDGALGPASGLSCLRTSLSFLHMTYAHGINATLSSDMIDGCLQEGAAWTTDLSNMGRGVPDMCALVDLPNRISYIKLGDTTSTCCVLSRIYGDSHF
FTVPDEGFMCTQIPARAFFDDVWMGREESYTIITVDSTGMAIYRQGNISFIFDPHGHGTIGQAVVVRVNTTDVYSYIASEYTHRPDNVESQWAAALVFFVTANDGPVSEEALSSA
VTLIYGSCDTYFTDEQYCEKLVTAQHPLLLSPPNSTTIVLNKSSIVPLHQNVGESVSLEATLHSTLTNTVALDPRCSYSEVDPWHAVLETTSTGSGVLDCRRRRRPSWTPPSSEE
NLACIDDGLVNNTHSTDNLHKPAKKVLKFKPTVDVPDKTQVAHVLPRLREVANTPDVVLNVSNVDTPESSPTFSRNMNVGSSLKDRKPFLFEQSGDVNMVVEKLLQHGHEISN
GYVQNAVGTLDTVITGHTNVPIWVTRPLVMPDEKDPLELFINLTILRLTGFVVENGTRTHHGATSVVSDFIGPLGEILTGFPSAAELIRVTSLILTNMPGAEFGLVAMRVQDTTGAL
HAELDVLEADLGGSSPIDLYSRLSTGLISILNSPIISHPGLFAELIPTRTGSLSERIRLLCELVSARETRYMREHTALVSSVKALENALRSTRNKIDAIQIPEVPQEPPEETDIPPEELIR
RVYEIRSEVTMLLTSAVTEYFTRGVLYSTRALIAEQSPRRFRVATASTAPIQRLLDSLPEFDAKLTAIISSLSIHPPPETIQNLPVVSLLKELIKEGEDLNTDTALVSWLSVVGEAQTA
GYLSRREFDELSRTIKTINTRATQRASAEAELSCFNTLSAAVDQAVKDYETYNNGEVKYPEITRDDLLATIVRATDDLVRQIKILSDPMIQSGLQPSIKRRLETRLKEVQTYANEAR
TTQDTIKSRKQAAYNKLGGLLRPVTGFVGLRAAVDLLPELASELDVQGALVNLRTKVLEAPVEIRSQLTGDFWALFNQYRDILEHPGNARTSVLGGLGACFTAIIEIVPIPTEYRPS
LLAFFGDVADVLASDIATVSTNPESESAINAVVATLSKATLVSSTVPALSFVLSLYKKYQALQQEITNTHKLTELQKQLGDDFSTLAVSSGHLKFISSSNVDDYEINDAILSIQTNVHA
LMDTVKLVEVELQKLPPHCIAGTSTLSRVVKDLHKLVTMAHEKKEQAKVLITDCERAHKQQTTRVLYERWTRDIIACLEAMETRHIFNGTELARLRDMAAAGGFDIHAVYPQARQ
VVAACETTAVTALDTVFRHNPYTPENTNIPPPLALLRGLTWFDDFSITAPVFTVMFPGVSIEGLLLLMRIRAVVLLSADTSINGIPNYRDMILRTSGDLLQIPALAGYVDFYTRSYDQ
FITESVTLSELRADIRQAAGAKLTEANKALEEVTHVRAHETAKLALKEGVFITLPSEGLLIRAIEYFTTFDHKRFIGTAYERVLQTMVDRDLKEANAELAQFRMVCQATKNRAIQILQ
NIVDTANATEQQEDVDFTNLKTLLKLTPPPKTIALAIDRSTSVQDIVTQFALLLGRLEEETGTLDIQAVDWMYQARNIIDSHPLSVRIDGTGPLHTYKDRVDKLYALRTKLDLLRRRI
ETGEVTWDDAWTTFKRETGDMLASGDTYATSVDSIKALQASASVVDMLCSEPEFFLLPVETKNRLQKKQQERKTALDVVLQKQRQFEETASRLRALIERIPTESDHDVLRMLLR
DFDQFTHLPIWIKTQYMTFRNLLMVRLGLYASYAEIFPPASPNGVFAPIPAMSGVCLEDQSRCIRARVAAFMGEASVVQTFREARSSIDALFGKNLTFYLDTDGVPLRYRVCYKS
VGVKLGTMLCSQGGLSLRPALPDEGIVEETTLSALRVANEVNELRIEYESAIKSGFSAFSTFVRHRHAEWGKTNARRAIAEIYAGLITTTLTRQYGVHWDKLIYSFEKHHLTSVMG
NGLTKPIQRRGDVRVLELTLSDIVTILVATTPVHLLNFARLDLIKQHEYMARTLRPVIEAAFRGRLLVRSLDGDPKGNARAFFNAAPSKHKLPLALGSNQDPTGGRIFAFRMADWK
LVKMPQKITDPFAPWQLSPPPGVKANVDAVTRIMATDRLATITVLGRMCLPPISLVSMWNTLQPEEFAYRTQDDVDIIVDARLDLSSTLNARFDTAPSNTTLEWNTDRKVITDAYI
QTGATTVFTVTGAAPTHVSNVTAFDIATTAILFGAPLVIAMELTSVFSQNSGLTLGLKLFDSRHMATDSGISSAVSPDIVSWGLRLLHMDPHPIENACLIVQLEKLSALIANKPLTNN

PPCLLLLDEHMNPSYVLWERKDSIPAPDYVVFWGPESLIDLPYIDSDEDSFPSCPDDPFYSQIIAGYAPQGPPNLDTTDFYPTEPLFKSPVQVVRSSKCKKMPVRPAQPAQPAQ
PAQPAQTVQPAQPIEPGTQIVVQNFKKPQSVKTTLSQKDIPLYVETESETAVLIPKQLTTSIKTTVCKSITPPNNQLSDWKNNPQQNQTLNQAFSKPILEITSIPTDDSISYRTWIEK
SNQTQKRHQNDPRMYNSKTVFHPVNNQLPSWVDTAADAPQTDLLTNYKTRQPSPNFPRDVHTWGVSSNPFNSPNRDLYQSDFSEPSDGYSSESENSIVLSLDEHRSCRVPR
HVRVVNADVVTGRRYVRGTALGALALLSQACRRMIDNVRYTRKLLMDHTEDIFQGLGYVKLLLDGTYI

>ORF23

MTQPASSRVVFDPSNPTTFSVEAIAAYTPVALIRLLNASGPLQPGHRVDIADARSIYTVGAAASAARARANHNANTIRRTAMFAETDPMTWLRPTVGLKRTFNPRIIRPQPPNPS
MSLGISGPTILPQKTQSADQSALQQPAALAFSGSSPQHPPPQTTSASVGQQQHVVSGSSGQQPQQGAQSSTVQPTTGSPPAAQGVPQSTPPPTQNTPQGGKGQTLSHTGQ
SGNASRSRRV

>ORF24 / ORF24N

MSRRTYVRSERRRGCGDNLLQRIRLVVPSALQCCDGDLPIFDPQRPPARCVFQFNGEDNVSEAFPVEYIMRLMANWAQVDCDPYIKIQNTGVSVLFQGFFFRPTNAPVAEVSI
DSNNVILSSTLSTGINLSALESIKRGGGIDRRPLQALMWVNCFVRMPYVQLSFRFMGPEDPSRTIKLMARATDAYMYKETGNNLDEYIRWRPSFRSPPENGSPNTSVQMQSDIK
PALPDTQTTRVWKLALPVANVTYALFIVIVLVVVLGAVLFWK

>ORF25

MYESENASEHHPELEDVFSENTGDSNPSMGSSDSTRSISGMRARDLITDTDVNLLNIDALESKYFPADSTFTLSVWFENLIPPEIEAILPTTDAQLNYISFTSRLASVLKHKESNDS
EKSAYVVPCEHSASVTRRRERFAGVMAKFLDLHEILKDA

>ORF26

MDRVESEEPMDGFESPVFSENTSSNSGWCSDAFSDSYIAYNPALLLKNDLLFSELLFASHLINVPRAIENNVTYEASSAVGVDNEMTSSTTEFIEEIGDVLALDRACLVCRTLDLY
KRKFGLTPEWVADYAMLCMKSLASPPCAVVTFSAAFEFVYLMDRYYLCRYNVTLVGSFARRTLSLLDIQRHFFLHVCFRTDGGLPGIRPPPGKEMANKVRYSNYSFFVQAVVR
AALLSISTSRLDETETRKSFYFNQDGLTGGPQPLAAALANWKDCARMVDCSSSEHRTSGMITCAERALKEDIEFEDILIDKLKKSSYVEAAWGYADLALLLLSGVATWNVDERTN
CAIETRVGCVKSYWQANRIENSRDVPKQFSKFTSEDACPEVAFGPILLTTLKNAKCRGRTNTECMLCCLLTIGHYWIALRQFKRDILAYSANNTSLFDCIEPVINAWSLDNPIKLKF
PFNDEGRFITIVKAAGSEAVYKHLFCDLLCALSELQTNPKILFAHPTTADKEVLELYKAQLAAQNRFEGRVCAGLWTLAYAFKAYQIFPRKPTANAAFIRDGGLMLRRHAISLVSLE
HTLSKYV

>ORF27

MHLKPTRFFHANQPPMPHSYEMEDLCFDDMQYRWSPSNTPYRSMSRRYKSVSRSGPSMRVRSRTPCRRQTIRGKLMSKERSVYRHYFNYIARSPPEELATVRGLIVPIIKTTP
VTLPFNLGQTVADNCLSLSGMGYHLGLGGYCPTCTASGEPRLCRTDRAALILAYVQQLNNIYEYRVFLASILALSDRANMQAASAEPLLSSVLAQPELFFMYHIMREGGMRDIRV
LFYRDGDAGGFMMYVIFPGKSVHLHYRLIDHIQAACRGYKIVAHVWQTTFLLSVCRNPEQQTETVVPSIGTSDVYCKMCDLNFDGELLLEYKRLYALFDDFVPPR

>ORF28

MAIRTGFCNPFLTQASGIKYNPRTGRGSNREFLHSYKTTMSSFQFLAPKCLDEDVPMEERKGVHVGTLSRPPKVYCNGKEVPILDFRCSSPWPRRVNIWGEIDFRGDKFDPRF
NTFHVYDIVETTEAASNGDVSRFATATRPLGTVITLLGMSRCGKRVAVHVYGICQYFYINKAEVDTACGIRSGSELSVLLAECLRSSMITQNDATLNGDKNAFHGTSFKSASPESF
RVEVIERTDVYYYDTQPCAFYRVYSPSSKFTNYLCDNFHPELKKYEGRVDATTRFLMDNPGFVSFGWYQLKPGVDGERVRVRPASRQLTLSDVEIDCMSDNLQAIPNDDSWP
DYKLLCFDIECKSGGSNELAFPDATHLEDLVIQISCLLYSIPRQSLEHILLFSLGSCDLPQRYVQEMKDAGLPEPTVLEFDSEFELLIAFMTLVKQYAPEFATGYNIVNFDWAFIMEK
LNSIYSLKLDGYGSINRGGLFKIWDVGKSGFQRRSKVKINGLISLDMYAIATEKLKLSSYKLDSVAREALNESKRDLPYKDIPGYYASGPNTRGIIGEYCIQDSALVGKLFFKYLPHL
ELSAVARLARITLTKAIYDGQQVRIYTCLLGLASSRGFILPDGGYPATFEYKDVIPDVGDVEEEMDEDESVSPTGTSSGRNVGYKGARVFDPDTGFYIDPVVVLDFASLYPSIIQAH
NLCFTTLTLNFETVKRLNPSDYATFTVGGKRLFFVRSNVRESLLGVLLKDWLAMRKAIRARIPGSSSDEAVLLDKQQAAIKVVCNSVYGFTGVAQGFLPCLYVAATVTTIGRQML
LSTRDYIHNNWAAFERFITAFPDIESSVLSQKAYEVKVIYGDTDSVFIRFKGVSVEGIAKIGEKMAHIISTALFCPPIKLECEKTFIKLLLITKKKYIGVIYGGKVLMKGVDLVRKNNCQ

FINDYARKLVELLLYDDTVSRAAAEASCVSIAEWNRRAMPSGMAGFGRIIADAHRQITSPKLDINKFVMTAELSRPPSAYINRRLAHLTVYYKLVMRQGQIPNVRERIPYVIVAPTD
EVEADAKSVALLRGDPLQNTAGKRCGEAKRKLIISDLAEDPIHVTSHGLSLNIDYYFSHLIGTASVTFKALFGNDTKLTERLLKRFIPETRVVNVKMLNRLQAAGFVCIHAPCWDNK
MNTEAEITEEEQSHQIMRRVFCIPKAILHQS

>ORF29

MENTQKTVTVPTGPLGYVYACRVEDLDLEEISFLAARSTDSDLALLPLMRNLTVEKTFTSSLAVVSGARTTGLAGAGITLKLTTSHFYPSVFVFHGGKHVLPSSAAPNLTRACNA
ARERFGFSRCQGPPVDGAVETTGAEICTRLGLEPENTILYLVVTALFKEAVFMCNVFLHYGGLDIVHINHGDVIRIPLFPVQLFMPDVNRLVPDPFNTHHRSIGEGFVYPTPFYNT
GLCHLIHDCVIAPMAVALRVRNVTAVARGAAHLAFDENHEGAVLPPDITYTYFQSSSSGTTTARGARRNDVNSTSKPSPSGGFERRLASIMAADTALHAEVIFNTGIYEETPTDIK
EWPMFIGMEGTLPRLNALGSYTARVAGVIGAMVFSPNSALYLTEVEDSGMTEAKDGGPGPSFNRFYQFAGPHLAANPQTDRDGHVLSSQSTGSSNTEFSVDYLALICGFGAPL
LARLLFYLERCDAGAFTGGHGDALKYVTGTFDSEIPCSLCEKHTRPVCAHTTVHRLRQRMPRFGQATRQPIGVFGTMNSQYSDCDPLGNYAPYLILRKPGDQTEAAKATMQDT
YRATLERLFIDLEQERLLDRGAPCSSEGLSSVIVDHPTFRRILDTLRARIEQTTTQFMKVLVETRDYKIREGLSEATHSMALTFDPYSGAFCPITNFLVKRTHLAVVQDLALSQCHC
VFYGQQVEGRNFRNQFQPVLRRRFVDLFNGGFISTRSITVTLSEGPVSAPNPTLGQDAPAGRTFDGDLARVSVEVIRDIRVKNRVVFSGNCTNLSEAARARLVGLASAYQRQE
KRVDMLHGALGFLLKQFHGLLFPRGMPPNSKSPNPQWFWTLLQRNQMPADKLTHEEITTIAAVKRFTEEYAAINFINLPPTCIGELAQFYMANLILKYCDHSQYLINTLTSIITGAR
RPRDPSSVLHWIRKDVTSAADIETQAKALLEKTENLPELWTTAFTSTHLVRAAMNQRPMVVLGISISKYHGAAGNNRVFQAGNWSGLNGGKNVCPLFTFDRTRRFIIACPRGGFI
CPVTGPSSGNRETTLSDQVRGIIVSGGAMVQLAIYATVVRAVGARAQHMAFDDWLSLTDDEFLARDLEELHDQIIQTLETPWTVEGALEAVKILDEKTTAGDGETPTNLAFNFDS
CEPSHDTTSNVLNISGSNISGSTVPGLKRPPEDDELFDLSGIPIKHGNITMEMI

>ORF30

MELDINRTLLVLLGQVYTYIFQVELLRRCDPRVACRFLYRLAANCLTVRYLLKLFLRGFNTQLKFGNTPTVCALHWALCYVKGEGERLFELLQHFKTRFVYGETKDSNCIKDYFVS
AFNLKTCQYHHELSLTTYGGYVSSEIQFLHDIENFLKQLNYCYIITSSREALNTLETVTRFMTDTIGSGLIPPVELFDPAHPCAICFEELCITANQGETLHRRLLGCICDHVTKQVRV
NVDVDDIIRCLPYIPDVPDIKRQSAVEALRTLQTKTVVNPMGAKNDTFDQTYEIASTMLDSYNVFKPAPRCMYAISELKFWLTSNSTEGPQRTLDVFVDNLDVLNEHEKHAELTAV
TVELALFGKTPIHFDRAFSEELGSLDAIDSILVGNRSSSPDSQIEALIKACYAHHLSSPLMRHISNPSHDNEAALRQLLERVGCEDDLTKEASDSATASECDLNDDSSITFAVHGW
ENLLSKAKIDAAERKRVYLEHLSKRSLTSLGRCIREQRQELEKTLRVNVYGEALLQTFVSMQNGFGARNVFLAKVSQAGCIIDNRIQEAAFDAHRFIRNTLVRHTVDAAMLPALTH
KFFELVNGPLFNHDEHRFAQPPNTALFFTVENVGLFPHLKEELAKFMGGVVGSNWLLSPFRGFYCFSGVEGVTFAQRLAWKYIRELVFATTLFTSVFHCGEVRLCRVDRLGKD
PRGCTSQPKGIGSSHGPLDGIYLTYEETCPLVAIIQSGETGIDQNTVVIYDSDVFSLLYTLMQRLAPDSTDPAFS

>ORF31 / ORF31N / *ORF31C*

MFVTAVVSVSPSSFYESLQVEPTQSEDITRSAHLGDGDEIREAIHKSQDAETKPTFYVCPPPTGSTIVRLEPTRTCPDYHLGKNFTEGIAVVYKENIAAYKFKATVYYKDVIVSTA
WAGSSYTQITNRYADRVPIPVSEITDTIDKFGKCSSKATYVRNNHKVEAFNEDKNPQDMPLIASKYNSVGSKAWHTTNDTYMVAGTPGTYRTGTSVNCIIEEVEARSIFPYDSFG
LSTGDIIYMSPFFGLRDGAYREHSNYAMDRFHQFEGYRQRDLDTRALLEPAARNFLVTPHLTVGWNWKPKRTEVCSLVKWREVEDVVRDEYAHNFRFTMKTLSTTFISETNEF
NLNQIHLSQCVKEEARAIINRIYTTRYNSSHVRTGDIQTYLARGGFVVVFQPLLSNSLARLYLQELVRENTNHSPQKHPTRNTRSRRSVPVELRANRTITTTSSVEFAMLQFTYDH
IQEHVNEMLARISSSWCQLQNRERALWSGLFPINPSALASTILDQRVKARILGDVISVSNCPELGSDTRIILQNSMRVSGSTTRCYSRPLISIVSLNGSGTVEGQLGTDNELIMSRD
LLEPCVANHKRYFLFGHHYVYYEDYRYVREIAVHDVGMISTYVDLNLTLLKDREFMPLQVYTRDELRDTGLLDYSEIQRRNQMHSLRFYDIDKVVQYDSGTAIMQGMAQFFQGL
GTAGQAVGHVVLGATGALLSTVHGFTTFLSNPFGALAVGLLVLAGLVAAFFAYRYVLKL*KTSPMKALYPLTTKGLKQLPEGMDPFAEKPNATDTPIEEIGDSQNTEPSVNSGFDP
DKFREAQEMIKYMTLVSAAERQESKARKKNKTSALLTSRLTGLALRNRRGYSRVRTENVTGV*

>ORF32

MESSNINALQQPSSIAHHPSKQCASSLNETVKDSPPAIYEDRLEHTPVQLPRDGTPRDVCSVGQLTCRACATKPFRLNRDSQYDYLNTCPGGRHISLALEIITGRWVCIPRVFPD
TPEEKWMAPYIIPDREQPSSGDEDSDTD

>ORF33 / ORF33,5

MAAEADEENCEALYVAGYLALYSKDEGELNITPEIVRSALPPTSKIPINIDHRKDCVVGEVIAIIEDIRGPFFLGIVRCPQLHAVLFEAAHSNFFGNRDSVLSPLERALYLVTNYLPSV
SLSSKRLSPNEIPDGNFFTHVALCVVGRRVGTVVNYDCTPESSIEPFRVLSMESKARLLSLVKDYAGLNKVWKVSEDKLAKVLLSTAVNNMLLRDRWDVVAKRRREAGIMGHV
YLQASTGYGLARITNVNGVESKLPNAGVINATFHPGGPIYDLALGVGESNEDCEKTVPHLKVTQLCRNDSDMASVAGNASNISPQPPSGVPTGGEFVLIPTAYYSQLLTGQTKN
PQVSIGAPNNGQYIVGPYGSPHPPAFPPNTGGYGCPPGHFGGPYGFPGYPPPNRLEMQMSAFMNALAAERGIDLQTPCVNFPDKTDVRRPGKRDFKSMDQRELDSFYSGES
QMDGEFPSNIYFPGEPTYITHRRRRVSPSYWQRRHRVSNGQHEELAGVVAKLQQEVTELKSQNGTQMPLSHHTNIPEGTRDPRISILLKQLQSVSGLCSSQNTTSTPHTDTVG
QDVNAVEASSKAPLIQGSTADDADMFANQMMVGRC

>ORF34

MTARYGFGSISFPNKCGIFLSTTKNFIAPNFPIHYWTAPAFELRGRMNPDLEKNTLTLKNAAAVAALDNLRGETITLPTEIDRRLKPLEEQLTRMAKVLDSLETAAAEAEEADAQSE
ECTRTEIIRNESIHPEVQIAKNDAPLQYDTNFQVDFITLVYLGRARGNNSPGIVFGPWYRTLQERLVLDRPVAARGVDCKDGRISRTFMNTTVTCLQSAGRMYVGDRAYSAFEC
AVLCLYLMYRTSNSVHEPQVSSFGNLIEHLPEYTETFVNYMTTHENKNSYQFCYDRLPRDQFHARGGRYDQGALTSHSVMDALIRLQVLPPAPGQFNPGVNDIIDRNHTAYVD
KIQQAAAAYLERAQNVFLMEDQTLLRLTIDTITALLLLRRLLWNGNVYGDKLKNNFQLGLIVSEATGTPTNNVILRGATGFDGKFKSGNNNFQFLCERYIAPLYTLNRTTELTEMFP
GLVALCLDAHTQLSRGSLGRTVIDISSGQYQDRLISLIALELEHRRQNVTSLPIAAVVSIHDSVMLQYERGLGMLMHQPRVRAALEESRRLAQFNVNSDYDLLYFVCLGVIPQFAS
TP

>ORF35

MSASRIRAKCFRLGQRCHTRFYDVLKKDIDNVRRGFADAFNPRLAKLLSPLSHVDVQRAVRISMSFEVNLGRRRPDCVCIIQTESSGAGKTVCFIVELKSCRFSANIHTPTKYHQ
FCEGMRQLRDTMALIKETTPTGSDEIMVTPLLVFVSQRGLNLLQVTRLPPKVIHGNLVMLASHLENVAEYTPPIRSVRERRRLCKKKIHVCSLAKKRAKSCHRSALTKFEENAAC
GVDLPLRRPSLGACGGILQSITGMFSHG

>ORF36

MSTDKTDVKMGVLRIYLDGAYGIGKTTAAEEFLHHFAITPNRILLIGEPLSYWRNLAGEDAICGIYGTQTRRLNGDVSPEDAQRLTAHFQSLFCSPHAIMHAKISALMDTSTSDLV
QVNKEPYKIMLSDRHPIASTICFPLSRYLVGDMSPAALPGLLFTLPAEPPGTNLVVCTVSLPSHLSRVSKRARPGETVNLPFVMVLRNVYIMLINTIIFLKTNNWHAGWNTLSFCN
DVFKQKLQKSECIKLREVPGIEDTLFAVLKLPELCGEFGNILPLWAWGMETLSNCSRSMSPFVLSLEQTPQHAAQELKTLLPQMTPANMSSGAWNILKELVNAVQDNTS

>ORF37 / ORF37N

MFALVLAVVILPLWTTANKSYVTPTPATRSIGHMSALLREYSDRNMSLKLEAFYPTGFDEELIKSLHWGNDRKHVFLVIVKVNPTTHEGDVGLVIFPKYLLSPYHFKAEHRAPFPA
GRFGFLSHPVTPDVSFFDSSFAPYLTTQHLVAFTTFPPNPLVWHLERAETAATAERPFGVSLLPARPTVPKNTILEHKAHFATWDALARHTFFSAEAIITNSTLRIHVPLFGSVWPI
RYWATGSVLLTSDSGRVEVNIGVGFMSSLISLSSGPPIELIVVPHTVKLNAVTSDTTWFQLNPPGPDPGPSYRVYLLGRGLDMNFSKHATVDICAYPEESLDYRYHLSMAHTEAL
RMTTKADQHDINEESYYHIAARIATSIFALSEMGRTTEYFLLDEIVDVQYQLKFLNYILMRIGAGAHPNTISGTSDLIFADPSQLHDELSLLFGQVKPANVDYFISYDEARDQLKTAY
ALSRGQDHVNALSLARRVIMSIYKGLLVKQNLNATERQALFFASMILLNFREGLENSSRVLDGRTTLLLMTSMCTAAHATQAALNIQEGLAYLNPSKHMFTIPNVYSPCMGSLRT
DLTEEIHVMNLLSAIPTRPGLNEVLHTQLDESEIFDAAFKTMMIFTTWTAKDLHILHTHVPEVFTCQDAAARNGEYVLILPAVQGHSYVITRNKPQRGLVYSLADVDVYNPISVVYL
SRDTCVSEHGVIETVALPHPDNLKECLYCGSVFLRYLTTGAIMDIIIIDSKDTERQLAAMGNSTIPPFNPDMHGDDSKAVLLFPNGTVVTLLGFERRQAIRMSGQYLGASLGGAFL
AVVGFGIIGWMLCGNSRLREYNKIPLT

>ORF38

MEFPYHSTVSYNGVTFYFNERATRAYFICGGCLISIPRKHGGEIAKFGHVVRGVGPGDRSVASYVRSELNRTGKTWAVSSNNNCVFLDRVALLAAGSGAVDRDLCGTFDVEVE
DPTLADYLVSLPVTHLTLVAGVDVTRENKLKLFPTPTAINTTNGFMYVPNEASFSLVYMRMLELPESLQELVSGLFDGTPEIRDALNGSNDDEKTSIIVSRRAADVVTEDVKADDV
PISGEPYSEKQPRRRKKSDHITLSNFVQIRTIPRVMDIWDPRHKATTHCIRALSCAVFFADEVIFKARKWPGLEDELNEARETIYTAVVAVYGERGELPFFGHAYGRDLTSCQRFV
IVQYILSRWEAFNCYAVIEDLTRSYVNALPSDDDTDQVAQDLIRTIVDTANSLLREVGFIGTLAETLLFLPLPQLPCYKETSHLAKKEGVRILRLAKTGVGLSDTVPVDVSVTERHEY

EISRYLDTLYSGDPCYNGAVRLCRLLGSSIPIALYYNTISGNAFEPYFAGRRYIAYLGALFFGRVHQTPFGDGKKTQR

>ORF39 / ORF39N

MNPPQARVSEQTKDLLSVMVNQHPEEDAKVCKSSDNSPLYNTMVMLSYGGDTDLLLSSACTRTSTVNRSAFTQHSVFYIISTVLIQPICCIFFFFYYKATRCMLLFTAGLLLTILH
HFRLIIMLLCVYRNIRSDLLPLSTSQQLLLGIIVVTRTMLFCITAYYTLFIDTRVFFLITGHLQSEVIFPDSVSKILPVSWGPSPAVLLVMAAVIYAMDCLVDTVSFIGPRVWVRVMLKT
SISF

>ORF40

MTTVSCPANVITTTESDRIAGLFNIPAGIIPTGNVLSTIEVCAHRCIFDFFKQIRSDDNSLYSAQFDILLGTYCNTLNFVRFLELGLSVACICTKFPELAYVRDGVIQFEVQQPMIARD
GPHPVDQPVHNYMVKRIHKRSLSAAFAIASEALSLLSNTYVDGTEIDSSLRIRAIQQMARNLRTVLDSFERGTADQLLGVLLEKAPPLSLLSPINKFQPEGHLNRVARAALLSDLK
RRVCADMFFMTRHAREPRLISAYLSDMVSCTQPSVMVSRITHTNTRGRQVDGVLVTTATLKRQLLQGILQIDDTAADVPVTYGEMVLQGTNLVTALVMGKAVRGMDDVARHLL
DITDPNTLNIPSIPPQSNSDSTTAGLPVNARVPADLVIVGDKLVFLEALERRVYQATRVAYPLIGNIDITFIMPMGVFQANSMDRYTRHAGDFSTVSEQDPRQFPPQGIFFYNKDGI
LTQLTLRDAMGTICHSSLLDVEATLVALRQQHLDRQCYFGVYVAEGTEDTLDVQMGRFMETWADMMPHHPHWVNEHLTILQFIAPSNPRLRFELNPAFDFFVAPGDVDLPGPQ
RPPEAMPTVNATLRIINGNIPVPLCPISFRDCRGTQLGLGRHTMTPATIKAVKDTFEDRAYPTIFYMLEAVIHGNERNFCALLRLLTQCIRGYWEQSHRVAFVNNFHMLMYITTYL
GNGELPEVCINIYRDLLQHVRALRQTITDFTIQGEGHNGETSEALNNILTDDTFIAPILWDCDALIYRDEAARDRLPAIRVSGRNGYQALHFVDMAGHNFQRRDNVLIHGRPVRGD
TGQGIPITPHHDREWGILSKIYYYIVIPAFSRGSCCTMGVRYDRLYPALQAVIVPEIPADEEAPTTPEDPRHPLHAHQLVPNSLNVYFHNAHLTVDGDALLTLQELMGDMAERTTAI
LVSSAPDAGAATATTRNMRIYDGALYHGLIMMAYQAYDETIATGTFFYPVPVNPLFACPEHLASLRGMTNARRVLAKMVPPIPPFLGANHHATIRQPVAYHVTHSKSDFNTLTYS
LLGGYFKFTPISLTHQLRTGFHPGIAFTVVRQDRFATEQLLYAERASESYFVGQIQVHHHDAIGGVNFTLTQPRAHVDLGVGYTAVCATAALRCPLTDMGNTAQNLFFSRGGVP
MLHDNVTESLRRITASGGRLNPTEPLPIFGGLRPATSAGIARGQASVCEFVAMPVSTDLQYFRTACNPRGRASGMLYMGDRDADIEAIMFDHTQSDVAYTDRATLNPWASQKH
SYGDRLYNGTYNLTGASPIYSPCFKFFTPAEVNTNCNTLDRLLMEAKAVASQSSTDTEYQFKRPPGSTEMTQDPCGLFQEAYPPLCSSDAAMLRTAHAGETGADEVHLAQYLI
RDASPLRGCLPLPR

>ORF41

MAMPFEIEVLLPGELSPAETSALQKCEGKIITFSTLRHRASLVDIALSSYYINGAPPDTLSLLEAYRMRFAAVITRVIPGKLLAHAIGVGTPTPGLFIQNTSPVDLCNGDYICLLPPVF
GSADSIRLDSVGLEIVFPLTIPQTLMREIIAKVVARAVERTAAGAQILPHEVLRGADVICYNGRRYELETNLQHRDGSDAAIRTLVLNLMFSINEGCLLLLALIPTLLVQGAHDGYVNL
LIQTANCVRETGQLINIPPMPRIQDGHRRFPIYETISSWISTSSRLGDTLGTRAILRVCVFDGPSTVHPGDRTAVIQV

>ORF42

MPRVLAHSDVTACSCYVLNKPVFITMDGAMRRTADLFMADSFVQEIVGGRKQNSGGVGFDRPLFTKTARERFILYRPSTVANCAILSSVLYVYVDPAFTSNTRASGTGVAIVGR
YKSDWIIFGLEHFFLRALTGTSSSEIGRCVTQCLGHILALHPNTFTNVHVSIEGNSSQDSAVAISLAIAQQFAVLEKGNVLSSAPVLLFYHSIPPGCSVAYPFFLLQKQKTPAVDYFV
KRFNSGNIIASQELVSLTVKLGVDPVEYLCKQLDNLTEVIKGGMGNLDTKTYTGKGTTGTMSDDLMVALIMSVYIGSSCIPDSVFMPIK

>ORF43 / ORF43C

MEAHLANETKHALWHNDHTKGLLHVVIPNAGLIAAGIDPALLILKKPGQRFKVEVQTRYHATGQCEPWCQVFAAYIPDNALTNLLIPKTEPFVSHVFSATHNSGGLILSLPVYLSP
GLFFDAFNVVAIRINTGNRKHRDICIMYAELIPNGTRYFADGQRVLLLCKQLIAYIRCTPRLASSIKIYAEHMVAAMGESHTSNGDNIGPVSSIIDLDRQLTSGGIDDSPAETRIQENN
RDVLELIKRAVNIVNSRHPVRPSSSRVASGLLQSAKGHGAQTSNTDPINNGSFDGVLEPPGQGGRFTGKKNNSSASIPPLQDVLLFTPASTEPQSLMEWFDICYAQLVSGDTPAD
FWKRRPLSIVPRHYAESPSPLIVVSYNGSSAWGGRITGSPILYHSAQAIIDAACINARVDNPQSLHVTARQELVARLPFLANVLNNQTPLPAFKPGAEMFLNQVFKQACVTSLTQ
GLITELQTNPTLQQLMEYDIADSSQTVIDEIVARTPDLIQTIVSVLTEMSMDAFYNSSLMYAVLAYLSSVYTRPQGGGYIPYLHASFPCWLGNRSIYLFDYYNSGGEILKLSKVPVP
VALEKVGIGNSTQLRGKFIRSADIVDIGICSKYLPGQCYAYICLGFNQQLQSILVLPGGFAACFCITDTLQAALPASLIGPILDRFCFSIPNPHK

>ORF44

MELQRIFPLYTATGAARKLTPEAVQRLCDALTLDMGLWKSILTDPRVKIMRSTAFITLRIAPFIPLQTDTTNIAVVVATIYITRPRQMNLPPKTFHVIVNFNYEVSYAMTATLRIYPVE
NIDHVFGATFKNPIAYPLPTSIPDPRADPTPADLTPTPNLSNYLQPPRLPKNPYACKVISPGVWWSDERRRLYVLAMEPNLIGLCPAGWHARILGSVLNRLLSHADGCDECNHRV
HVGALYALPHVTNHAEGCVCWAPCMWRKAGQRELKVEVDIGATQVLFVDVTTCIRITSTKNPRITANLGDVIAGTNASGLSVPVNSSGWQLYMFGETLSRAIINGCGLLQRICFP
ETQRLSGEPEPTTT

>ORF45

MSLIMFGRTLGEESVRYFERLKRRRDERFGTLESPTPCSTRQGSLGNATQIPFLNFAIDVTRRHQAVIPGIGTLHNCCEYIPLFSATARRAMFGAFLSSTGYNCTPNVVLKPWRY
SVNANVSPELKKAVSSVQFYEYSPEEAAPHRNAYSGVMNTFRAFSLSDSFCQLSTFTQRFSYLVETSFESIEECGSHGKRAKVDVPIYGRYKGTLELFQKMILMHTTHFISSVLL
GDHADRVDCFLRTVFNTPSVSDSVLEHFKQKSTVFLVPRRHGKTWFLVPLIALVMATFRGIKVGYTAHIRKATEPVFEGIKSRLEQWFGANYVDHVKGESITFSFTDGSYSTAVF
ASSHNTNVSVL

>ORF46

MSGHTPTYASHRRNRVKLVEAHNRAGLFKERTLDLIRGGASVQDPAFVYAFTAAKEACADLNNQLRSAARIASVEQKIRDIQSKVEEQTSIQQILNTNRRYIAPDFIRGLDKTEDD
NTDNIDRLEDAVGPNIEHENHTWFGEDDEALLTQWMLTTHPPTSKYLQLQDLCVPTTIPTDMNQMQPQPISKNENPPTPHTDV

>ORF47

MDADDTPPNLQISPTAGPLRSHHNTDGHEPNATAADQQERESTNPTHGCVNHPWANPSTATCMESPERSQQTSLFLLKHGLTRDPIHQRERVDVFPQFNKPPWVFRISKLSR
LIVPIFTLNEQLCFSKLQIRDRPRFAGRGTYGRVHIYPSSKIAVKTMDSRVFNRELINAILASEGSIRAGERLGISSIVCLLGFSLQTKQLLFPAYDMDMDEYIVRLSRRLTIPDHIDRK
IAHVFLDLAQALTFLNRTCGLTHLDVKCGNIFLNVDNFASLEITTAVIGDYSLVTLNTYSLCTRAIFEVGNPSHPEHVLRVPRDASQMSFRLVLSHGTNQPPEILLDYINGTGLTKYT
GTLPQRVGLAIDLYALGQALLEVILLGRLPGQLPISVHRTPHYHYYGHKLSPDLALDTLAYRCVLAPYILPSDIPGDLNYNPFIHAGELNTRISRNSLRRIFQCHAVRYGVTHSKLFE
GIRIPASLYPATVVTSLLCHDNSEIRSDHPLLWHDRDWIGST

>ORF48

MARSGLDRIDISPQPAKKIARVGGLQHPFVKTDINTINVEHHFIDTLQKTSPNMDCRGMTAGIFIRLSHMYKILTTLESPNDVTYTTPGSTNALFFKTSTQPQEPRPEELASKLTQD
DIKRILLTIESETRGQGDNAIWTLLRRNLITASTLKWSVSGPVIPPQWFYHHNTTDTYGDAAAMAFGKTNEPAARAIVEALFIDPADIRTPDHLTPEATTKFFNFDMLNTKSPSLLV
GTPRIGTYECGLLIDVRTGLIGASLDVLVCDRDPLTGTLNPHPAETDISFFEIKCRAKYLFDPDDKNNPLGRTYTTLINRPTMANLRDFLYTIKNPCVSFFGPSANPSTREALITDHV
EWKRLGFKGGRALTELDAHHLGLNRTISSRVWVFNDPDIQKGTITTIAWATGDTALQIPVFANPRHANFKQIAVQTYVLSGYFPALKLRPFLVTFIGRVRRPHEVGVPLRVDTQA
AAIYEYNWPTIPPHCAVPVIAVLTPIEVDVPRVTQILKDTGNNAITSALRSLRWDNLHPAVEEESVDCANGTTSLLRATEKPLL

>ORF49

MGQSSSSGRGGICGLCKRYNELVTCNGETVALNSEFFEDFDFDENVTEDADKSTQRRPRVIDVTPKRKPSGKSSHSKCAKC

>ORF50 / ORF50C

MGTQKKGPRSEKVSPYDTTTPEVEALDHQMDTLNWRIWIIQVMMFTLGAVMLLATLIAASSEYTGIPCFYAAVVDYELFNATLDGGVWSGNRGGYSAPVLFLEPHSVVAFTYYT
ALTAMAMAVYTLITAAIIHRETKNQRVRQSSGVAWLVVDPTTLFWGLLSLWLLNAVVLLLAYKQIGVAATLYLGHFATSVIFTTYFCGRGKLDETNIKAVANLRQQSVFLYRLAGPT
RAVFVNLMAALMAICILFVSLMLELVVANHLHTGLWSSVSVAMSTFSTLSVVYLIVSELILAHYIHVLIGPSLGTLVACATLGTAAHSYMDRLYDPISVQSPRLIPTTRGTLACLAVFS
VVMLLLRLMRAYVYHRQKRSRFYGAVRRVPERVRGYIRKVKPAHRNSRRTNYPSQGYGYVYENDSTYETDREDELLYERSNSGWE

>ORF51

MSPNTGESNAAVYASSTQLARALYGGDLVSWIKHTHPGISLELQLDVPVKLIKPGMSQTRPVTVVRAPMGSGKTTALLEWLQHALKADISVLVVSCRRSFTQTLIQRFNDAGLS
GFVTYLTSETYIMGFKRLIVQLESLHRVSSEAIDSYDVLILDEVMSVIGQLYSPTMRRLSAVDSLLYRLLNRCSQIIAMDATVNSQFIDLISGLRGDENIHTIVCTYAGVGFSGRTCTI

LRDMGIDTLVRVIKRSPEHEDVRTIHQLRGTFFDELALRLQCGHNICIFSSTLSFSELVAQFCAIFTDSILILNSTRPLCNVNEWKHFRVLVYTTVVTVGLSFDMAHFHSMFAYIKPM
SYGPDMVSVYQSLGRVRLLLLNEVLMYVDGSRTRCGPLFSPMLLNFTIANKFQWFPTHTQITNKLCCAFRQRCANAFTRSNTHLFSRFKYKHLFERCSLWSLADSINILQTLLAS
NQILVVLDGMGPITDVSPVQFCAFIHDLRHSANAVASCMRSLRQDNDSCLTDFGPSGFMADNITAFMEKYLMESINTEEQIKVFKALACPIEQPRLVNTAILGACIRIPEALEAFDV
FQKIYTHYASGWFPVLDKTGEFSIATITTAPNLTTHWELFRRCAYIAKTLKWNPSTEGCVTQVLDTDINTLFNQHGDSLAQLIFEVMRCNVTDAKIILNRPVWRTTGFLDGCHNQC
FRPIPTKHEYNIALFRLIWEQLFGARVTKSTQTFPGSTRVKNLKKKDLETLLDSINVDRSACRTYRQLYNLLMSQRHSFSQQRYKITAPAWARHVYFQAHQMHLAPHAEAMLQL
ALSELSPGSWPRINGAVNFESL

>ORF52

MDATQITLVRESGHICAASIYTSWTQSGQLTQNGLSVLYYLLCKNSCGKYVPKFAEITVQQEDLCRYSRHGGSVSAATFASICRAASSAALDAWPLEPLGNADTWRCLHGTALA
TLRRVLGFKSFYSPVTFETDTNTGLLLKTIPDEHALNNDNTPSTGVLRANFPVAIDVSAVSACNAHTQGTSLAYARLTALKSNGDTQQQTPLDVEVITPKAYIRRKYKSTFSPPIER
EGQTSDLFNLEERRLVLSGNRAIVVRVLLPCYFDCLTTDSTVTSSLSILATYRLWYAAAFGKPGVVRPIFAYLGPELNPKGEDRDYFCTVGFPGWTTLRTQTPAVESIRTATEMY
METDGLWPVTGIQAFHYLAPWGQHPPLPPRVQDLIGQIPQDTGHADATVNWDAGRISTVFKQPVQLQDRWMAKFDFSAFFPTIYCAMFPMHFRLGKIVLARMRRGMGCLKPA
LVSFFGGLRHILPSIYKAIIFIANEISLCVEQTALEQGFAICTYIKDGFWGIFTDLHTRNVCSDQARCSALNLAATCERAVTGLLRIQLGLNFTPAMEPVLRVEGVYTHAFTWCTTGS
WLWNLQTNTPPDLVGVPWRSQAARDLKERLSGLLCTATKIRERIQENCIWDHVLYDIWAGQVVEAARKTYVDFFEHVFDRRYTPVYWSLQEQNSETKAIPASYLTYGHMQDK
DYKPRQIIMVRNPNPHGPPTVVYWELLPSCACIPPIDCAAHLKPLIHTFVTIINHLLDAHNDFSSPSLKFTDDPLASYNFLFL

>ORF53

MQRIRPYWIKFEQTGGAGMADGMSGINIPSILGCSVTIDNLLTRAEEGLDVSDVIEDLRIQAIPRFVCEAREVTGLKPRFLANSVVSLRVKPEHQETVLVVLNGDSSEVSCDRYYM
ECVTQPAFRGFIFSVLTAVEDRVYTVGVPPRLLIYRMTLFRPDNVLDFTLCVILMYLEGIGPSGASPSLFVQLSVYLRRVECQIGPLEKMRRFLYEGVLWLLNTLMYVVDNNPFTK
TRVLPHYMFVKLLNPQPGTAPNIIKAIYSCGVGQRFDLPHGTPPCPDGVVQVPPGLLNGPLRDSEYQKSVYFWWLNRTMVTPKNVQLFETYKNSPRVVK

>ORF54

MAEITSLFNNSSGSEEKRIASSVSIDQGLNGSNPNDQYKNMFDIYWNEYAPDIGFCTFPEEDGWMLIHPTTQSMLFRKILAGDFGYTDGQGIYSAVRSTETVIRQVQATVLMNAL
DATRYEDLAADWEHHIQQCNLHAGALAERYGLCGESEAVRLAHQVFETWRQTLQSSLLEFLRGITGCLYTSGLNGRVGFAKYVDWIACVGIVPVVRKVRSEQNGTPAPLNTYM
GQAAELSQMLKVADATLARGAAVVTSLVECMQNVAIMDYDRTRLYYNYNRRLIMAKDDVTGMKGECLVVWPPVVCGEGVVFDSPLQRLSGEVLACYALREHARVCQVLNTAP
LRVLIGRRNEDDRSHSTRAVDRIMGENDTTRAGSAASRLVKLIVNLKNMRHVGDITETVRSYLEETGNHILEGSGSVDTSQPGFGKANQSFNGGAMSGTTNVQSAFKTSVVNSI
NGMLEGYVNNLFKTIEGLKDVNSDLTERLQFKEGELKRLREERVKIKPSKGSHITMAEETRIADLNHEVIDLTGIIGDDAYIANSFQSRYIPPYGDDIKRLSELWKQELVRCFKLHR
VNNNQGQEISVSYSNASISLLVAPYFSFILRATRLGFLVTQSEVHRSEEELCQAIFKKARTESYLSQIRILYEMQVRAEVIKRGPRRTPSPSWGLPDPTEDDERIPEPNKINNQYM
HVGYKNLSHFMKGHPPERLRVHKVNAADSTLLDKIRANRRRGDGRWDVRNKYTQHFRLQRNDRQLTNTSRRGVGCERRDRRS

>ORF55

MKRSISVDSSSPKNVFNPETPNGFDDSVYLNFTSMHSIQPILSRIRELAAITIPKERVPRLCWFKQLLELQAPPEMQRNELPFSVYLISGNAGSGKSTCIQTLNEAIDCIITGSTRVA
AQNVHAKLSTAYASRPINTIFHEFGFRGNHIQAQLGRYAYNWTTTPPSIEDLQKRDIVYYWEVLIDITKRVFQMGDDGRGGTSTFKTLWAIERLLNKPTGSMSGTAFIACGSLPAF
TRSNVIVIDEAGLLGRHILTAVVYCWWLLNAIYQSPQYINGRKPVIVCVGSPTQTDSLESHFQHDMQRSHVTPSENILTYIICNQTLRQYTNISHNWAIFINNKRCQEDDFGNLLKT
LEYGLPITEAHARLVDTFVVPASYINNPANLPGWTRLYSSHKEVSAYMSKLHAHLKLSKNDHFSVFALPTYTFIRLTAFDEYRKLTGQPGLSVEHWIRANSGRLHNYSQSRDHD
MGTVKYETHSNRDLIVARTDITYVLNSLVVVTTRLRKLVIGFSGTFQSFAKVLRDDSFVKARGETSIEYAYRFLSNLIFGGLINFYNFLLNKNLHPDKVSLAYKRLAALTLELLSGTN
KAPLHEAAVNGAGAGIDCDGAATSADKAFCFTKAPESKVTASIPEDPDDVIFTALNDEVIDLVYCQYEFSYPKSSNEVHAQFLLMKAIYDGRYAILAELFESSFTTAPFSAYVDNV
NFNGSELLIGNVRGGLLSLALQTDTYTLLGYTFAPVPVFVEELTRKKLYRETTEMLYALHVPLMVLQDQHGFVSIVNANVCEFTESIEDAELAMATTVDYGLSSKLAMTIARSQGL
SLEKVAICFTADKLRLNSVYVAMSRTVSSRFLKMNLNPLRERYEKSAEISDHILAALRDPNVHVVY

>ORF56 / ORF56C

MKNPQKLAITFLPLYVIPTYTLCIKALYKNTHAGLLFSFLGFVLNTPAMSISGPPTTFILYRLHGVRRVLHWTLPDHEQTLYAFTGGSRSMAVKTDARCDTMSGGMIVLQHTHTVT LLTIDCSTDFSSYAFTHRDFHLQDKPHATFAMPFMSWVGSDPTSQLYSNVGGVLSVITEDDLSMCISIVIYGLRVNRPDDQTTPTPTPHQYTSQRRQPETNCPSSPQPAFFTSD DDVLSLILRDAANA
>ORF57
MDVRERNVFGNASVATPGEHQKFVRELILSGHNNVVLQTYTGKWSDCRKHGKSVMYNTGEARHPTCKAHQR
>ORF58
MFSELPPSVPTALLQWGWGLHRGPCSIPNFKQVASQHSVQNDFTENSVDANEKFPIGHAGCIEKTKDDYVPFDTLFMVSSIDELGRRQLTDTIRRSLVMNACEITVACTKTAAF SGRGVSRQKHVTLSKNKFNPSSHKSLQMFVLCQKTHAPRVRNLLYESIRARRPRRYYTRSTDGKSRPLVPVFVYEFTALDRVLLHKENTLTDQPINTENSGHGRTRT
>ORF59
MDVSGEPTVCSNAYANEMKLSDSKDIYVLAHPVTKKTRKRPRGLPLGVKLDPPTFKLNNMSHHYDTETFTPVSSQLDSVEVFSKFNISPEWYDLLSDELKEPYAKGIFLEYNRLL NSGEEILPSTGDIFAWTRFCGPQSIRVVIIGQDPYPTAGHAHGLAFSVKRGITPPSSLKNIFAALMESYPNMTPPTHGCLESWARQGVLLLNTTLTVRRGTPGSHVYLGWGRLVQ RVLQRLCENRTGLVFMLWGAHAQKTTQPNSRCHLVLTHAHPSPLSRVPFRNCRHFVQANEYFTRKGEPEIDWSVI
>ORF60 / ORF60C
MASHKWLLQIVFLKTITIAYCLHLQDDTPLFFGAKPLSDVSLIITEPCVSSVYEAWDYAAPPVSNLSEALSGIVVKTKCPVPEVILWFKDKQMAYWTNPYVTLKGLAQSVGEEHKS GDIRDALLDALSGVWVDSTPSSTNIPENGCVWGADRLFQRVCQ
>ORF61
MDTILAGGSGTSDASDNTCTICMSTVSDLGKTMPCLHDFCFVCIRAWTSTSVQCPLCRCPVQSILHKIVSDTSYKEYEVHPSDDDGFSEPSFEDSIDILPGDVIDLLPPSPGPSRE SIQQPTSRSSREPIQSPNPGPLQSSAREPTAESPSDSQQDSIQPPTRDSSPGVTKTCSTASFLRKVFFKDQPAVRSATPVVYGSIESAQQPRTGGQDYRDRPVSVGINQDPRT MDRLPFRATDRGTEGNARFPCYMQPLLGWLDDQLAELYQPEIVEPTKMLILNYIGIYGRDEAGLKTSLRCLLHDSTGPFVTNMLFLLDRCTDPTRLTMQTWTWKDTAIQLITGPI VRPETTSTGETSRGDERDTRLVNTPQKVRLFSVLPGIKPGSARGAKRRLFHTGRDVKRCLTIDLTSESDSACKGSKTRKVASPQGESNTPSTSGSTSGSLKHLTKKSSAGKAG KGIPNKMKKS
>ORF62
MDTPPMQRSTPQRAGSPDTLELMDLLDAAAAAAEHRARVVTSSQPDDLLFGENGVMVGREHEIVSIPSVSGLQPEPRTEDVGEELTQDDYVCEDGQDLMGSPVIPLAEVFHT RFSEAGAREPTGADRSLETVSLGTKLARSPKPPMNDGETGRGTTPPFPQAFSPVSPASPVGDAAGNDQREDQRSIPRQTTRGNSPGLPSVVHRDRQTQSISGKKPGDEQAG HAHASGDGVVLQKTQRPAQGKSPKKKTLKVKVPLPARKPGGPVPGPVEQLYHVLSDSVPAKGAKADLPFETDDTRPRKHDARGITPRVPGRSSGGKPRAFLALPGRSHAPDP IEDDSPVEKKPKSREFVSSSSSSSSWGSSSEDEDDEPRRVSVGSETTGSRSGREHAPSPSNSDDSDSNDGGSTKQNIQPGYRSISGPDPRIRKTKRLAGEPGRQRQKSFSLP RSRTPIIPPVSGPLMMPDGSPWPGSAPLPSNRVRFGPSGETREGHWEDEAVRAARARYEASTEPVPLYVPELGDPARQYRALINLIYCPDRDPIAWLQNPKLTGVNSALNQFY QKLLPPGRAGTAVTGSVASPVPHVGEAMATGEALWALPHAAAAVAMSRRYDRAQKHFILQSLRRAFASMAYPEATGSSPAARISRGHPSPTTPATQAPDPQPSAAARSLSVC PPDDRLRTPRKRKSQPVESRSLLDKIRETPVADARVADDHVVSKAKRRVSEPVTITSGPVVDPPAVITMPLDGPAPNGGFRRIPRGALHTPVPSDQARKAYCTPETIARLVDDPL FPTAWRPALSFDPGALAEIAARRPGGGDRRFGPPSGVEALRRRCAWMRQIPDPEDVRLLIIYDPLPGEDINGPLESTLATDPGPSWSPSRGGLSVVLAALSNRLCLPSTHAWA GNWTGPPDVSALNARGVLLLSTRDLAFAGAVEYLGSRLASARRRLLVLDAVALERWPRDGPALSQYHVYVRAPARPDAQAVVRWPDSAVTEGLARAVFASSRTFGPASFARI ETAFANLYPGEQPLCLCRGGNVAYTVCTRAGPKTRVPLSPREYRQYVLPGFDGCKDLARQSRGLGLGAADFVDEAAHSHRAANRWGLGAALRPVFLPEGRRPGAAGPEAGD VPTWARVFCRHALLEPDPAAEPLVLPPVAGRSVALYASADEARNALPPIPRVMWPPGFGAAETVLEGSDGTRFVFGHHGGSERPSETQAGRQRRTADDREHALELDDWEVG CEDAWDSEEGGGDDGDAPGSSFGVSIVSVAPGVLRDRRVGLRPAVKVELLSSSSSSEDEDDVWGGRGGRSPPQSRG

>ORF63
MFCTSPATRGDSSESKPGASVDVNGKMEYGSAPGPLNGRDTSRGPGAFCTPGWEIHPARLVEDINRVFLCIAQSSGRVTRDSRRLRRICLDFYLMGRTRQRPTLACWEELLQ LQPTQTQCLRATLMEVSHRPPRGEDGFIEAPNVPLHRSALECDVSDDGGEDDSDDDGSTPSDVIEFRDSDAESSDGEDFIVEEESEESTDSCEPDGVPGDCYRDGDGCNTPS PKRPQRAIERYAGAETAEYTAAKALTALGEGGVDWKRRRHEAPRRHDIPPPHGV
>ORF64
MNLCGSRGEHPGGEYAGLYCTRHDTPAHQALMNDAERYFAAALCAISTEAYEAFIHSPSERPCASLWGRAKDAFGRMCGELAADRQRPPSVPPIRRAVLSLLREQCMPDPQS HLELSERLILMAYWCCLGHAGLPTIGLSPDNKCIRAELYDRPGGICHRLFDAYLGCGSLGVPRTYERS
>ORF65 / ORF65N
MAGQNTMEGEAVALLMEAVVTPRAQPNNTTITAIQPSRSAEKCYYSDSENETADEFLRRIGKYQHKIYHRKKFCYITLIIVFVFAMTGAAFALGYITSQFVG
>ORF66
MNDVDATDTFVGQGKFRGAISTSPSHIMQTCGFIQQMFPVEMSPGIESEDDPNYDVNMDIQSFNIFDGVHETEAEASVALCAEARVGINKAGFVILKTFTPGAEGFAFACMDSK TCEHVVIKAGQRQGTATEATVLRALTHPSVVQLKGTFTYNKMTCLILPRYRTDLYCYLAAKRNLPICDILAIQRSVLRALQYLHNNSIIHRDIKSENIFINHPGDVCVGDFGAACFPV DINANRYYGWAGTIATNSPELLARDPYGPAVDIWSAGIVLFEMATGQNSLFERDGLDGNCDSERQIKLIIRRSGTHPNEFPINPTSNLRRQYIGLAKRSSRKPGSRPLWTNLYEL PIDLEYLICKMLSFDARHRPSAEVLLNHSVFQTLPDPYPNPMEVGD
>ORF67 / ORF67N / *ORF67C*
MFLIQCLISAVIFYIQVTNALIFKGDHVSLQVNSSLTSILIPMQNDNYTEIKGQLVFIGEQLPTGTNYSGTLELLYADTVAFCFRSVQVIRYDGCPRIRTSAFISCRYKHSWHYGNSTD RISTEPDAGVMLKITKPGINDAGVYVLLVRLDHSRSTDGFILGVNVYTAGSHHNIHGVIYTSPSLQNGYSTRALFQQARLCDLPATPKGSGTSLFQHMLDLRAGKSLEDNPWLHE DVVTTETKSVVKEGIENHVYPTDMSTLPEKSLNDPPENLLIIIPIVASVMILTAMVIVIV*ISVKRRRIKKHPIYRPNTKTRRGIQNATPESDVMLEAAIAQLATIREESPPHSVVNPFVK*
>ORF68 / ORF68F / *ORF68C*
MGTVNKPVVGVLMGFGIITGTLRITNPVRASVLRYDDFHIDEDKLDTNSVYEPYYHSDHAESSWVNRGESSRKAYDHNSPYIWPRNDYDGFLENAHEHHGVYNQGRGIDSGER LMQPTQMSAQEDLGDDTGIHVIPTLNGDDRHKIVNVDQRQYGDVFKGDLNPKPQGQRLIEVSVEENHPFTLRAPIQRIYGVRYTETWSFLPSLTCTGDAAPAIQHICLKHTTCFQ DVVVDVDCAENTKEDQLAEISYRFQGKKEADQPWIVVNTSTLFDELELDPPEIEPGVLKVLRTEKQYLGVYIWNMRGSDGTSTYATFLVTWKGDEKTRNPTPAVTPQPRGAEF HMWNYHSHVFSVGDTFSLAMHLQYKIHEAPFDLLLEWLYVPIDPTCQPMRLYSTCLYHPNAPQCLSHMNSGCTFTSPHLAQRVASTVYQNCEHADNYTAYCLGISHMEPSFGL ILHDGGTTLKFVDTPESLSGLYVFVVYFNGHVEAVAYTVVSTVDHFVNAIEERGFPPTAGQPPATTKPKEITPVNPGTSPLLRYAAWTGGLAAVVLLCLVIFLIC*TAKRMRVKAYR VDKSPYNQSMYYAGLPVDDFEDSESTDTEEEFGNAIGGSHGGSSYTVYIDKTR*
>ORFS/L_MKKVSV / ORFS/L_C
MKKVSVCLCGRATGGLVGPPPENNPPPVSGRPADPGRGGQPSRGPLEREKKKRPHLPARLRGDHRGGWDFLPGNPPPPAFNKTRAFCVHPSFTARMATVHYSRRPGTPPV TLTSSPSMDDVATPIPYLPTYAEAVADAPPPYRSRESLVFSPPLFPHVENGTTQQSYDCLDCAYDGIHRLQLAFLRIRKCCVPAFLILFGILTLTAVVVAIVAVFPEEPPNSTT

Supplementary Table S5: All VZV proteins and fragments used in this study.
All proteins have been cloned as full-length ORFs (except ORF22 - 8kb; this was cloned only in half: ORF22N - 4kb). Underlined are fragments of proteins (N: N-terminal; C: C-terminal; F: fragments somewhere in between).If there are multiple fragments, one is underlined, the other underlined and in italics, e.g. ORF15 (normal), ORF15N (underlined), ORF15F (underlined italics). Find the file in the internet under: [http://www.biomedcentral.com/content/supplementary/ 1477-5956-8-8-S1.XLS

NN-Network					Combined Network				
Node ID	node degree	% nodes removed	average degree	rel. average deg.	Node ID	node degree	% nodes removed	average degree	rel. average deg.
all		0.00%	2.364	1	all		0.00%	2.138	1.000000
ORF25	33	1.82%	2.636	1.11505922	ORF25	46	1.49%	2.418	1.130964
ORF60	18	3.64%	2.687	1.13663283	ORF24	29	2.99%	2.446	1.144060
ORF27	18	5.45%	2.843	1.20262267	ORF60	28	4.48%	2.476	1.158092
ORF39	12	7.27%	2.959	1.25169205	ORF38	24	5.97%	2.659	1.243686
ORF33	12	9.09%	3.048	1.2893401	ORF18	23	7.46%	2.697	1.261459
ORF65	11	10.91%	3.084	1.30456853	ORF9a	22	8.96%	2.751	1.286717
ORF33.5	10	12.73%	3.037	1.28468697	ORF19	21	10.45%	2.775	1.297942
ORF23	10	14.55%	2.862	1.2106599	ORF39	21	11.94%	2.835	1.326006
ORF22	10	16.36%	3.241	1.37098139	ORF26	21	13.43%	2.877	1.345650
ORF62	9	18.18%	3.527	1.49196277	ORF27	20	14.93%	2.792	1.305893
ORF38	9	20.00%	2.89	1.22250423	ORF42	20	16.42%	2.844	1.330215
ORF21	9	21.82%	2.857	1.20854484	ORF62	18	17.91%	3.0009	1.403601
ORF19	9	23.64%	3.224	1.36379019	ORF33	18	19.40%	3.118	1.458372
ORF16	9	25.45%	2	0.84602369	ORF68	17	20.90%	3.21	1.501403
ORF24	8	27.27%	1.806	0.76395939	ORF3	16	22.39%	3.57	1.669785
ORF68	7	29.09%	1.829	0.77368866	ORF65	15	23.88%	3.679	1.720767
ORF42	7	30.91%	1.767	0.74746193	ORF23	15	25.37%	3.894	1.821328
ORF3	7	32.73%	1.609	0.68062606	ORF41	14	26.87%	4.11	1.922357
ORF18	7	34.55%	1.3125	0.55520305	ORF12	14	28.36%	4.297	2.009822
ORF9a	6	36.36%	1.286	0.54399323	ORF56	14	29.85%	4.523	2.115529
ORF41	6	38.18%	1.308	0.55329949	ORF9	13	31.34%	4.024	1.882133
ORF34	6	40.00%	1.273	0.53849408	ORF21	13	32.84%	2.971	1.389616
ORF8	5	41.82%	1.3	0.5499154	ORF16	13	34.33%	3.036	1.420019

NN-Network					Combined Network				
Node ID	node degree	% nodes removed	average degree	rel. average deg.	Node ID	node degree	% nodes removed	average degree	rel. average deg.
ORF58	5	43.64%	1.3	0.5499154	ORF43	12	35.82%	2.529	1.182881
ORF56	5	45.45%	1.3	0.5499154	ORF33.5	12	37.31%	2.25	1.052385
ORF50	5	47.27%	1.143	0.48350254	ORF57	11	38.81%	2.288	1.070159
ORF12	5	49.09%	1.143	0.48350254	ORF22	10	40.30%	2.288	1.070159
ORF61	4	50.91%	1	0.42301184	ORF28	10	41.79%	2.316	1.083255
ORF44	4	52.73%	1	0.42301184	ORF64	10	43.28%	2.345	1.096819
ORF43	4	54.55%	1	0.42301184	ORF53	10	44.78%	2	0.935454
ORF1	4	56.36%	1	0.42301184	ORF50	9	46.27%	1.72	0.804490
S/L	3	58.18%	1	0.42301184	ORF29	8	47.76%	1.72	0.804490
ORF7	3	60.00%	1	0.42301184	ORF4	8	49.25%	1.75	0.818522
ORF64	3	61.82%	1	0.42301184	ORF58	8	50.75%	1.75	0.818522
ORF36	3	63.64%	1	0.42301184	ORF1	7	52.24%	1.75	0.818522
ORF26	3	65.45%	1	0.42301184	ORF67	7	53.73%	1.773	0.829280
ORF66	2	67.27%	1	0.42301184	ORF8	7	55.22%	1.81	0.846586
ORF57	2	69.09%	1	0.42301184	ORF15	6	56.72%	1.81	0.846586
ORF55	2	70.91%	1	0.42301184	ORF44	6	58.21%	1.417	0.662769
ORF52	2	72.73%	1	0.42301184	ORF61	6	59.70%	1.143	0.534612
ORF46	2	74.55%	1	0.42301184	ORF34	6	61.19%	1.143	0.534612
ORF4	2	76.36%	1	0.42301184	ORF37	5	62.69%	1.143	0.534612
ORF32	2	78.18%	1	0.42301184	S/L	5	64.18%	1.167	0.545837
ORF15	2	80.00%	1	0.42301184	ORF32	5	65.67%	1.2	0.561272
ORF11	2	81.82%	1	0.42301184	ORF66	4	67.16%	1.2	0.561272
ORF67	1	83.64%	1	0.42301184	ORF46	4	68.66%	1.2	0.561272
ORF53	1	85.45%	1	0.42301184	ORF36	3	70.15%	1.2	0.561272

NN-Network					Combined Network				
Node ID	node degree	% nodes removed	average degree	rel. average deg.	Node ID	node degree	% nodes removed	average degree	rel. average deg.
ORF51	1	87.27%	1	0.42301184	ORF10	3	71.64%	1.2	0.561272
ORF49	1	89.09%	1	0.42301184	ORF7	3	73.13%	1.2	0.561272
ORF45	1	90.9%	N/A		ORF35	3	74.63%	1.25	0.584659
ORF31	1	92.7%	N/A		ORF11	3	76.12%	1	0.467727
ORF30	1	94.5%	N/A		ORF49	3	77.61%	1	0.467727
ORF28	1	96.4%	N/A		ORF31	3	79.10%	1	0.467727
ORF2	1	98.2%	N/A		ORF2	2	80.60%	1	0.467727
ORF10	1	100.0%	N/A		ORF13	2	82.1%	N/A	
					ORF14	2	83.6%	N/A	
					ORF55	2	85.1%	N/A	
					ORF48	2	86.6%	N/A	
					ORF52	2	88.1%	N/A	
					ORF6	2	89.6%	N/A	
					ORF51	1	91.0%	N/A	
					ORF20	1	92.5%	N/A	
					ORF63	1	94.0%	N/A	
					ORF30	1	95.5%	N/A	
					ORF59	1	97.0%	N/A	
					ORF45	1	98.5%	N/A	
					ORF17	1	100.0%	N/A	

Supplementary Table S6: Attack tolerance of the NN- and combinatorial PPI network.
The node degree of every protein in the VZV NN- and combinatorial network was calculated. The average degree was calculated after each node was removed according to the node degree in decreasing order. The average degree of the whole network was set to 1 and for every node removed the respective relative average degree was calculated.

NN Network		Combined Network	
Degree	Number of Nodes	Degree	Number of Nodes
1	10	1	6
2	9	2	8
3	2	3	5
4	6	4	4
5	5	5	3
6	4	6	3
7	4	7	2
8	1	8	4
9	3	9	1
10	5	10	4
11	1	11	2
13	2	12	2
18	1	13	2
19	1	14	3
34	1	15	3
		17	2
		18	1
		19	1
		21	2
		22	3
		23	1
		24	2
		29	1
		30	1
		47	1

Supplementary Table S7: Node degree distribution of the NN- and combinatorial PPI network.
The degree of a protein is synonymous to the number of interaction partners connected by an edge in the network graph. The number of nodes is the amount of proteins that share the same number of PPIs. Characteristic for scale-free networks is that many proteins have only a few interaction partners while a decreasing portion of proteins show many interactions, so called hub-proteins.

PLATE	ROW	COL	POS	CLASS	GENE_ID_X	SYMBOL_BAIT	ACCESSION I	GENE_ID_Y	SYMBOL_PREY	ACCESSION II
Plate1	A	1	1	gold_positive_fl	3937	LCP2	BC016618	4690	NCK1	BC006403
Plate1	B	1	2	gold_negative	4670	HNRPM	BC000138	8743	TNFSF10	BC032722
Plate1	C	1	3	gold_positive_fl	2171	FABP5	BC019385	6278	S100A7	BC034687
Plate1	D	1	4	gold_negative	6472	SHMT2	BC013677	246329	STAC3	BC008069
Plate1	E	1	5	gold_positive_fl	7189	TRAF6	BC031052	257397	TAB3	BC032526
Plate1	F	1	6	gold_negative	2010	EMD	BC000738	55156	ARMC1	BC011607
Plate1	G	1	7	gold_negative	5162	PDHB	BC000439	51530	ZC3HC1	BC011551
Plate1	H	1	8	gold_positive_fl	6277	S100A6	BC009017	6285	S100B	BC001766
Plate1	A	2	9	gold_negative	2287	FKBP3	BC016288	4835	NQO2	BC006096
Plate1	B	2	10	gold_positive_fl	9402	GRAP2	BC025692	27040	LAT	BC011563
Plate1	C	2	11	gold_negative	5865	RAB3B	BC005035	91653	BOC	BC034614
Plate1	D	2	12	gold_positive_fl	3937	LCP2	BC016618	9402	GRAP2	BC025692
Plate1	E	2	13	gold_positive_fl	27258	LSM3	BC007055	57819	LSM2	BC009192
Plate1	F	2	14	gold_negative	2584	GALK1	BC001166	56922	MCCC1	BC004187
Plate1	G	2	15	gold_negative	2820	GPD2	BC019874	79680	FLJ21125	BC011679
Plate1	H	2	16	gold_negative	5711	PSMD5	BC014478	55356	SLC22A15	BC026358
Plate1	A	3	17	gold_positive_fl	5747	PTK2	BC028733	6714	SRC	BC011566
Plate1	B	3	18	gold_positive_fl	3434	IFIT1	BC007091	3646	EIF3S6	BC008419
Plate1	C	3	19	gold_negative	5537	PPP6C	BC006990	59348	ZNF350	BC009921
Plate1	D	3	20	gold_negative	975	CD81	BC002978	10577	NPC2	BC002532
Plate1	E	3	21	gold_positive_fl	5710	PSMD4	BC002365	5886	RAD23A	BC014026
Plate1	F	3	22	gold_negative	685	BTC	BC011618	84726	KIAA0515	BC012289
Plate1	G	3	23	gold_positive_fl	4171	MCM2	BC007670	4172	MCM3	BC001626
Plate1	H	3	24	gold_positive_fl	578	BAK1	BC004431	598	BCL2L1	BC019307
Plate1	A	4	25	gold_positive_fl	4999	ORC2L	BC014834	55388	MCM10	BC009108
Plate1	B	4	26	gold_positive_fl	2908	NR3C1	BC015610	5970	RELA	BC011603
Plate1	C	4	27	gold_positive_fl	1017	CDK2	BC003065	1163	CKS1B	BC007751
Plate1	D	4	28	gold_negative	977	CD151	BC013302	55255	WDR41	BC040241
Plate1	E	4	29	gold_positive_fl	572	BAD	BC001901	598	BCL2L1	BC019307
Plate1	F	4	30	gold_positive_fl	3091	HIF1A	BC012527	7157	TP53	BC003596
Plate1	G	4	31	gold_positive_fl	5576	PRKAR2A	BC002763	7430	VIL2	BC013903
Plate1	H	4	32	gold_positive_fl	896	CCND3	BC011616	1021	CDK6	BC027989
Plate1	A	5	33	gold_positive_fl	627	BDNF	BC029795	4909	NTF5	BC012421
Plate1	B	5	34	gold_negative	5152	PDE9A	BC009047	161253	FLJ38964	BC035663

PLATE	ROW	COL	POS	CLASS	GENE_ID_X	SYMBOL_BAIT	ACCESSION I	GENE_ID_Y	SYMBOL_PREY	ACCESSION II
Plate1	C	5	35	gold_positive_fl	2246	FGF1	BC032697	2260	FGFR1	BC015035
Plate1	D	5	36	gold_negative	821	CANX	BC003552	29028	ATAD2	BC019909
Plate1	E	5	37	gold_positive_fl	4085	MAD2L1	BC000356	8379	MAD1L1	BC009964
Plate1	F	5	38	gold_positive_fl	3726	JUNB	BC004250	10538	BATF	BC032294
Plate1	G	5	39	gold_negative	2639	GCDH	BC002579	84240	ZCCHC9	BC032736
Plate1	H	5	40	gold_negative	2898	GRIK2	BC037954	151188	ARL6IP6	BC028741
Plate1	A	6	41	gold_negative	5995	RGR	BC011349	55324	ABCF3	BC009253
Plate1	B	6	42	gold_positive_fl	1649	DDIT3	BC003637	2353	FOS	BC004490
Plate1	C	6	43	gold_negative	3920	LAMP2	BC002965	7327	UBE2G2	BC001738
Plate1	D	6	44	gold_positive_fl	2919	CXCL1	BC011976	3579	IL8RB	BC037961
Plate1	E	6	45	gold_negative	5008	OSM	BC011589	146542	LOC146542	BC018997
Plate1	F	6	46	gold_negative	4331	MNAT1	BC000820	29926	GMPPA	BC007456
Plate1	G	6	47	gold_negative	2173	FABP7	BC012299	6811	STX5A	BC002645
Plate1	H	6	48	gold_negative	572	BAD	BC001901	8817	FGF18	BC006245
Plate1	A	7	49	gold_positive_fl	4097	MAFG	BC012327	4779	NFE2L1	BC010623
Plate1	B	7	50	gold_negative	6317	SERPINB3	BC005224	10671	DCTN6	BC013175
Plate1	C	7	51	gold_negative	3059	HCLS1	BC016758	6297	SALL2	BC024245
Plate1	D	7	52	gold_positive_fl	2908	NR3C1	BC015610	3320	HSPCA	BC023006
Plate1	E	7	53	gold_negative	5031	P2RY6	BC009391	7726	TRIM26	BC032297
Plate1	F	7	54	gold_negative	2277	FIGF	BC027948	7597	ZNF46	BC035804
Plate1	G	7	55	gold_negative	3705	ITPK1	BC018192	80723	TMEM22	BC022557
Plate1	H	7	56	gold_negative	5627	PROS1	BC015801	10494	STK25	BC007852
Plate1	A	8	57	gold_positive_fl	3065	HDAC1	BC000301	7704	ZBTB16	BC026902
Plate1	B	8	58	gold_positive_fl	467	ATF3	BC006322	1649	DDIT3	BC003637
Plate1	C	8	59	gold_negative	5156	PDGFRA	BC015186	80762	NDFIP1	BC004317
Plate1	D	8	60	gold_negative	3109	HLA-DMB	BC027175	5664	PSEN2	BC006365
Plate1	E	8	61	gold_positive_fl	3958	LGALS3	BC001120	3959	LGALS3BP	BC002403
Plate1	F	8	62	gold_negative	6529	SLC6A1	BC033904	7104	TM4SF4	BC001386
Plate1	G	8	63	gold_positive_fl	2237	FEN1	BC000323	5111	PCNA	BC000491
Plate1	H	8	64	gold_positive_fl	207	AKT1	BC000479	8115	TCL1A	BC003574
Plate1	A	9	65	gold_positive_fl	3481	IGF2	BC000531	3487	IGFBP4	BC016041
Plate1	B	9	66	gold_negative	5502	PPP1R1A	BC022470	414918	MGC33692	BC031069
Plate1	C	9	67	gold_negative	4841	NONO	BC012141	7439	VMD2	BC015220
Plate1	D	9	68	gold_positive_fl	5824	PEX19	BC000496	8799	PEX11B	BC011963

PLATE	ROW	COL	POS	CLASS	GENE_ID_X	SYMBOL_BAIT	ACCESSION I	GENE_ID_Y	SYMBOL_PREY	ACCESSION II
Plate1	E	9	69	gold_positive_fl	4686	NCBP1	BC001450	22916	NCBP2	BC001255
Plate1	F	9	70	gold_negative	2555	GABRA2	BC022488	9050	PSTPIP2	BC035395
Plate1	G	9	71	gold_positive_fl	2175	FANCA	BC008979	2189	FANCG	BC000032
Plate1	H	9	72	gold_positive_fl	1026	CDKN1A	BC000312	8900	CCNA1	BC036346
Plate1	A	10	73	gold_positive_fl	6500	SKP1A	BC009839	6502	SKP2	BC001441
Plate1	B	10	74	gold_positive_fl	868	CBLB	BC032851	2885	GRB2	BC000631
Plate1	C	10	75	gold_negative	2107	ETF1	BC014269	55716	LIMR	BC015015
Plate1	D	10	76	gold_negative	2167	FABP4	BC003672	2641	GCG	BC005278
Plate1	E	10	77	gold_negative	347	APOD	BC007402	4589	MUC7	BC025688
Plate1	F	10	78	gold_positive_fl	1104	CHC1	BC007300	5901	RAN	BC014901
Plate1	G	10	79	gold_positive_fl	2885	GRB2	BC000631	5747	PTK2	BC028733
Plate1	H	10	80	gold_positive_fl	4771	NF2	BC003112	9146	HGS	BC003565
Plate1	A	11	81	gold_negative	445	ASS	BC009243	64785	FLJ13912	BC005879
Plate1	B	11	82	gold_negative	410	ARSA	BC014210	1627	DBN1	BC007281
Plate1	C	11	83	gold_negative	293	SLC25A6	BC014775	7760	ZNF213	BC007287
Plate1	D	11	84	gold_positive_fl	1054	CEBPG	BC013128	2353	FOS	BC004490
Plate1	E	11	85	gold_negative	760	CA2	BC011949	5802	PTPRS	BC029496
Plate1	F	11	86	gold_positive_fl	5598	MAPK7	BC007404	5607	MAP2K5	BC008838
Plate1	G	11	87	gold_negative	1104	CHC1	BC007300	89857	KLHL6	BC032348
Plate1	H	11	88	gold_positive_fl	331	BIRC4	BC032729	836	CASP3	BC016926
Plate1	A	12	89	gold_negative	3718	JAK3	BC028068	51382	ATP6V1D	BC001411
Plate1	B	12	90	gold_positive_fl	6271	S100A1	BC014392	6285	S100B	BC001766
Plate1	C	12	91	gold_positive_fl	5894	RAF1	BC018119	5906	RAP1A	BC014086
Plate1	D	12	92	gold_negative	1058	CENPA	BC000881	53938	PPIL3	BC007693
Plate2	A	1	1	gold_positive_fl	3937	LCP2	BC016618	7409	VAV1	BC013361
Plate2	B	1	2	gold_negative	3119	HLA-DQB1	BC012106	220202	ATOH7	BC032621
Plate2	C	1	3	gold_negative	5371	PML	BC000080	23678	SGKL	BC015326
Plate2	D	1	4	gold_negative	4781	NFIB	BC001283	114609	TIRAP	BC032474
Plate2	E	1	5	gold_negative	6257	RXRB	BC001167	6373	CXCL11	BC005292
Plate2	F	1	6	gold_negative	4060	LUM	BC007038	55757	UGCGL2	BC032302
Plate2	G	1	7	gold_positive_fl	5781	PTPN11	BC008692	10818	FRS2	BC021562
Plate2	H	1	8	gold_positive_fl	1647	GADD45A	BC011757	5111	PCNA	BC000491
Plate2	A	2	9	gold_positive_fl	4000	LMNA	BC003162	4001	LMNB1	BC012295
Plate2	B	2	10	gold_positive_fl	2885	GRB2	BC000631	7409	VAV1	BC013361

PLATE	ROW	COL	POS	CLASS	GENE_ID_X	SYMBOL_BAIT	ACCESSION I	GENE_ID_Y	SYMBOL_PREY	ACCESSION II
Plate2	C	2	11	gold_negative	1315	COPB	BC037280	23516	SLC39A14	BC015770
Plate2	D	2	12	gold_negative	4660	PPP1R12B	BC034430	387032	ZNF307	BC014031
Plate2	E	2	13	gold_positive_fl	2175	FANCA	BC008979	2176	FANCC	BC015748
Plate2	F	2	14	gold_negative	3628	INPP1	BC015496	134510	MGC10067	BC013425
Plate2	G	2	15	gold_positive_fl	3065	HDAC1	BC000301	5925	RB1	BC039060
Plate2	H	2	16	gold_negative	3704	ITPA	BC010138	284403	C19orf14	BC017261
Plate2	A	3	17	gold_positive_fl	5159	PDGFRB	BC032224	5781	PTPN11	BC008692
Plate2	B	3	18	gold_negative	1748	DLX4	BC016145	117177	RAB3IP	BC015548
Plate2	C	3	19	gold_negative	389	RHOC	BC007245	54830	FLJ20130	BC016327
Plate2	D	3	20	gold_negative	821	CANX	BC003552	22899	ARHGEF15	BC036749
Plate2	E	3	21	gold_negative	1854	DUT	BC033645	91442	MGC32020	BC020247
Plate2	F	3	22	gold_positive_fl	8767	RIPK2	BC004553	10392	CARD4	BC040339
Plate2	G	3	23	gold_positive_fl	5195	PEX14	BC006327	5824	PEX19	BC000496
Plate2	H	3	24	gold_negative	4605	MYBL2	BC007585	55068	FLJ10094	BC024178
Plate2	A	4	25	gold_positive_fl	375	ARF1	BC011358	23647	ARFIP2	BC000392
Plate2	B	4	26	gold_positive_fl	4000	LMNA	BC003162	5925	RB1	BC039060
Plate2	C	4	27	gold_positive_fl	5824	PEX19	BC000496	8504	PEX3	BC015506
Plate2	D	4	28	gold_negative	10	NAT2	BC015878	3301	DNAJA1	BC008182
Plate2	E	4	29	gold_positive_fl	6118	RPA2	BC012157	6119	RPA3	BC005264
Plate2	F	4	30	gold_negative	1667	DEFA1	BC027917	158427	C9orf97	BC022958
Plate2	G	4	31	gold_negative	5718	PSMD12	BC019062	9419	CRIPT	BC006980
Plate2	H	4	32	gold_positive_fl	5979	RET	BC004257	10818	FRS2	BC021562
Plate2	A	5	33	gold_negative	949	SCARB1	BC022087	112885	PHF21B	BC012187
Plate2	B	5	34	gold_positive_fl	331	BIRC4	BC032729	842	CASP9	BC006463
Plate2	C	5	35	gold_positive_fl	387	RHOA	BC005976	392	ARHGAP1	BC018118
Plate2	D	5	36	gold_negative	4336	MOBP	BC022471	64432	MRPS25	BC003590
Plate2	E	5	37	gold_positive_fl	392	ARHGAP1	BC018118	663	BNIP2	BC002461
Plate2	F	5	38	gold_negative	5711	PSMD5	BC014478	93426	C10orf94	BC034821
Plate2	G	5	39	gold_negative	1209	CLPTM1	BC004865	83746	L3MBTL2	BC017191
Plate2	H	5	40	gold_positive_fl	331	BIRC4	BC032729	840	CASP7	BC015799
Plate2	A	6	41	gold_negative	355	FAS	BC012479	27258	LSM3	BC007055
Plate2	B	6	42	gold_negative	5367	PMCH	BC018048	79608	RIC3	BC022455
Plate2	C	6	43	gold_positive_fl	5824	PEX19	BC000496	9409	PEX16	BC004356
Plate2	D	6	44	gold_positive_fl	7157	TP53	BC003596	7329	UBE2I	BC000427

PLATE	ROW	COL	POS	CLASS	GENE_ID_X	SYMBOL_BAIT	ACCESSION I	GENE_ID_Y	SYMBOL_PREY	ACCESSION II
Plate2	E	6	45	gold_negative	947	CD34	BC039146	90203	C20orf161	BC019823
Plate2	F	6	46	gold_positive_fl	207	AKT1	BC000479	5170	PDPK1	BC033494
Plate2	G	6	47	gold_positive_fl	567	B2M	BC032589	3105	HLA-A	BC008611
Plate2	H	6	48	gold_positive_fl	914	CD2	BC033583	965	CD58	BC005930
Plate2	A	7	49	gold_positive_fl	567	B2M	BC032589	3107	HLA-C	BC004489
Plate2	B	7	50	gold_negative	4693	NDP	BC029901	11163	NUDT4	BC012069
Plate2	C	7	51	gold_positive_fl	1398	CRK	BC008506	5159	PDGFRB	BC032224
Plate2	D	7	52	gold_negative	3006	HIST1H1C	BC002649	56654	NPDC1	BC004217
Plate2	E	7	53	gold_negative	3146	HMGB1	BC003378	7004	TEAD4	BC015497
Plate2	F	7	54	gold_negative	4600	MX2	BC035293	79023	NUP37	BC000861
Plate2	G	7	55	gold_negative	5991	RFX3	BC022191	57715	SEMA4G	BC020960
Plate2	H	7	56	gold_positive_fl	6500	SKP1A	BC009839	8945	BTRC	BC027994
Plate2	A	8	57	gold_negative	1482	NKX2-5	BC025711	55454	GALNACT-2	BC030268
Plate2	B	8	58	gold_negative	4907	NT5E	BC015940	257397	TAB3	BC032526
Plate2	C	8	59	gold_positive_fl	5879	RAC1	BC004247	23647	ARFIP2	BC000392
Plate2	D	8	60	gold_negative	653	BMP5	BC027958	79892	C10orf119	BC004183
Plate2	E	8	61	gold_negative	1264	CNN1	BC022015	27315	FRAG1	BC009930
Plate2	F	8	62	gold_positive_fl	2065	ERBB3	BC002706	3084	NRG1	BC007675
Plate2	G	8	63	gold_negative	1315	COPB	BC037280	51440	HPCAL4	BC030827
Plate2	H	8	64	gold_positive_fl	2885	GRB2	BC000631	27040	LAT	BC011563
Plate2	A	9	65	gold_negative	4171	MCM2	BC007670	91584	DKFZp434G0625	BC028744
Plate2	B	9	66	gold_negative	2841	GPR18	BC008569	92906	HNRPLL	BC017480
Plate2	C	9	67	gold_positive_fl	835	CASP2	BC002427	8738	CRADD	BC017042
Plate2	D	9	68	gold_negative	5202	PFDN2	BC012464	126070	ZNF440	BC035760
Plate2	E	9	69	gold_negative	4129	MAOB	BC022494	10664	CTCF	BC014267
Plate2	F	9	70	gold_negative	90	ACVR1	BC033867	55280	CWF19L1	BC008746
Plate2	G	9	71	gold_positive_fl	8743	TNFSF10	BC032722	8795	TNFRSF10B	BC001281
Plate2	H	9	72	gold_negative	318	NUDT2	BC004926	60672	FLJ12438	BC008068
Plate2	A	10	73	gold_positive_fl	1081	CGA	BC010957	93659	CGB5	BC006290
Plate2	B	10	74	gold_positive_fl	4088	SMAD3	BC000414	4089	SMAD4	BC002379
Plate2	C	10	75	gold_positive_fl	5144	PDE4D	BC036319	10399	GNB2L1	BC014788
Plate2	D	10	76	gold_negative	5935	RBM3	BC006825	10291	SF3A1	BC007684
Plate2	E	10	77	gold_positive_fl	4089	SMAD4	BC002379	55802	DCP1A	BC007439
Plate2	F	10	78	gold_positive_fl	5530	PPP3CA	BC025714	5534	PPP3R1	BC027913

PLATE	ROW	COL	POS	CLASS	GENE_ID_X	SYMBOL_BAIT	ACCESSION I	GENE_ID_Y	SYMBOL_PREY	ACCESSION II
Plate2	G	10	79	gold_positive_fl	1027	CDKN1B	BC001971	8900	CCNA1	BC036346
Plate2	H	10	80	gold_negative	2811	GP1BA	BC027955	5819	PVRL2	BC003091
Plate2	A	11	81	gold_positive_fl	4086	SMAD1	BC001878	4089	SMAD4	BC002379
Plate2	B	11	82	gold_negative	5089	PBX2	BC003111	50853	VILL	BC000243
Plate2	C	11	83	gold_positive_fl	2962	GTF2F1	BC013007	2963	GTF2F2	BC001771
Plate2	D	11	84	gold_positive_fl	4171	MCM2	BC007670	4174	MCM5	BC000142
Plate2	E	11	85	gold_positive_fl	1810	DR1	BC002809	10589	DRAP1	BC010025
Plate2	F	11	86	gold_negative	539	ATP5O	BC021233	29121	CLEC2D	BC019883
Plate2	G	11	87	gold_negative	705	BYSL	BC007340	22889	KIAA0907	BC027182
Plate2	H	11	88	gold_positive_fl	3040	HBA2	BC008572	3043	HBB	BC007075
Plate2	A	12	89	gold_positive_fl	4999	ORC2L	BC014834	5000	ORC4L	BC014847
Plate2	B	12	90	gold_negative	1152	CKB	BC010002	3050	HBZ	BC027892
Plate2	C	12	91	gold_negative	2065	ERBB3	BC002706	285237	MGC26717	BC024188
Plate2	D	12	92	gold_positive_fl	567	B2M	BC032589	3106	HLA-B	BC013187

Supplementary Table S8: Insert and layout informations on hsPRS-v1 and hsRRS-v1.
Each line contains a protein pair tested in bait-X/prey-Y and bait-Y/prey-X direction.