**Article**

**Network of interactions between ZIKA virus non-structural proteins and human host proteins**

**Volha A. Golubeva** 1*, Thales C. Nepomuceno 1,2*, Giuliana de Gregoriis 2, Rafael D. Mesquita 3, Xueli Li 1, Patricia P. Garcez 4, Guilherme Suarez-Kurtz 2, Victoria Izumi 5, John Koomen 6, Marcelo A. Carvalho 2,7#, Alvaro N.A. Monteiro 1*

1 Cancer Epidemiology Program, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA
2 Divisão de Pesquisa Clínica, Instituto Nacional de Câncer, Rio de Janeiro 20230-130, Brazil
3 Departamento de Bioquímica, Instituto de Química, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
4 Institute of Biomedical Science, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
5 Proteomics and Metabolomics Core, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA
6 Chemical Biology and Molecular Medicine Program, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA
7 Instituto Federal do Rio de Janeiro - IFRJ, Rio de Janeiro 20270-021, Brazil.
* Both authors contributed equally
# Correspondence: alvaro.monteiro@moffitt.org or marcelo.carvalho@ifrj.edu.br; Tel.: 813-7456321 (A.N.A.M.) or Tel.: +55 21 2566-7774. (M.A.C.);

**Abstract:** The Zika virus (ZIKV) is a mosquito-borne Flavivirus and can be transmitted through an infected mosquito bite or through human-to-human interaction by sexual activity, blood transfusion, breastfeeding or perinatal exposure. After the 2015-2016 outbreak in Brazil, a strong link between ZIKV infection and microcephaly emerged. ZIKV specifically targets human neural progenitor cells, suggesting that proteins encoded by ZIKV bind and inactivate host cell proteins leading to microcephaly. Here, we present a systematic annotation of interactions between human proteins and the seven non-structural ZIKV proteins corresponding to a Brazilian isolate. The interaction network was generated by combining tandem-affinity purification followed by mass spectrometry with yeast two-hybrid screens. We identified 150 human proteins, involved in distinct biological processes, as interactors to ZIKV non-structural proteins. Our interacting network is composed of proteins that have been previously associated with microcephaly in human genetic disorders and/or animal models. This study builds on previously published interacting networks of ZIKV and genes related to autosomal recessive primary microcephaly to generate a catalog of human cellular targets of ZIKV proteins implicated in processes related to microcephaly in humans. Collectively, this data can be used as a resource for future characterization of ZIKV infection biology and help create a basis for the discovery of drugs which may disrupt the interaction and reduce the health damage to the fetus.

**Keywords:** ZIKV, protein-protein interaction, non-structural viral proteins, network

1. **Introduction**

Zika virus (ZIKV) is a neurotropic arthropod-borne virus belonging to Flaviviridae family, along with other Flaviviruses capable of infecting central nervous system, such as West Nile Virus, St. Louis Encephalitis Virus, and Japanese Encephalitis Virus. It is commonly transmitted though the bite of an infected Aedes species mosquito (A. aegypti or A. albopictus). These mosquitoes become infected when they feed on the infected person. Importantly, besides the mosquito bites,
human-to-human modes of transmission have also been documented, including sexual activity, blood transfusions and mother to fetus [1]. Since its first confirmed human infection in the 1960s, there were three documented Zika virus (ZIKV) outbreaks worldwide. The first two occurred in Micronesia and French Polynesia in 2007 and 2013, respectively. The most recent one (2015-2016) started in the Northeastern region of Brazil and rapidly spread through South America, the Caribbean and Mexico. By July 2016, locally transmitted cases of Zika infection were first reported in the US (Florida). According to the World Health Organization (WHO), 73 different countries had reported ZIKV infections by February of 2016 [2,3]. According to CDC, there have been no recorded local transmissions of the Zika virus in the continental United States in 2018 and 2019. However, with the globally increasing rate of travelling and the historical ability of viruses to acquire genetically modified virulence, the search for effective methods of Zika prevention and control remains important.

ZIKV infections in adults have been associated with neurological conditions such as Guillain-Barré syndrome, acute flaccid paralysis, and meningoencephalitis [4-7]. The Brazilian outbreak was the first time that ZIKV infection (presented in pregnant women) was correlated to congenital microcephaly in newborns [8-10]. Both in vitro and in vivo models have demonstrated that ZIKV has a tropism toward human neural progenitor cells [11-13]. In these cells, ZIKV infection is followed by apoptosis, corroborating the hypothesis of ZIKV as the etiological agent of these neurological disorders [4,5,11-13]. Further, independent studies have shown that the microcephaly and neural development-associated phenotypes seem to be a distinct feature of the Asian lineage [12,14,15]. However, the precise molecular mechanism(s) underlying these ZIKV-related manifestations is not understood.

ZIKV is a Baltimore class IV arbovirus from the Flaviviridae family. The ZIKV genome encodes a polyprotein that is processed by both viral and host proteases into ten proteins. Three of them (the capsid, pre-membrane and envelope) are responsible for the structural organization of the virus. The other seven are non-structural (NS) proteins (NS1, NS2A NS2B, NS3, NS4A, NS4B, and NS5) responsible for regulatory function, viral replication and subvert host responses [16].

The identification of virus-host protein-protein interaction is essential to better understand viral pathogenesis and to identify cellular mechanisms that could be pharmacologically targeted [17]. To gain further insight into the ZIKV pathogenesis, we generated a virus-host protein-protein interaction network focused on the interactions mediated by the non-structural proteins encoded by the Brazilian ZIKV genotype. Here we present a network composed of proteins related to neuron projection development, microcephaly-associated disorders, and by protein complexes linked to replication and infection of other members of the Flaviviridae family. In addition, we integrate our dataset with previously published ZIKV protein interaction networks, high-throughput protein-protein interaction data, and the historical ability of viruses to acquire genetically modified virulence, the search for effective methods of Zika prevention and control remains important.

2. Materials and Methods

2.1. ZIKV NS proteins cDNA constructs

We generated the cDNA of seven individual NS proteins (NS1, NS2A, NS3, NS2B, NS4A, NS4B, and NS5) corresponding to a ZIKV Brazilian isolate from the state of Pernambuco (GenBank AMD16557.1) [21]. We used the virus isolate as template for PCR to generate cDNAs for NS1, NS2A, NS2B, NS4A, and NS4B. We generated the Brazilian genotype cDNAs for NS3 and NS5 using site-directed mutagenesis with the pLV_Zika_Flag_NS3 and pLV_Zika_NS5_Flag plasmids as template (Addgene # 79634 and # 79639, respectively). Primer sequences will be provided upon request. All sequences were confirmed by Sanger sequencing.

ZIKV cDNAs coding for NS proteins were cloned into pGBK7 (Clontech) in frame with the GAL4 DNA binding domain (DBD) for Y2H assays, and into pNTAP (Agilent) vector for the TAP-MS assays. The GST-tagged baits used for Y2H validations were generated by subcloning the
cDNA from the isolated pGADT7 (Clontech) plasmid into the pDEST27 vector using Gateway recombination cloning according to the manufacturer’s instructions (ThermoFisher).

To validate Y2H interactions, recovered Y2H plasmids containing prey cDNAs were amplified by PCR using primers containing attB sites. PCR products were cloned into pDONR221 for Gateway recombination cloning (Invitrogen) and subsequently into pDEST27, to produce an N-terminal GST fusion. Primer sequences will be provided upon request.

2.2. Y2H Library Screening

To identify direct human brain protein targets of ZIKV NS proteins, we used the MATCHMAKER Gold Y2H system (Clontech). Seven ZIKV viral proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) were transformed into the Saccharomyces cerevisiae strain Y2HGol (Clontech) alone or co-transformed together with empty pGADT7 vector and tested for auto-activation and toxicity (defined by low transformation efficiency, small colony phenotype, or inability to grow in liquid culture) as previously described by our group [22].

All bait proteins were expressed in Y2HGol and did not induce toxic effects on the yeast cell cycle or survival (Figure S3A-B). Y2HGol transformants expressing each bait were mated to Y187 strain expressing a pre-transformed human brain normalized cDNA library (Matchmaker® Gold Yeast Two-Hybrid System; Cat.no. 630489; Clontech) for 20 h. The mated cultures were then plated on quadruple dropout medium (SD -Trp/-Leu/-His/-Ade) and incubated for 8-12 days (NS5 was screened twice). For every screen, more than 1 X 10^6 transformants were screened (Table S1). Yeast mini-prep DNA was used to recover pGADT7 fusions from each positive clone (Clontech Yeast Plasmid Isolation Kit), amplified by KOD polymerase chain reaction (PCR) and Sanger sequenced using a T7 primer. Out of frame clones were discarded and in-frame clones were kept for further analysis (Table S2).

2.3. Validation of Y2H interactions

Protein-protein interactions identified in Y2H screens were validated by expression in Human Embryonic Kidney (HEK) 293FT cells and protein pull-downs. HEK293FT cells were co-transfected with pDEST27 containing prey fusions to GST, and pNTAP containing bait fusions to SBP and CBP.

Cells were collected after 24 h and lysed in 1% CHAPS lysis buffer [1% CHAPS, 150 mM NaCl, 10 mM HEPES (pH 7.4) with protease and phosphatase inhibitors]. Cellular lysates were subjected to affinity purification of the TAP-tagged NS constructs using streptavidin-conjugated agarose beads, which were washed four times with 1% CHAPS lysis buffer and then analyzed by Western blot using anti-GST (GE27-4577-01; Sigma Aldrich) and anti-CBP tag antibodies (GenScript; Cat.no. A00635).

2.4. Tandem Affinity Purification coupled to Mass Spectrometry

HEK293FT cells were transfected using the calcium phosphate method with the SBP-CBP-tagged NS or control (GFP) vectors (Figure 1A). HEK cells have been previously used as a model for characterizing host-pathogen protein-protein interactions (PPI) [19,23, 24].

About 1 x 10^8 cells were used for the purification of the protein complexes using the InterPlay TAP purification kit (Stratagene) as described previously [22].

A nanoflow ultra-high-performance liquid chromatograph (RSLC, Dionex) coupled to an electrospray bench top orbitrap mass spectrometer (Q-Exact plus, Thermo) was used for liquid chromatography-tandem mass spectrometry (LC-MS/MS) peptide sequencing experiments. Samples were loaded onto a pre-column and washed for 8 minutes with aqueous 2% acetonitrile and 0.04% trifluoroacetic acid. Trapped peptides were eluted onto an analytical column, (C18 PepMap100, Thermo) and separated using a 90-minute gradient delivered at 300 nl/min. Sixteen tandem mass spectra were collected in a data-dependent manner following each survey scan using a 15 second exclusion for previously sampled peptide peaks (QExact, Thermo).
2.5. Analysis of proteomics data

We used Scaffold (Version 4.8.5) to obtain the original samples report of all TAP-MS based peptide and protein identifications. Peptide identifications were retained if they satisfied a minimum of 95.0% threshold. Protein identifications were accepted if they met greater than 50.0% threshold with a minimum of 2 identified peptides.

To further analyze the original Scaffold mass spectrometry data, we used APOSTL, an integrative Galaxy pipeline for affinity proteomics data [25]. The following global cutoffs were applied to 7,996 interactions and generated a list of 88 high confidence interactions: SaintScore cutoff: 0.5; FoldChange cutoff: 0; NSAF Score cutoff: 0.0000025.

APOSTL also interrogates the CRAPome database (http://crapome.org/), which contains common contaminants in affinity purification–mass spectrometry data [26]. Seventeen proteins displayed a CRAPome score >90% and were candidates to be called non-specific interactors. However, two hits were plausible due to being previously implicated in microcephaly (RAB18 and NEDD1), two hits were found associated with ZIKV NS proteins in a previously published independent study (AHCYL1 and GET4) [19]. Moreover, only 2 out 17 displayed multiple interactions, suggesting that the other 15 proteins do not constitute non-specific interactors in our assay. Therefore, we decided to retain all proteins, but have indicated high CRAPome score in Figure 2 when appropriate.

2.6. Generation of the microcephaly-associated PIN

We generated a microcephaly-associated PIN by searching NCBI’s ENTREZ Gene using [microcephaly] AND [Homo sapiens] as a query. This exercise led to 277 genes which were manually curated to remove those without an OMIM designation (i.e. pseudogenes and partially characterized loci) with a final tally of 261 genes. These genes were used as input to BisoGenet [27], which adds edges between the input nodes, to generate a microcephaly-associated network with 370 interactions (Table S13). BisoGenet settings were ‘input nodes only’ (Method) and checking ‘protein-protein interactions’ only leaving ‘Protein DNA interaction’ and ‘microRNA silencing interaction’ unchecked. Significant interaction clusters were identified using ClusterONE (Version 1.0) [28] using the following settings: ‘minimum size’ = 5; ‘minimum density’ = 0.5; ‘edge weights’ = unweighted. GO enrichment of clusters was done using BINGO as a Cytoscape plug-in.

2.7. Network generation and GO analysis

Network graphics were generated with Cytoscape version 3.7.1 [30]. Each NS integrated dataset was analyzed using WebGestalt to determine enrichment of GO terms. For each bait set (all proteins that interact with each NS bait), the number of genes in the set that were scored for a term was obtained. The number was then divided by the number of genes in the GO database for that term to obtain an enrichment ratio. Enrichment ratios that had a 0.0 value were replaced by 0.3 (half of the lowest non-zero value, 0.6) in the complete set. Bait sets were clustered with Cluster 3.0 using Correlation (uncentered) metric of similarity with no filtering but log2-transformed to depict increase and decrease changes as numerically equal but with opposite sign. The clustering method chosen was complete linkage. It was visualized using Java TreeView v 1.1.6r4 [31].

2.8. Mitocheck analysis and clustering

The Mitocheck phenotype database (20,921 genes), which scores 14 mitosis-related phenotypes in a binary form (presence = 1; absence = 0) from RNA interference screens was downloaded from http://www.mitocheck.org/ as a tab-delimited file in which genes are represented in rows and phenotypes in columns.

The enrichment and clustering (for each bait set) were performed as described above. Further, we also deconvoluted the integrated dataset to genes that were positive to at least one of the 14
phenotypes (Table S9). These new data set were clustered with Cluster 3.0 using Correlation (uncentered) metric of similarity with no filtering but log2-transformed to depict increase and decrease changes as numerically equal but with opposite sign. The clustering method chosen was complete linkage. It was visualized using Java TreeView v 1.1.6r4 [31].

3. Results

3.1. Yeast-two hybrid screenings

To build the first pair-wise protein-protein interaction database of ZIKV NS proteins encoded by the Brazilian genotype, we performed stringent Yeast-Two Hybrid (Y2H) screenings using ZIKV NS proteins as baits to screen a normalized human brain cDNA library (Figure 1A and Table S1). The Y2H screenings generated a protein-protein interaction network (PIN) ranging from 1 (NS2A) to 56 (NS3) interactions, totaling 99 unique protein hits and 109 bait-prey interactions (Figure 1A-B and Table S2).

To validate our Y2H PIN, a subset of 38 bait-prey interactions (38.4% of the Y2H PIN) was tested for interaction in human HEK293FT cells. All NS coding sequences were cloned in an eukaryotic expression vector in frame with the streptavidin and calmodulin binding proteins (SBP and CBP, respectively). Candidate interactor cDNAs were expressed in frame with Glutathione-S-Transferase (GST). CBP pulldown assays were performed against GST-tagged preys in HEK293FT cells, and 76.3% of interactions were confirmed (Figure 1C).

3.2. Tandem affinity purification followed by mass spectrometry

To further characterize the NS-mediated protein interactions, we expressed all baits as fusions to SBP and CBP in HEK293FT cells (Figure 1A). We then performed tandem affinity purification coupled to mass spectrometry (TAP-MS) which resulted in a high-confidence PIN with interactions ranging from 8 (NS2B and NS5) to 27 (NS2A), totaling 62 unique protein hits and 89 bait-prey interactions (Figures 1A and 2A-B; Table S3).

3.3. Merged ZIKV PIN

We combined the final Y2H and TAP-MS networks to generate the Merged ZIKV Network containing 157 nodes (including the baits) with 189 interactions between NS proteins and human host proteins (Figure 3A) (Table S4). Only 3 out of 151 hits were common to both Y2H and MS-based networks (BCLAF1, AHCYL1, and COPB1) indicating a limited overlap between methods.

3.4. Gene ontology

Gene ontology (GO) enrichment analysis of the Merged ZIKV PIN identified a subset of proteins mainly involved in 13 non-redundant biological processes (Figure S1). Among the hits identified, 13 are members of the proteasome complex (11 unique to the 26S subunit) and five members of the chaperonin-containing TCP1 (CCT) complexes (8.7% and 3.3% of our PIN, respectively) (Figure S1 and Table S5). GO enrichment analysis of cellular components revealed an enrichment of peptidase complex, chaperone complex and ficolin-1-rich granules (Table S6).

Next, we applied unsupervised clustering of bait sets according to their GO enrichment ratios for biological processes and cellular components (Figure 3B-C and Table S7-8). Interestingly, protein bait sets were clustered into two major groups (NS1, NS2A and NS3 versus NS2B, NS4A, NS4B and NS5).
Figure 1. Yeast two-hybrid PIN. A. Schematic representation of the ZIKV virus (ZIKV) genome and Non-structural (NS) constructs used on our protein-protein interaction screenings. Grey boxes represent the GAL4 DNA binding domain (DBD) and the streptavidin binding protein (SBP)-calmodulin binding protein (CBP) tag used on the yeast two-hybrid (Y2H) and tandem affinity purification (TAP) assays, respectively. The number of hits identified by each assay and bait are summarized on the right. B. Network of the interactions identified by Y2H screens. The color legend is depicted on the upper left-hand corner. C. Streptavidin pulldown of TAP-tagged NS constructs from 293FT cells, followed by Western blotting with the indicated antibodies. The individual percentage of hits validated by bait is depicted on the right. Red dots indicate the expected band size. Red asterisk indicates expected band size for Streptavidin pulldown assay.
Figure 2. TAP-MS PIN. A. Protein interaction profile of TAP-MS screenings plotted based on Normalized Spectral Abundance Factor (NSAF) [32] and specificity based on fold change of spectral counts between TAP-tagged NS proteins and TAP-tagged GFP (negative control). Node sizes and colors are based on the spectral sum and CRAPome PCT score, respectively. B. Network of the interactions identified by TAP-MS screens. The color legend is depicted on the upper left-hand corner.
Figure 3. Merged ZIKV PIN. A. Network of the interactions identified by Y2H and TAP-MS screens. The color legend is depicted on the upper left-hand corner. B-C. Clustering of bait sets according to GO enrichment ratio for biological processes (B) and cellular component (C). D. Phenoclusters (clustering of bait sets according to enrichment or depletion of Mitocheck phenotype classes. Clustering and visualization were performed using Cluster v3.0 software and TreeView v1.1.6r4, respectively.

3.5. Phenoclusters

Autosomal recessive primary microcephaly (MCPH) development is intrinsically associated with impaired mitosis [33]. Therefore, we used data from the Mitocheck project database (http://www.mitocheck.org/) [34,35] to determine the enrichment (or depletion) ratio of our bait sets for each mitotic phenotype scored in Mitocheck (Table S9). We then used the enrichment ratios to cluster bait sets according to their functional similarities (Figure 3D). Bait sets clustered around two large components according to their involvement in mitotic processes. One cluster (NS1, NS2A, and NS3 bait sets) presented enrichment of mitotic phenotypes while the second (NS2B, NS4A, NS4B, and NS5) did not, suggesting that NS1, NS2A, and NS3 are more likely to disrupt cellular mitotic processes (Figure 3D).

Finally, to identify individual preys more likely to be involved in mitotic processes we clustered all preys according to their Mitocheck enrichment rations (Figure S2) and identified a cluster of 9 proteins (CEP192, FAM184A, PAPSS1, EFTUD2, ZNF155, BAG6, SELENOP, KIF4A and PHPT1) with phenotypes consistent with centrosomal abnormalities (Table S10). This analysis
reflected the clustering pattern for NS1, NS2A and NS3 bait sets obtained when clustering for GO biological processes and cellular components (Figure 3B-D).

3.6. Integration with other ZIKV PINs

Our work builds on three previous physical interaction networks of host and ZIKV proteins [18-20]. We integrated our PIN with the published networks to evaluate the level of overlap between the four PINs (Figure 4A; Table S11). No common hit was shared by all four PINs and pair-wise overlaps ranged from 1 to 50 hits suggesting that ZIKV-host protein interacting network is still far from reaching saturation and that each network represent a distinct set of interactions with the integration of all datasets being important for the accurate description of the ZIKV-host interaction network (Figure 4A-B).

We identified five highly internally connected clusters among the integrated PIN (Figure 4C). All five clusters contained components of the ZIKV PIN from this study. Four of them were enriched in proteins involved in: a) Anaphase promoting complex (APC)-dependent proteasomal ubiquitin-dependent protein catabolic process (GO31145); b) Protein amino acid N-linked glycosylation via asparagine (GO18279); c) Protein folding (GO6457); d) Regulation of transcription (GO45449); and e) Histone H2B ubiquitination (GO33523).

Finally, 14 unique nodes of our merged ZIKV PIN (9.3% of the data set) have been shown to be important for proper replication of different Flavivirus (Table S12) [36-38], suggesting that our network also contains proteins that could explain the mechanisms of ZIKV replication and help identify therapeutic targets.

Figure 4. Integration of ZIKV PINs. A. Network of the interactions identified by this study, Scaturro et al. 2018, Shah et al. 2018, and Coyaud et al. 2018. The color legend is depicted on the upper left-hand corner. B. Clustering of bait sets according to overlapping protein complexes among the integrated network using ClusterONE (Version 1.0) [28]. Gene ontology of these networks were obtained by the Cytoscape plugin, BINGO [29]. GO accession numbers represent the following biological process: GO31145 - Anaphase promoting complex (APC)-dependent proteasomal ubiquitin-dependent protein catabolic process, GO18279 - Protein amino acid N-linked glycosylation via asparagine, GO6457 - Protein folding, GO45449 - Regulation of transcription, and GO33523 - Histone H2B ubiquitination. Proteins identified in this study are represented as orange
3.7. Integration with a microcephaly-associated network

We generated a new network composed of microcephaly-associated genes/proteins (Table S13) and our merged ZIKV PIN using BisoGenet to impute known interactions between nodes in this network (see Experimental Procedures) (Figure 5A). Four highly cohesive (i.e. highly connected internally, but only sparsely with the rest of the network) clusters emerged (Figure 5B). Of those, only three contained components of the ZIKV PIN: Anaphase promoting complex (APC)-dependent proteasomal ubiquitin-dependent protein catabolic process (GO31145), Centrosome duplication (GO7099), and COPI coating of Golgi vesicle (GO48205). Finally, we identified four nodes common to both PINs (CEP192, ASXL1, VARS, and EFTUD2). A similarly limited overlap between the microcephaly-associated and merged ZIKV PIN was also obtained with the three other previously determined ZIKV PINs (Figure 5C).

Figure 5. Integration of Microcephaly-associated PIN with the merged ZIKV PIN. A. Network of the interactions identified by this study, and proteins related to the Microcephalic phenotype (see experimental procedures) B. Clustering of bait sets according to overlapping protein complexes among the integrated network using ClusterONE (Version 1.0) [28]. Gene ontology of these networks were obtained by the Cytoscape plugin, BINGO [29]. Proteins identified in this study are represented as orange nodes. C. Venn diagrams represent the overlap between the Microcephaly-associated PIN and the individual ZIKV PINs.
4. Discussion

In humans, ZIKV infection was correlated with congenital microcephaly in newborns and with other neurological conditions in adults [4-7,11-13]. Still, little is known about the molecular mechanism of ZIKV infection and how it relates to neurological disorders. Here, we present a human host protein-ZIKV (Brazilian genotype) NS protein interaction network. This network was obtained by a combination of Yeast two-hybrid (Y2H) screens and tandem affinity purification coupled to mass spectrometry (TAP-MS). Y2H screens primarily reveal direct pair-wise interactions and are capable of detecting transient interactions, while TAP-MS will reveal proteins engaged in stable complexes, which may lead to the identification of indirect interactions [39-42]. The use of both methods results in a more comprehensive landscape of ZIKV protein-protein interactions.

The merged network combining two complementary methods (Y2H and TAP-MS) contains 157 nodes and 189 interactions with a limited overlap between the two methods, consistent with other previously determined PIN [22,41,42]. Further, the subset of Y2H interactions validated in human cells displayed a false positive rate of ~24% in line with other published Y2H screens [22,43-46]. These results suggest that this PIN contains high-confidence interactions.

Twenty-nine human proteins interacted with more than one ZIKV protein (19.3% of all hits). Similar relatively high levels of promiscuity of human proteins in relation to their viral interactors were also found in previous studies. Scaturro et al. [18], and Coyaud et al. [19] had 10.5% and 36% of all hits interacting to more than one ZIKV protein. Although Shah et al. [20] had a much lower (0.3% of all hits) rate of human proteins interacting to more than one ZIKV protein, a high level of promiscuity of human proteins is also apparent across studies, where the same human proteins are often found interacting with distinct ZIKV proteins. For example, all human proteins shared between Shah et al. [20] and Scaturro et al. [18], or between Shah et al. and this study, were found to interact with different ZIKV NS proteins. In addition, comparisons across other studies showed consistently high levels of discordance in bait interactions (Figure 4B). These data suggest that different ZIKV NS proteins have common targets in the human proteome. However, it is unclear why different studies detected exclusive interactions with different baits. It is conceivable that several ZIKV NS proteins interact with large protein complexes, such as the 26S subunit of the proteasome complex, via different targets; furthermore, differences in the biology of the cells providing the proteome (i.e. levels of protein expression and formation of specific protein complexes), the biochemical methods, or the filtering criteria for significant interactions may also determine which interactions are robust enough to result in detection. Alternatively, some could represent spurious interactions detected due to the overexpression of the baits; however, the limited number of proteins in our dataset with high CRAPome [47] scores indicating consistent recovery in affinity proteomics as non-specific background would suggest otherwise.

Further, we identified multiple components of CCT (Chaperonin Containing TCP1 or TriC-TCP-1 Ring Complex) complex as targets. This complex plays a role in trafficking of telomerase and small Cajal body (CB) RNAs through the proper folding of the telomerase cofactor, TCAB1 [48]. CBs are transcription-dependent nuclear compartments and play a critical role in neuron biology through snRNP and snoRNP assembly [49]. Interestingly, Coyaud et al. [19] demonstrated that ZIKV NS5 expression leads to an increase in the absolute number of CBs per cell, but to a reduction of the volume of these CBs, suggesting that NS5 expression could lead to CB fragmentation. Our data points to the interaction of NS1 with multiple components of the CCT complex, suggesting that NS1 could also play a role in CB stability and in neural disorders. Additionally, it has already been shown that the Dengue virus (DENV) infection occurs in an NS1/CCT-dependent manner [50].

Centrosomal abnormalities lead to impaired mitosis, which is a hallmark of MCPH. In fact, our data set present multiple proteins related to phenotypes associated with impaired mitosis (Figure 3B and 3C). Furthermore, our PIN shares 24 (16% of all unique hits) known interaction partners of 14 (out of 18) MCPH loci plus CEP63 (Table S14).

In that context, CEP192 (identified as an NS5 interaction partner by Y2H) plays a central role in the initial steps of centriole duplications through the interaction and recruitment of CEP152 (MCPH9) and PLK4, respectively. Which is necessary for the proper recruitment of SAS6 (MCPH14), STIL...
(MCPH7) and CENPJ (MCPH6) [51-56]. Our data suggest that NS3 could interfere with centriole duplication and consequently could be important for the ZIKV-associated microcephaly phenotype. Furthermore, GO enrichment analysis and Mitocheck phenoclusters suggest that NS1, NS2A and NS3 target host factors implicated in mitotic phenotypes.

In summary, the data presented here together with three previously published studies [18-20] provide a valuable resource to dissect the mechanistic underpinnings of central nervous system perturbations caused by ZIKV infection and to identify potential pharmacological targets.

**Supplementary Materials:** The following are available online, Figure S1: GO enrichment of the merged network, Figure S2: Phenoclusters of individual preys, Figure S3: Y2H bait expression, control transformations and matings, Table S1. Yeast two-hybrid screening data, Table S2. Yeast two-hybrid hits, Table S3. TAP-MS hits - APOSTL output, Table S4. Merged PIN (Y2H+TAP/MS), Table S5. GO (Cellular component) enrichment membership, Table S6. GO (Biological Process) enrichment membership, Table S7. Bait-specific GO (Biological Process) enrichment ratios, Table S8. Bait-specific GO (Cellular component) enrichment ratios, Table S9. Phenoclusters of bait sets, Table S10. Phenoclusters of individual preys, Table S11. Integrated ZIKV PIN, Table S12. Flavivirus replication factors (functional screens) intersection with Merged ZIKV PIN, Table S13. Microcephaly-associated genes, Table S14. Merged ZIKV PIN and MCPH subnetwork.

**Author Contributions:** Conceptualization, Marcelo Carvalho and Alvaro Monteiro; Data curation, Volha Golubeva, Thales Nepomuceno, Rafael Mesquita, Marcelo Carvalho and Alvaro Monteiro; Formal analysis, Volha Golubeva, Thales Nepomuceno and Alvaro Monteiro; Funding acquisition, Marcelo Carvalho and Alvaro Monteiro; Investigation, Volha Golubeva, Thales Nepomuceno, Giuliana de Gregoriis, Rafael Mesquita, Xueli Li, Patricia Garcez, Victoria Izumi, John Koomen and Alvaro Monteiro; Methodology, Volha Golubeva, Thales Nepomuceno, Xueli Li and Victoria Izumi; Project administration, Alvaro Monteiro; Resources, Patricia Garcez, Guilherme Suarez-Kurtz, Marcelo Carvalho and Alvaro Monteiro; Supervision, Marcelo Carvalho and Alvaro Monteiro; Validation, Volha Golubeva, Thales Nepomuceno, Giuliana de Gregoriis and Xueli Li; Visualization, Thales Nepomuceno and Alvaro Monteiro; Writing – original draft, Volha Golubeva, Thales Nepomuceno, Marcelo Carvalho and Alvaro Monteiro; Writing – review & editing, Volha Golubeva, Thales Nepomuceno, Rafael Mesquita, Patricia Garcez, Guilherme Suarez-Kurtz, John Koomen, Marcelo Carvalho and Alvaro Monteiro.

**Funding:** This work was supported by Florida Department of Health Zika Research Grant Initiative pilot award 7ZK29 and in part by the Proteomics and Metabolomics Core at the Moffitt Cancer Center through its NCI CCSG grant (P30-CA76292), and by Moffitt’s Center for Immunization and Infection Research in Cancer (CIIRC). T.C.N. is a Fulbright Scholar.

**Acknowledgments:** We thank the individuals and families who have generously donated their time, samples, and information to facilitate research on ZIKV. We also thank Brent Kuenzi for help with proteomic data analysis.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest related to the contents of this article.
References


