Proteomic analysis of virus-host interactions in an infectious context using recombinant viruses

Komarova Anastassia V.¹, Chantal Combredet¹, Laurène Meyniel-Schicklin²,³,⁴, Manuel Chapelle⁵, Grégory Caignard¹, Jean-Michel Camadro⁵, Vincent Lotteau²,³,⁴, Pierre-Olivier Vidalain¹*, Frédéric Tangy¹*

From the ¹Unité de Génomique Virale et Vaccination, Institut Pasteur, CNRS URA 3015, Paris, France. ²INSERM U851, Lyon, France. ³IFR128-BioSciences, Université Lyon 1, France. ⁴Université de Lyon, France. ⁵Plateforme de Protéomique Structurale et Fonctionnelle, Institut Jacques Monod, Université Paris-Diderot, CNRS UMR 7592, Paris, France.

* To whom correspondence should be addressed: Pierre-Olivier Vidalain or Frédéric Tangy. E-mail: vidalain@pasteur.fr or ftangy@pasteur.fr. Phone: +33 (1) 45 68 87 70; Fax: +33 (1) 40 61 31 67

RUNNING TITLE: Analysis of V protein complexes within measles virus infected cells
ABBREVIATIONS

AA amino acid
CH Red fluorescent protein Cherry
CHD4 Chromodomain-helicase-DNA-binding protein 4
GATAD2A Transcriptional repressor p66-alpha
HCV Hepatitis C virus
HDAC Histone deacetylase complex
IFIH1 Interferon-induced helicase C domain-containing protein 1
IFN Interferon
IKK-α Inhibitor of nuclear factor kappa-B kinase subunit alpha
IRF7 interferon regulatory factor 7
JAK Tyrosine-protein kinase
LGP2 Probable ATP-dependent RNA helicase DHX58
Luc Firefly luciferase protein
MBD3 Methyl-CpG-binding domain protein 3
MOI multiplicity of infection
MTA Metastasis-associated protein
MV Measles virus
N nucleoprotein
RBBP7 Histone-binding protein
rMV2/Luc recombinant MV with Luc inserted in an additional transcription unit between P and M ORFs
rMV2/STrEP-CH recombinant measles virus with Cherry protein possessing N-terminal STrEP-tag inserted in an additional transcription unit between P and M ORFs
rMV2/CH-STrEP recombinant measles virus with Cherry protein possessing C-terminal STreP-tag inserted in an additional transcription unit between P and M ORFs

rMV2/STrEP-V recombinant measles virus with V protein possessing N-terminal STreP-tag inserted in an additional transcription unit between P and M ORFs

rMV2/N-STrEP recombinant measles virus with N protein possessing C-terminal STreP-tag inserted in an additional transcription unit between P and M ORFs

RNP ribonucleoprotein

STAT Signal transducers and activators of transcription protein

TEV Tobacco Etch Virus

TSA Trichostatin A
RNA viruses exhibit small-sized genomes encoding few proteins, but still establish complex networks of interactions with host cell components to achieve replication and spreading. Ideally, these virus-host protein interactions should be mapped directly in infected cell culture, but such a high standard is often difficult to reach when using conventional approaches. We thus developed a new strategy based on recombinant viruses expressing tagged viral proteins to capture both direct and indirect physical binding partners during infection. As a proof of concept, we engineered a recombinant measles virus (MV) expressing one of its virulence factor, the MV-V protein, with a One-STrEP amino-terminal tag. This allowed virus-host protein complex analysis directly from infected cells by combining modified tandem affinity chromatography and mass spectrometry analysis. Using this approach, we established a prosperous list of 245 cellular proteins interacting either directly or indirectly with MV-V, and including 4 of the 9 already known partners of this viral factor. These interactions were highly specific of MV-V since they were not recovered when the nucleoprotein MV-N, instead of MV-V, was tagged. Besides key components of the antiviral response, cellular proteins from mitochondria, ribosomes, endoplasmic reticulum, protein phosphatase 2A and histone deacetylase complex were identified for the first time as prominent targets of MV-V and the critical role of the later protein family in MV replication was addressed. Most interestingly, MV-V showed some preferential attachment to essential proteins in the human interactome network, as assessed by centrality and interconnectivity measures. Furthermore, the list of MV-V interactors also showed a massive enrichment for well-known targets of other viruses. Altogether, this clearly supports our approach based on reverse
genetics of viruses combined with high-throughput proteomics to probe the interaction network that viruses establish in infected cells.
INTRODUCTION

RNA viruses are responsible for numerous human diseases like flu, AIDS, hepatitis C (HCV), dengue, measles, yellow fever, and others that still represent major public health threats. Despite small genomes encoding only few viral proteins, RNA viruses establish a complex network of interactions with host cell components to block cellular defense mechanisms and hijack host cell machinery (for review see (1)). Deciphering these interactions is essential to reach a comprehensive understanding of the viral infection process. To obtain this information at a system level, high-throughput technologies like yeast two-hybrid or MS-based protein-complex analysis are currently used. However, these strategies have significant limitations.

In the yeast two-hybrid system, viral proteins are used as baits to screen host cDNA or ORFeome libraries. This technology was used to establish a first draft of HCV and influenza virus infection networks, and allowed the identification of several cellular pathways as major targets of these viruses (2, 3). Although highly tractable, the yeast two-hybrid technique can be criticized since an artificial interaction assay is performed in a heterologous system. For MS-based proteomics, single tagged viral proteins or viral protein complexes are expressed in host cells, then isolated together with interacting cellular factors using one-step or tandem affinity purification protocols. For instance, the ribonucleoprotein and polymerase complexes of influenza A virus have been rebuilt in human cells by transient co-expression of tagged viral proteins, and then purified to identify cellular binding partners (4). Although this approach is somewhat more relevant in its design than the yeast two-hybrid system, it can be argued that virus-host interactions are detected in non-infected cells. Alternatively, virions or virus ribonucleoproteins (RNPs or
ribonucleocapsid) can be purified from infected cell cultures or tissues using gel exclusion chromatography or appropriate density gradients, and then analyzed for co-purified cellular proteins by MS (5, 6, 7). However, this approach is only applicable to viral proteins that assemble into high molecular weight multi-protein complexes. Thus, innovative strategies are needed to detect virus-host protein interactions in infected cell cultures.

The study of RNA viruses has greatly benefited from the ability to engineer viral genomes and generate modified viruses by reverse genetics (for review see (8-10). Interestingly, this technology may be used to generate recombinant viruses carrying amino acid (AA) tags in fusion with one or several of their proteins, provided that the tags do not interfere with viral protein functions. Then, tagged viral proteins can be purified directly from infected cells and co-purified cellular interactors identified by MS-based analysis. With such tagged viruses in hands, virus-host protein interactions could be identified all along a virus life cycle, within different cell types and under various culture conditions. As a proof-of-concept, we developed and describe in this report such technology for measles virus (MV).

MV is a negative-strand RNA virus from Paramyxoviridae family that encodes six structural proteins and two non-structural virulence factors, among which MV-V has been extensively studied (11). Although MV-V is not essential to viral replication in vitro, this protein is required for MV propagation in vivo (12, 13). MV-V is encoded by the P locus, and is composed of an N-terminal PNT region and C-terminal VCT region. MV-V is best known for its ability to block type I interferon (IFN-α/β) pathway by different mechanisms, thus impairing antiviral immune response in MV-infected cells. Accounting for this activity, MV-V has been shown to interact directly with IFIH1
and LGP2 antiviral RNA helicases, JAK1 and IKK-α kinases, and IRF7, STAT1, STAT2, p53, and p73 transcription factors (14-19). Thus, MV-V interacts with multiple cellular components and moonlights between different functions that are essential to viral inhibition of host cell immunity. Consequently, MV-V represents a prototypical viral virulence factor that exerts multiple regulatory functions, thus suggesting numerous dynamic interactions with human proteome. We used a reverse genetic system developed in our laboratory (20) to produce a recombinant MV strain expressing the MV-V protein with a One-STrEP amino-terminal tag. This allowed MV-V purification directly from infected cells using a modified one-step purification approach. Co-purified cellular factors were identified by direct MS-based analysis. Already known MV-V cellular partners were recovered with this system but more importantly, a specific and prosperous list of new virus-host interactions was found. To our knowledge, this is the first report describing a combination of reverse genetic with proteomic analysis to map virus-host interactions directly from virus-infected cells. In the future, the same method could be applied to any virus with a negative-strand RNA genome.

EXPERIMENTAL PROCEDURES

Cells, plasmid constructions, and rescue of recombinant viruses. HEK-293T (human embryonic kidney) and Vero cells (African green monkey kidney cells) were maintained in DMEM-Glutamax (Gibco-BRL) supplemented with 10% heat-inactivated foetal calf serum (FCS, Invitrogen, Frederick, MD). Stable helper 293-T7-NP cells were used for viral rescue as previously described (21), and grown in DMEM 10% FCS. pTM-MVSchw plasmid, which contains an infectious MV cDNA
corresponding to the anti-genome of the Schwarz vaccine strain, has been previously described (20, 21).

A two-step PCR-based strategy was used to produce coding sequences for MV-V, MV-N, and the Red fluorescent protein Cherry (CH) with either N- or C-terminal One-STrEP-tags. First, the One-STrEP-tag coding sequence was amplified by PCR from pEXPR-IBA105 using different primer pairs to add a flexible Tobacco Etch Virus (TEV) protease linker (ENLYFQS) either at the N- or C-terminus of this sequence. In parallel, DNA fragments corresponding to MV-V, MV-N, or CH coding sequences were amplified using pTM-MVSchw or pmCherry (Clontech) as a template. PCR products were finally combined, and then amplified in a second PCR reaction to recover expected fusion sequences, i.e. MV-V and CH with a N-terminal One-STrEP tag, and MV-N and CH with a C-terminal One-STrEP tag (supplementary Fig. 1). These four amplicons contained unique BsiWI and BssHII sites at their extremities for subsequent cloning in pTM2-MVSchw, but were first cloned in pCR2.1-TOPO plasmid (Invitrogen) and sequenced. These plasmids were designated pTOPO/STrEP-V, pTOPO/STrEP-CH, pTOPO/N-STrEP and pTOPO/CH-STrEP. Finally, sequences were introduced in the additional transcription unit of pTM2-MVSchw vector after BsiWI/BssHII digestion. The resulting plasmids were designated pTM2-MV/STrEP-V, pTM2-MV/STrEP-CH, pTM2-MV/N-STrEP and pTM2-MV/CH-STrEP. The coding sequence of firefly luciferase (Luc) from Photinus pyralis was introduced in pTM2-MVSchw using the same approach. The DNA fragment encoding for Luc was amplified by PCR using pISRE-Luc (Stratagene) as a template. This DNA fragment contained unique MluI and BssHII restriction sites at its extremities for subsequent cloning in pTM2-MVSchw. The resulting plasmid was designated pTM2-MV/Luc. All MV insertions respect the “rule of six”, which stipulates
that the number of nucleotides of MV genome must be a multiple of six (22). Recombinant viruses were rescued, and virus titers and single-step growth curves were determined as previously described (21).

A simple PCR amplification strategy was used to subclone STrEP-V, STrEP-PNT, CH-STrEP coding sequences in a mammalian expression vector. STrEP-V and STrEP-PNT were amplified from pTOPO/STrEP-V, whereas CH-STrEP was amplified from pTOPO/CH-STrEP. All three DNA fragments contained unique EcoRI and Sall restriction sites at their extremities for subsequent cloning in pCI-neo (Promega). The resulting plasmids were designated as pCI-neo/STrEP-V, pCI-neo/STrEP-PNT, pCI-neo/CH-STrEP.

Antibodies and western blots assays. Vero cells were infected in 35 mm dishes with MV recombinant viruses at a multiplicity of infection (MOI) of 1. At 24 h post-infection, cells were lysed in 500 µl of lysis buffer (20 mM MOPS-KOH pH7.4, 120 mM of KCl, 0.5 % Igepal, 2 mM β-Mercaptoethanol), supplemented with Complete Protease Inhibitor Cocktail (Roche). Cell lysates were incubated on ice for 20 min, then clarified by centrifugation at 16,000g for 15 min. Protein extracts were resolved by SDS-PAGE gel electrophoresis on 4-12% NuPAGE Bis-Tris gels with MOPS running buffer (Invitrogen) and transferred to cellulose membranes with the IBlot Dry blotting system (Invitrogen). To detect One-STrEP-tag and MV proteins, membranes were blotted with either a mouse anti-N mAb (clone 25; kindly provided by Pr. Chantal Rabourdin-Combe (23), or a rabbit polyclonal anti-V rAb kindly provided by Dr. Kaoru Takeuchi (24), or Streptavidin-HRP (Invitrogen, 19534-050). Peroxidase activity was visualized with an ECL Plus Western Blotting Detection System (GEHealthcare, RPN2132). STAT1 and STAT2 proteins were revealed using
clone-1 (Ref. G16920) and clone 22 (Ref. S21220), respectively (mouse monoclonal antibodies; BD Biosciences). Secondary anti-mouse HRP-conjugated antibodies were from GE-Healthcare.

*Modified One-STRep tag purification.* 8×10^7 HEK-293T cells were either infected with MV recombinant viruses at an MOI of 1 or transfected using jetPRIME (Polyplus transfection) with 40 µg of pCI-neo/STRep-V or pCI-neo/STRep-PNT or pCI-neo/CH-STRep. 24 h post-infection or post-transfection, cells were lysed in 4 ml of lysis buffer. Cell lysates were incubated on ice for 20 min, and then clarified by centrifugation at 16,000g for 15 min. Protein extracts were incubated for 2 h on a spinning wheel at 4°C with 300 µl of StrepTactin Sepharose High Performance (GE Healthcare, 28935599). Beads were then washed twice for 5 min on a spinning wheel with 10 ml of washing buffer (20 mM MOPS-KOH pH7.4, 120 mM of KCl, 2 mM β-Mercaptoethanol), supplemented with Complete Protease Inhibitor Cocktail (Roche). Protein complexes were eluted from StrepTactin beads with 0.04 mg/ml of TEV protease (kindly provided by Dr. Nicolas Wolff, Unité de Résonance Magnétique Nucléaire des Biomolécules, Institut Pasteur) in 2 ml of washing buffer over night on a spinning wheel at 4°C. Beads were centrifuged at 1,600g for 5 min at 4°C to recover eluted protein complexes. Beads were washed ones with 2 ml of washing buffer to recover a maximum of protein complexes. Finally, eluted proteins were precipitated overnight at 4°C with TCA (12% final concentration). Protein pellets were washed twice with ice-cold acetone, and resuspended in deionized water (Mili-Q). Protein concentration was determined using the Amidoschwarz staining reaction (25).

*SDS-PAGE analysis of affinity purified protein complexes.* XT Sample buffer
(Bio Rad, 161-0791) and XT Reducing Agent (BioRad, 161-0792) were added to 10 µg of protein. After denaturation for 5 min at 100°C, protein samples were loaded on a precast 4-12% acrylamide gel (BioRad). Gels were stained with Proteo Silver Plus Silver Stain Kit (Sigma, Prot-SIL2).

Identification of virus-host protein partners by mass spectrometry. Equal protein concentrations from 0.1 µg to 1 µg were used in each nano-LC-MS/MS experiment. Proteins were digested overnight at 37°C with 20 µl of a trypsin solution prepared as follow: HPLC water, 10 ng/µl trypsin sequencing grade (Roche), 10% HPLC grade ACN, 25 mM ammonium carbonate (Sigma). Resulting peptide extracts were speed vac dried and dissolved in 12 µl HPLC grade water, 0.1% formic acid. 5 µl were used for nano-LC-MS/MS analysis using a nano-chromatography system (Easy nLC, Proxeon) connected online to a LTQ Velos Orbitrap (ThermoFisher) mass spectrometer. A 2 cm long, 5 µm particle size C18 Easy column (Proxeon) was used for peptide trapping and dessalting. A 10 cm long and 3 µm particle size C18 Easy column (Proxeon) was used for peptides separation. The peptide elution gradient was from 100% buffer A (HPLC grade water, 0.1% formic acid) to 35% buffer B (HPLC grade ACN, 0.1% formic acid) in 60 min using a constant flow of 300 nl/min. MS spectra were acquired on the Orbitrap analyzer at resolution mode R=30000. After each MS spectrum, an automatic selection of 20 most intense precursor ions was activated with 15 sec dynamic exclusion delay to acquire MS/MS spectra on the LTQ Velos analyser, using CID fragmentation mode at 35% relative resonant activation energy for 40 msec.

Raw data were pre-processed using ProteomeDiscoverer version 1.2 (ThermoFisher). Pre-processing consisted of MS/MS spectrum averaging of spectra
from equal precursor molecular weight, with 5ppm mass tolerance and 90s time window. No thresholding was applied neither to MS ions intensities nor MS/MS fragment ions intensities.

All spectra were analyzed twice using Mascot version 2.1 (Matrix Science). A first query was performed against a non-redundant database of 20,352 human protein sequences from Swissprot (rel. 57.15) to which corresponding random decoy entries were added. Then, second search was performed against a database made of the 90 MV protein sequences found in the SwissProt database. Each sequence was randomized 11 times and added as decoy to this database, thus resulting in a 1,080 entries database suitable for robust and consistent calculation. The sequence of the red fluorescent protein Cherry was also included. Mascot was run in MS/MS ion search mode with the following parameter settings: no fixed modification, variable modification (oxydation on methionine), precursor mass tolerance 5 ppm, fragment ions mass tolerance 0.4 Da, 2 missed cleavages and trypsin as digestion enzyme. Additional filtering was applied after protein and peptide identification for further analysis, using following criteria: for single and multiple peptide hit proteins each Mascot ionscore $\geq 30$. Such criteria allowed us to obtain a false-positive rate below 1% for each search, based on numbering of identified decoy entries. For proteins identified on the basis of one unique peptide spectrum see supplementary Fig. 2.

Protein interaction network analysis. Interactions between cellular proteins that co-purified with MV-V were retrieved from APID (“Agile Protein Interaction DataAnalyzer”) using APID2NET plug-in integrated in Cytoscape software, an advanced network analysis tool (26, 27). The obtained interaction network was subsequently analyzed with BINGO2 plug-in to determine statistical enrichments for
Gene Ontology (GO) categories (28). Finally, protein nodes were coloured using the GOLorize plug-in to emphasize enrichments for specific GO terms (29).

For network topological analyses, and because APID dataset is not accessible as a whole, a human interactome network was reconstructed from IntAct database (December 2010, #19850723). Intact database is part of APID dataset, and represents 25.7% of interaction data retrieved by APID2NET. Virus-host protein interaction data were from VirHostNet database (December 2010, #18984613). The R (http://www.r-project.org) statistical environment was used to perform statistical analysis, and the igraph R package (http://cneurocvs.rmki.kfki.hu/igraph/) was used to compute network metrics. In our topological analyses, the degree $k$ of a node in a graph is defined as the number of edges that are incident to this node. It is a local centrality measure that takes into account only direct 1-hop neighbours, i.e. direct interacting partners. The betweenness $b$ of a node in a graph can be defined by the number of shortest paths going through this node, and is normalized by twice the total number of proteins pairs in the graph. The equation used to compute the betweenness centrality $b$ for a node $v$ is:

$$b(v) = \frac{1}{n \times (n-1)} \times \sum_{i,j,v \in V} \frac{g_{ij}(v)}{g_{ij}}$$

where $g_{ij}$ is the number of shortest paths going from node $i$ to $j$, $i$ and $j \in V$ and $g_{ij}(v)$ the number of shortest paths from $i$ to $j$ that pass through the node $v$. 

14
Statistical significance for the interconnectivity of MV-V interactors was assessed by a random resampling testing procedure (n = 10,000 permutations). For each permutation, we randomly extracted a list of proteins equal in size to the MV-V interactor list from the human interactome, and the number of shared interactions was assessed. The randomization procedure was weighted and corrected according to the connectivity of proteins in order to prevent inspection bias on highly studied proteins. A theoretical distribution was computed for the 10,000 resampled values. From this distribution, an empirical p-value was computed by counting the number of resampled values greater than the value observed for MV-V interactors.

Statistical significance for the number of proteins targeted by at least one viral protein was evaluated by comparison with the distribution obtained by random simulation using R. We randomly drawn as many proteins as those identified as MV-V interactors from the human proteome, and we numbered the number of proteins targeted by at least one viral protein. This procedure was repeated 10,000 times. From this distribution, an empirical p-value was computed by counting the number of resampled values greater than the value observed for MV-V interactors.

siRNA-mediated Histone deacetylase complex down-regulation. Synthetic double-stranded siRNA were purchased from Applied Biosystems (Catalog numbers: s17391, s17629, s11847, s2984, s28737, s224320, s223571, s10457) to knock-down the expression of MTA1, MTA2, RBBP7, CHD4, MBD3, GATAD2A, ZBTB43 (CT1) and PDK2 (CT2), respectively. A standard protocol of forward siRNA transfection was used. One day before transfection HEK-293T cells were plated in either 96-well white microplates (for subsequent luciferase and cell viability assays) or in 96-well transparent microplates (for Imaging). $2 \times 10^4$ cells per well in 0,08 ml DMEM
medium with 10% fetal bovine serum, but without antibiotics were seeded. Transfections with siRNAs were carried out in triplicates next day using Lipofectamine RNAi-MAX (Invitrogen) according to the manufacturers' protocol. For each well 6 pmol siRNA and 0.25 µl RNAi-MAX transfection reagent combined with 20µl of Opti-MEM serum-free medium (Invitrogen) were added. 24 h post-transfection cells were infected with recombinant MVs expressing Luc or CH proteins (rMV2/Luc or rMV2/STrEP-CH) at an MOI of 0.5, or mock-treated to determine cell viability. For recombinant MV expressing Luc, luciferase activity was determined 48h post-infection using the Bright-Glo Luciferase assay system (Promega). Cell viability was determined at the same time point using the CellTiter-Glo Luminescent Cell Viability Assay (Promega). Luminescence was measured using a 2300 EnSpire Multilable Plate reader (PerkinElmer). Fluorescent and light microscopies were performed 48h post-infection with rMV2/STrEP-CH to determine CH protein expression levels and cell-to-cell fusion events in siRNA-transfected HEK-293T cells.

*Inhibition of Histone deacetylase activity with Trichostatin A (TSA).* 4 × 10⁴ cells per well were plated in either 96-well white microplates (for subsequent luciferase and cell viability assays) or in 96-well transparent microplates (for Imaging). Next day, cells were treated or not with TSA (Sigma, dose range from 1 to 0.06 µM), and infected with recombinant rMV2/Luc or rMV2/STrEP-CH at an MOI of 0.2 or mock-treated to determine cell viability. These experiments were carried out in duplicates. Luciferase assays were performed as described above 48h post-infection or 48h after TSA treatment. Fluorescent and light microscopies were performed 72 h post-infection with rMV2/STrEP-CH to determine CH protein expression level in HEK-293T cells in the presence or absence of TSA.
**Fluorescence microscopy imaging.** Fluorescence corresponding to the mCherry reporter (excitation/emission: 587/610 nm) was observed using the DM IRB inverted research microscope (Leica) supplied with the green filter N2.1 (Leica; BP 515-560 nm) and imaging was performed using the QImaging Retiga EXi Fast 1394 CCD camera.

**RESULTS**

*Generation of recombinant MV expressing a tagged MV-V.* To co-purify MV-V with associated protein complexes from infected cells, MV-V has to be expressed from MV genome with a fusion tag allowing its capture. To reach this goal, we took advantage of a reverse genetic system for MV based on the Schwarz vaccine strain that has been developed in our lab (20). A sequence encoding for MV-V protein with an amino-terminal One-StrEP tag was inserted into the pTM-Schw plasmid containing MV genome, in an additional transcription unit located downstream of the P gene (Fig. 1). According to previous studies, known MV-V functions are not affected by the addition of an amino-terminal tag (14, 15, 17, 30). The engineered recombinant virus was designated rMV2/StrEP-V. As a negative control, we also produced a recombinant virus expressing the CH protein instead of MV-V (designated rMV2/StrEP-CH). In each construct, a flexible TEV protease coding linker was inserted between the One-StrEP tag and MV-V or CH proteins to provide a strategy for releasing protein complexes after purification. Recombinant viruses encoding One-StrEP-tagged MV-V or CH proteins were successfully rescued by transfecting the rMV2/StrEP-V and rMV2/StrEP-CH plasmids into helper cells and propagation on Vero cells as previously described (20, 21).
Recombinant viruses properly express tagged MV-V and CH proteins, and replicate at high titers. To detect the expression of MV-V and CH-tagged proteins from rMV2/STrEP-V and rMV2/STrEP-CH, Vero cells were infected with recombinant viruses, and protein expression was determined 24 h post-infection by western blot analysis. High levels of tagged MV-V and CH proteins were detected in infected cells (Fig. 2A). In rMV2/STrEP-V infected cells, a second One-STrEP-tagged protein with a lower molecular weight than One-STrEP-tagged MV-V was also detected by Streptavidin-HRP blotting. This protein is most likely produced from One-STrEP-tagged MV-V encoding mRNA molecules where one additional non-templated guanine nucleotide has been inserted by the viral polymerase at the editing site (31). This results in a truncated MV-V protein encompassing the amino-terminal PNT region plus five specific C-terminal amino acids. CH expression from rMV2/STrEP-CH was also assessed by fluorescent microscopy (Fig. 2B). Altogether, this demonstrates that additional transcriptional units within rMV2/STrEP-V and rMV2/STrEP-CH are properly expressed.

We also tested the potential impact of such additional transcriptional units on viral replication by determining single-step growth curves of rMV2/STrEP-V and rMV2/STrEP-CH viruses on Vero cells (Fig. 2C). The growth of recombinant viruses was similar to that of unmodified MV (Schwarz), and titers were comparable. In conclusion, both recombinant viruses efficiently propagated through the cell monolayer, with no detectable interference with viral replication, and properly expressed tagged MV-V or CH proteins.

Purification of MV-V and associated protein complexes. The whole purification protocol is presented as a flow chart in Fig. 3. The MV-V protein was co-purified
together with its cellular interacting partners from HEK-293T cells infected with rMV2/STrEP-V recombinant virus at an MOI of 1. One-STrEP-tagged MV-V protein and associated cellular factors were isolated from total cell extracts collected 24 h post-infection by single-step affinity purification using Strep-Tactin Sepharose beads. MV-V and interacting partners were subsequently released from the beads by TEV cleavage and protein samples were resolved by SDS-PAGE electrophoresis (Fig. 4A). Many proteins were co-purified with MV-V, whereas in contrast, only a limited number of proteins were co-purified with CH from the rMV2/STrEP-CH negative control. Most importantly, we confirmed that two well-characterized MV-V binding partners, STAT1 and STAT2, were detected among cellular proteins that co-purified with STrEP-V but not STrEP-CH (Fig. 4A lower panel). Thus, protein samples were analyzed directly by nano-LC-MS/MS, without the need to separate proteins by SDS-PAGE electrophoresis and perform in-gel digestions.

Identification of virus-host protein partners by MS. To obtain a comprehensive list of both direct and indirect interactors of MV-V protein, HEK-293T cells were infected with either rMV2/STrEP-V or rMV2/STrEP-CH at an MOI of 1. Tagged MV-V or CH proteins were purified 24 h post-infection by the affinity chromatography presented above, and co-purified cellular proteins were directly identified by nano-LC-MS/MS analysis. The same experiment was repeated three times, and only cellular proteins that were co-purified with MV-V but never found in association with the control CH protein were considered (Fig. 4B and supplementary Table 1). Interestingly, we found large overlaps between all three experiments, thus demonstrating the robustness of this experimental approach. A list of 245 unique MV-V partners was generated that contains only cellular proteins found at least in two independent experiments (supplementary Table 2). Within this list, best-known direct
interactors of MV-V were identified, including STAT1, STAT2 and IFIH1, confirming the sensitivity of this method. We also found p53 that was previously reported to interact directly with the C-terminal VCT region of MV-V (19). In contrast, we failed to identify five known interactors of MV-V: p73, LGP2, IKK-α, IRF7 or JAK1. A weak expression in HEK-293T cells (e.g. IRF7) or a strong association with membrane-associated receptors (e.g. JAK1) could explain their absence from this dataset. Altogether, these data show that our mapping strategy is sensitive, and provides reproducible data.

Specificity of our virus-host protein mapping approach. Bait-prey interactions identified using our approach represent both direct and indirect interactors of MV-V protein. In addition to cellular partners, MV-V was also found to interact with viral proteins, in particular MV nucleoprotein N (MV-N, supplementary Table 3). This interaction has been previously reported, and is mediated by the N-terminal PNT region of MV-V (32). This suggests that some of the 245 cellular proteins co-purified with MV-V could be indirect interactions mediated by N as a bridge. To address this question, we engineered a new virus expressing the MV-N protein with a One-STrEP C-terminal tag (designated rMV2/N-STrEP). Previous reports have demonstrated that known MV-N functions are not affected by the addition of C-terminal tag (33). As a negative control, we also produced a recombinant virus expressing the CH protein with a C-terminal tag (designated rMV2/CH-STrEP) (Fig. 5A). The expression of tagged MV-N and CH proteins was validated by western blot (Fig. 5B). Recombinant rMV2/N-STrEP and rMV2/CH-STrEP were also shown to replicate at high titers (Fig. 5C). HEK-293T cells were infected as described above with rMV2/N-STrEP or rMV2/CH-STrEP viruses, and co-purified protein complexes were similarly analyzed. Protein profiles obtained by SDS-PAGE electrophoresis were obviously different from
those observed with MV-V with only few co-purified proteins specific of MV-N (Fig. 5D). This result was confirmed by nano-LC-MS/MS analysis where only 8 unique cellular partners were identified for MV-N (supplementary Table 4 and 5) and only 2 of them were in the list of the unique MV-V interactors (Fig. 5E). Altogether, this demonstrates that most of MV-V partners, which can be either direct or indirect interactors, are highly specific of this viral protein since they were not identified with MV-N in a parallel experiment.

**Specificity of MV-V protein interaction with cellular partners in the absence of viral replication.** To determine if the interactions reported above between MV-V and host proteins are dependent on the micro-environment induced by MV infection, we expressed the STrep-V protein alone in HEK-293T cells and characterized its binding partners by affinity chromatography and MS analysis. In addition, full-length MV-V and the PNT region alone were compared for their ability to bind cellular partners, since this N-terminal MV-V fragment is intrinsically disordered and thus potentially responsible for interactions with multiple protein partners (34, 35). We generated plasmids encoding for full-length MV-V or its PNT region alone fused to the amino-terminal One-STrep tag (designated pCI-neo/STrep-V or pCI-neo/STrep-PNT, Fig. 6A). As a negative control, we also produced a plasmid encoding for a One-STrep-tagged CH protein (designated pCI-neo/CH-STrep). These plasmids were transiently transfected in HEK-293T cells. Then, MV-V and PNT protein complexes were isolated from total cell extracts collected 24h post-transfection, and the single-step affinity purification was performed as described above. Protein profiles obtained by SDS-PAGE electrophoresis for MV-V and PNT co-purified protein complexes differed from one another (Fig. 6B). Many proteins were co-purified with MV-V, but only a limited number of cellular proteins were co-purified with
the PNT region alone or the CH-STrEP negative control (Fig. 6B). Using nano-LC-MS/MS analysis, two lists of cellular binding partners were established, where only cellular proteins that were co-purified with either MV-V or PNT but not found in association with the control CH protein were considered (Supplementary Tables 6). In total, 323 cellular proteins were found to co-purify with MV-V when expressed alone by transient transfection. This list was compared to the list of 245 MV-V partners generated in the viral context (Fig. 6C and supplementary Table 7). A large overlap was observed between the two lists, with 57.9% of the interactions recovered from MV-infected cells that were also found when expressing MV-V alone. These results show a strong correlation between the two interactions lists, supporting the fact that most interactions detected in MV-infected cells are dependent on MV-V protein expression alone. Interestingly, only 20 interactions were identified with the PNT region alone, and most of them were common to full-length MV-V. STAT1 was detected among cellular proteins that co-purified with both STrEP-V and STrEP-PNT, thus confirming our previous results showing that the PNT region is responsible for the direct MV-V interaction with STAT1 (15, 17). These results further support the high specificity of our purification strategy since only the full-length MV-V protein, and not its N-terminal PNT domain expressed in the same conditions outcomes in a large list of cellular partners. They also demonstrate that the C-terminal region plays a critical role in MV-V ability to form large complexes with numerous host factors.

**Functional analysis of MV-V interaction network.** We then analyzed MV-V cellular interactors for functional enrichment using protein interaction databases, Gene Ontology annotation, and literature. We first performed a query on protein interaction databases using APID2NET search tool (26). The 245 human cellular proteins that co-purified with MV-V were found connected by no less than 766
protein-protein interactions. This interactome network was displayed using Cytoscape application (Fig. 7), and exhibits a core component of 147 cellular proteins, whereas 98 proteins were either isolated from the core component (79 proteins) or simply not referenced in APID2NET (19 proteins). We then analyzed this network for functional enrichment using Gene Ontology (GO) database and BINGO2 plugin for Cytoscape (28, 29). As expected and because of MV-V binding to STAT1, STAT2, IFIH1 and p53, this network was statistically enriched for cellular components involved in “response to biotic stimulus”, a GO term encompassing cellular “response to viruses”. A significant enrichment was also detected for protein chaperones, suggesting that MV-V triggers an unfolded protein response upon accumulation in infected cells. More surprisingly, MV-V infection network was also enriched for several components of the histone deacetylase complex, suggesting yet uncharacterized interactions with chromatin remodeling and gene silencing machineries. Interactions with multiple components of the protein phosphatase 2A complex also suggest MV-V implications in serine/threonine kinase signalling pathways. Interestingly, this latter interaction was recently confirmed by yeast two-hybrid approach (Caignard G. & al., manuscript in preparation), thus demonstrating MV-V direct binding to protein phosphatase 2A. MV-V infection network was highly enriched for mitochondria, reticulum and ribosomal proteins. Although this unravels close interactions between these cellular components and MV-V, functional consequences on virus replication cycle will require further investigations. Importantly, the profile of cellular modules that co-purified with MV-V was highly conserved, whether this viral protein was expressed from MV genome in infected cells or by transient transfection of plasmid DNA (compare Fig. 7 and Fig. 8). Thus,
MV-V alone has the capacity to interact with these different functional modules, independently of other viral factors.

*Topological analysis of MV-V interaction network.* We then analyzed MV-V infection network for specific topological features. First, we asked whether host proteins that co-purified with MV-V are central in the human interactome network. Local (degree) and global (betweenness) centrality measures were calculated. Briefly, the degree (k) of a protein in a network corresponds to its number of direct partners and is therefore a measure of local centrality. Betweenness (b) is a global measure of centrality, as it measures the number of shortest paths (the minimum distance between two proteins in the network) that pass through a given protein. The degree distribution of cellular proteins co-purified with MV-V was plotted, and compared to the distribution obtained for the whole human interactome network (Fig. 9A). These distributions were significantly distinct (U-test; p-value < 2.2 x 10^{-16}), and MV-V cellular interactors showed an average degree (k = 21.19) that is much higher than the average degree of cellular proteins in the human interactome (k = 7.39). Thus, cellular proteins that co-purified with MV-V exhibit much more cellular partners than normally expected by chance. We also calculated and plotted the betweenness centrality distribution for the cellular proteins that co-purified with MV-V (Fig. 9B). This distribution was significantly different from the one obtained for the whole human interactome network (U-test; p-value < 2.2 x 10^{-16}), and MV-V cellular interactors showed an average betweenness centrality coefficient (b = 3.64 x 10^{-4}) that is much higher than the average betweenness centrality coefficient of cellular proteins in the human interactome (b = 1.06 x 10^{-4}). Thus, betweenness centrality measures show that MV-V interacting partners are enriched for vertices that connect multiple modules in the human interactome. Furthermore, MV-V interacting proteins are
highly interconnected, since they share 2.5 more interactions than expected for a random protein set (Fig. 9C; p-value < 1 x 10^{-5}). Thus, MV-V interacting proteins correspond to well-defined functional modules composed of large protein complexes. This conclusion is clearly in agreement with the GO term analysis described above where MV-V is shown to target cellular organelles like mitochondria or multi-protein complexes like ribosomes, histone deacetylases or protein phosphatase 2A. Altogether, topological analyses of MV-V interaction network demonstrate that MV-V preferentially interacts, either directly or indirectly, with host proteins that are central in the human interactome network, and represent well-defined functional modules as assessed by the enrichment for proteins sharing interactions. This is clearly in the line of previous reports showing that viral proteins preferentially bind host proteins with such characteristics (1-3, 36).

**MV-V interactors are well-known targets of viruses in general.** We thus queried the VirHostNet database using our list of MV-V interactors, and identified a total of 220 distinct virus-host interactions involving 50 different virus species. Of the 245 cellular proteins that co-purified with MV-V, 44.9% correspond to known targets of at least one other virus (110 out of 245, see Fig. 9D and supplementary Table 8). This corresponds to a strong enrichment when compared to the 5.4% of human proteins identified as viral targets in VirHostNet. Statistical significance was assessed by Fisher’s exact test (p-value < 2.2 x 10^{-16}), but also by comparing our dataset with 10,000 random sets of 245 human proteins (p-value < 1 x 10^{-5}; Fig. 9D). This demonstrates that MV-V cellular targets mostly correspond to host factors that are critical to viral replication in general, and thus represent common targets for several viruses. Most importantly, this clearly provides a strong argument to support the
quality of our MV-V infection network obtained using a combination of reverse genetic with proteomic analysis.

Functional evidence of MV-V interaction with histone deacetylase complex proteins. MV-V interaction with cellular translational machinery and mitochondria proteins is not surprising since many viruses hijack these cellular machineries to achieve successful replication. However, we were more surprised by the enrichment of MV-V infection network for several histone deacetylase complex proteins (HDAC). As a proof of concept for the functional validation of this virus-host interaction map built in a viral context, we examined a potential role of HDACs in MV replication. First, we engineered a recombinant virus expressing the firefly luciferase protein (Luc) in the additional transcription unit between MV-P and MV-M proteins. This virus was designated rMV2/Luc and expressed Luc when replicating in target cells. Then, HEK-293T cells were infected with rMV2/Luc at an MOI of 0.2 in the absence or presence of TSA. Treatment with TSA increased MV replication in a dose-dependent manner as assessed by Luc expression (3.2-fold increase with 1 µM of TSA; Fig. 11A), despite a substantial effect on cell viability at highest TSA concentrations (supplementary Fig. 3). MV replication in the presence or absence of TSA was also assessed by fluorescent microscopy using the rMV2/STrEP-CH recombinant virus (supplementary Fig. 4). TSA treatment increased CH expression, and much larger virus-induced syncytia were observed. Thus, the HDAC inhibitor TSA increased MV replication in target cells.

Metastasis-associated proteins (MTA1 and MTA2), Chromodomain-helicase-DNA-binding protein 4 (CHD4 or Mi2-beta), Histone-binding protein (RBBP7), Methyl-CpG-binding domain protein 3 (MBD3), and Transcriptional repressor p66-alpha
(GATAD 2A) are six HDAC components found to interact with MV-V protein in an infectious context (Fig. 7, supplementary Table 2). To further support the biological relevance of these interactions, we knocked-down these proteins and then determined MV replication. HEK-293T cells were transfected with MTA1, MTA2, CHD4, RBBP7, MBD3 or GATAD2A-specific siRNAs, or control siRNAs (CT1 and CT2). Transfected cells were cultured for 24h, infected with rMV2/Luc (MOI 0.5), and then lysed 48h post-infection. Transfection of these different siRNAs had no effect on cell viability (supplementary Fig. 5). Silencing of RBBP7, MBD3, GATAD2A positively regulated MV replication as assessed by luciferase expression (Fig. 11B). These results were confirmed by fluorescent microscopy when using rMV2/STrEP-CH to infect HEK-293T cells. Higher expression levels of CH protein and larger syncytia were observed in cells where RBBP7 or GATAD2A expression was suppressed (Fig. 11C). Thus, siRNA-mediated silencing of HDAC components that are specific MV-V interactors enhances MV replication, suggesting a negative role of deacetylation on viral replication.

DISCUSSION

To map virus-host protein interactions, we developed a new method based on recombinant viruses expressing tagged viral proteins to capture their physical binding partners during infection. More specifically, a reverse genetic system was used to engineer MV genome, and rescue virus strains that encode for tagged MV-V or MV-N proteins. This allowed virus-host protein complex analysis directly from infected cells by combining modified tandem affinity chromatography and mass spectrometry analysis. Our modified One-STrEP tag protocol resulted in highly purified protein
complexes compatible with a direct analysis of their components by nano-LC-MS/MS. Most importantly, we show that reproducible lists of cellular partners were obtained with this approach as assessed by convergent results obtained from three biological replicates. Finally, we established a list of 245 direct or indirect binding partners for MV-V, providing new hypothesis to investigate the biology of this virus (supplementary Table 2).

Among the 245 cellular proteins that co-purified with MV-V, we were able to recover 4 of the 9 cellular proteins previously described as direct binding partners of this virulence factor, including STAT1, STAT2, IFIH1 and p53. Until now, these interactions were essentially documented in cells transiently or stably transfected with MV-V (15, 17-19, 37, 38). To our knowledge, this is the first report demonstrating that all these interactions actually occur in MV-infected cells, in highly relevant conditions where MV-V is encoded by the viral genome itself. Finding previously described interactions nicely supports the relevance of our system. However, we did not recover all previously reported interactions, but this could be explained by technical or biological limitations. Indeed, some previously reported interactions could be too unstable to withstand the stringent conditions that were used to purify MV-V complexes. In addition, some cellular proteins with unusual physical properties can be difficult to detect by MS analysis. Finally, some interactions could only occur in specific cell types or culture conditions but we should be able to address this point in the future. Indeed, our approach is suitable to probe viral protein partners in different cell types or culture conditions all along a virus replication cycle. This represents a major advantage of our strategy and in the future, we should be able to identify interactions that only occur in specific cell types, culture conditions, or infected
tissues. We could even use the same strategy to investigate virus-host protein interactions directly in vivo by processing tissue samples from infected animals.

In addition to previously reported interactions, our strategy led to identify a prosperous list of 245 cellular proteins that co-purified with MV-V (supplementary Table 2). Although MV-V exhibits a large intrinsically disordered N-terminal region that could favor direct binding to numerous partners (for review see (35), most of the interactions identified are likely to be indirect. Indeed, our interaction mapping strategy does not discriminate between direct and indirect protein-protein interactions. In the future, our approach should be combined with other technologies that essentially detect binary protein-protein interactions like Nucleic Acid Programmable Protein Array (NAPPA), Luminescence-based Mammalian IntERActome (LUMIER), Mammalian Protein-Protein Interaction Trap (MAPPIT), or Protein Complementation Assay (PCA) (for review see (39).

Strong bioinformatic and experimental arguments support the high relevance of our list of MV-V interactors. First and as aforementioned, this list of 245 cellular proteins was based on the overlap of at least 2 out of 3 biological replicates, and thus corresponds to reproducible interaction data. In addition, host proteins that co-purified with CH protein from control viruses were systematically discarded to eliminate sticky proteins and potential false positives. To further establish the specificity of our MV-V interactor list, we also built a recombinant virus expressing MV-N with a One-STrEP tag. In the same experimental conditions that were applied to identify MV-V binding partners, only 8 cellular proteins co-purified with MV-N. Interestingly, only two cellular proteins interacted both with MV-N and MV-V (Fig. 5E, supplementary Table 5). This is not surprising since MV-V was previously reported to
interact with MV-N via its N-terminal PNT region (32), and this is confirmed by the presence of MV-N in protein complexes co-purified with MV-V (supplementary Table 3). We also established that a majority of cellular proteins that co-purified with MV-V in infected cells also co-purified with this viral protein when expressed by transient transfection, whereas only a minor subset was co-purified when expressing the PNT region of MV-V alone (Fig. 6C and supplementary Table 7). This demonstrates that our list of MV-V binding partners, although it clearly corresponds both to direct and indirect interactions, is highly specific of this viral protein.

Statistical analyses also support the high relevance of our list of MV-V interactors. Indeed, cellular proteins that co-purified with MV-V are significantly enriched for central proteins with high connectivity and betweenness coefficients, and exhibit strong interconnections in the human interactome network (Fig. 9). These topological features are clearly distinct from those obtained by random sampling, thus demonstrating that our list of MV-V interactors is skewed for host proteins exhibiting specific characteristics. Furthermore, this topological bias was previously reported in other studies that aimed at mapping virus-host protein interactions in a large-scale setting (2, 3, 36). Indeed, viral proteins in general have strong tendency to target host proteins that are central in the human interactome network, and represent key components of essential cellular modules as assessed by their topological features. As a corollary, we found that a large fraction MV-V interacting partners correspond to known targets of other viruses (Fig. 10 and supplementary Table 8). Again, this strong enrichment nicely confirms the relevance of our strategy, and supports the idea that MV-V binds directly or indirectly to a kernel of cellular complexes that are hijacked by most, if not all, viruses.
In addition to previously reported interactions with key components of the host antiviral response including STAT1, STAT2, IFIH1, and p53, MV-V was found to target essential components of ribosomes, reticulum, and mitochondria. As part of the protein synthesis machinery, we identified a set of ribosomal proteins, translation factors, chaperones, and aminoacyl-tRNA synthetases amongst MV-V associated proteins (Fig. 7 and supplementary Table 2). These host proteins might be recruited by MV-V to optimize translation of nascent viral mRNA molecules. Conversely, interactions with host translation machinery could prevent the synthesis of antiviral factors of which expression is controlled at translational level (40). MV-V interactions with endoplasmic reticulum components is also new, and could participate in the processing, maturation and transport of viral proteins from perinuclear regions to the cell surface. The enrichment for mitochondrial proteins is also very important. For the moment, the role of mitochondria in the replication cycle of non-segmented negative-strand RNA viruses is less well studied than for positive-stranded RNA viruses. We identified numerous components of mitochondrial protein precursors import complex (TOM, TIM, VDAC), solute carrier group of membrane transport protein members (SLC), component of the respiratory chain (cytochrome oxidase) and ATP synthetase, membrane enzymes and proteins involved in cell cycle control (prohibitin). Interestingly, mitochondrial proteins like prohibitin, prohibitin 2 and VDAC1 have been shown to interact with nonstructural protein 2 of severe respiratory syndrome coronavirus (SARS-CoV) (41). In addition, major changes in mitochondrial protein expression profiles have been recently observed in cells infected with human respiratory syncytial virus (RSV), another member of Paramyxoviridae family (42). Munday et al. used a stable isotope labelling with amino acids in cell culture (SILAC) coupled to LC-MS/MS analyses. They showed
substantial changes in the abundance and localization of mitochondrial proteins such as the accessory subunit of the mitochondrial membrane respiratory chain NADH dehydrogenase (NDUFB) and components of mitochondrial import complex: TOM20, TOM22, TOM40 and TOM70, VDAC proteins, prohibitin, prohibitin 2 (42). Interestingly, we found NDUFA4, NDUFA8, Tom22 and Tom70 in the list of MV-V interactors. MV-V association with mitochondria components could rely on IFIH-1 binding. MV-V interaction with IFIH1 has been shown to prevent its oligomerisation, but we still ignore whether this prevents IFIH1 interaction with MAVS (43). This latter protein is anchored at the membrane of mitochondria, and is used as a scaffold for downstream signalling components of the innate antiviral response (44). Thus, IFIH1 could be an intermediate that bridges MV-V to mitochondrial membranes. MV-V was also found to interact with multi-protein complexes like the histone deacetylase machinery or protein phosphatase 2A. Interestingly, we have experimental evidences showing MV-V direct binding to subunits of protein phosphatase 2A complex (Caignard et al., manuscript in preparation).

To further demonstrate the functional relevance of the interactors found with MV-V, we selected a subset of proteins and decided to focus on HDAC components. Gene repression is commonly associated with regions of deacetylated histones. Maintenance of deacetylated histones and silent chromatin is an active process mediated by HDAC enzymes. Interestingly, it has been shown that HDAC components counteracting the opposite action on cytokine-induced transcription and histone deacetylase activity is required for optimal cellular defense against virus infection (45-47). This activity can be exceeded on transcriptional level by activation IFN-β and interferon-stimulated gene promoters or on post-translational level by post-translational modification of transcription factors. Consequently, blocking of HDAC
activity by specific inhibitors such as TSA has been shown to increase virus cytopathic effect or virus replication for HCV, encephalomyocarditis, vesicular stomatitis, Sendai and Newcastle disease viruses. In the present report, we validated the stimulation of MV replication by TSA (Fig. 11A and supplementary Fig. 4). Additionally, this correlated with direct and indirect positive effects on MV replication of another HDAC inhibitor such as valproic acid sodium salt (48). Furthermore, the knockdown of RBBP7, GATAD2A or MBD3, three HDAC components that specifically interacted with MV-V in a viral context, also increased MV replication (Fig. 11B). The fact that HDAC inhibitors and the silencing of the components of histone deacetylase machinery provided analogous effect on MV-replication suggests that in addition to already known MV-V interactions with STAT1, STAT2, JAK1, IFIH1, LGP2, IKK-α, IRF7, these new MV-V cellular partners (RBBP7, GATAD2A, MBD3, MTA1, MTA2 and CHD4) may also be important for MV-mediated inhibition of the host innate antiviral response.

Altogether, our data demonstrate the interest of combining virus reverse genetics together with protein complex analysis by nano-LC-MS/MS to decipher virus-host protein interactions. Direct and indirect interactions cannot be distinguished in this approach, but it is extremely valuable because interactions are identified directly from infected cells, in a highly relevant and physiological setting. It also provides a global overview of cellular pathways and functional modules targeted by a viral protein, instead of a limited list of direct interactors that sometimes resist to interpretations. This nicely complements other strategies, including yeast two-hybrid technology or protein arrays, which allow the identification of direct protein interactions but are performed in an artificial context that calls for additional validation steps. In the future, the same strategy could be applied to any other negative-strand
RNA viruses since a RNA virus could easily develop direct interactions with 60-80 different host proteins according to recent estimations, and only a few percents have been characterized (1). Whatever we do, mapping these interactions is a prerequisite to reach a system level understanding of viral infection process.

ACKNOWLEDGMENTS

We would like to thank Dr. Kaoru Takeuchi and Dr. Chantal Rabourdin-Combe for anti-V rAb polyclonal antibody and anti-N mAb monoclonal antibody. We thank Thibaut Jacob for updating GO annotation in GOlorize plug-in. We thank all members of the Unité de Génomique Virale et Vaccination for their support. We acknowledge the financial support of Institut Pasteur and CNRS. Dr. AK is supported by a “Bourse Roux” (Institut Pasteur).
REFERENCES


21. Guerbois, M., Moris, A., Combredet, C., Najburg, V., Ruffie, C., Fevrier, M.,
    vaccine expressing HIV-1 Gag virus like particles covered with gp160DeltaV1V2 is
    strongly immunogenic. Virology 388, 191-203.
    replication of Sendai virus defective interfering RNA. Journal of virology 67, 4822-
    4830.
    Measles virus V protein blocks interferon (IFN)-alpha/beta but not IFN-gamma
    signaling by inhibiting STAT1 and STAT2 phosphorylation. FEBS Lett 545, 177-182.
    interactome graphic analyzer. Bioinformatics 23, 2495-2497.
27. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D.,
    for integrated models of biomolecular interaction networks. Genome Res 13, 2498-
    2504.
    assess overrepresentation of gene ontology categories in biological networks.
    Bioinformatics 21, 3448-3449.
29. Garcia, O., Saveanu, C., Cline, M., Fromont-Racine, M., Jacquier, A.,


**FIGURE LEGENDS**

Fig. 1. **Schematic representation of recombinant virus genomes.** MV negative-sense RNA genome is displayed with its 3’ end on the left, with the six genes indicated by capital letters and depicted as white rectangles. The additional transcriptional units encoding for One-STrEP-tagged MV-V (STrEP-V) or CH protein (STrEP-CH) are inserted between P and M genes. The black oval represents the One-STrEP-tag sequence and the black rectangle corresponds to the Tobacco Etch Virus (TEV) protease coding linker. T7p is a T7 RNA polymerase promoter sequence, hh is a hammerhead ribozyme, T7t is a T7 RNA Polymerase termination signal, ℓ is a hepatitis delta ribozyme.

Fig. 2. **Characterization of recombinant viruses expressing STrEP-V or STrEP-CH proteins.** A, Vero cells were infected with the native MV Schwarz strain, or the rMV2/STrEP-CH expressing the One-STrEP-tagged CH protein, or the rMV2/STrEP-V expressing the One-STrEP-tagged MV-V protein. Expression of native and One-STrEP-tagged MV-V proteins was determined by western blot using anti-V monoclonal antibodies (top panel). Expression of One-STrEP-tagged CH or MV-V proteins was determined by western blot using Streptavidin-HRP conjugates (lower panel). B, Fluorescent microscopy showing efficiency of CH protein expression in HEK293T 24 hours post-infection by rMV2/STrEP-CH. Fluorescence images were taken with a 10x objective. White scale bar correspond to 100 µm. C, Single-step growth curves obtained for rMV2/STrEP-V and rMV2/STrEP-CH. MV Schwarz strain was used as a control. Vero cells in 6-well dishes were infected with MV Schwarz, rMV2/STrEP-V, or rMV2/STrEP-CH at an MOI of 0.1. Cell-associated virions were recovered at each time point, and titers were determined using the TCID50 method.
Fig. 3. **Summary of the protocol used to purify STrEP-V and associated cellular proteins from infected cell lysates.** HEK-293T cells were infected with rMV2/STrEP-V or rMV2/STrEP-CH at an MOI of 1. 24 h post-infection, STrEP-V was co-purified with interacting cellular proteins using StrepTactin Sepharose beads. After two subsequent washing steps, protein complexes were released from the beads by specific cleavage of a flexible linker by the TEV protease. Protein complexes were precipitated and quantified. Finally, purified virus-host protein complexes were analyzed by direct nano-LC-MS/MS. White circles represent StrepTactin Sepharose beads, white rhombus represent unbound proteins and contaminants, and grey rhombus correspond to MV-V interacting partners. A white lightning indicates the cleavage site for the TEV protease. Other indications are as in Fig. 1.

Fig. 4. **Purified protein complex analysis.** A, Purified protein complexes were resolved by SDS-PAGE and silver staining (top panel). Same protein samples were resolved by SDS-PAGE followed by western blot analysis to detect MV-V interactions with STAT1 and STAT2 proteins (lower panel). B, Venn diagram showing overlaps between three lists of MV-V interacting partners established from three independent experiments. For each experiment, a negative control was obtained by processing rMV2/STrEP-CH infected cells. Any host protein that co-purified with STrEP-CH in at least one of the three experiments was systematically discarded from the list of MV-V interacting partners. A list of MV-V interacting partners that contains 245 cellular proteins was established from the overlap between at least two independent experiments.
Fig. 5. **MV-V and MV-N exhibit very distinct lists of cellular binding partners.**


Fig. 6. **Specificity of MV-V interaction with cellular partners in the absence of viral replication.**


Fig. 7. **A map of the human protein interaction sub-network targeted by MV-V.**

From the list of cellular interactors of MV-V in infected cells identified by nano-LC-MS/MS analysis, we queried interaction databases using APID2NET search tool. Retrieved interactions were displayed using Cytoscape network analysis tool. Statistical enrichment for specific GO terms was determined using BINGO2 plug-in. A selection of statistically significant GO terms (p-value < 0.01) was displayed on the map by colouring the corresponding nodes with GOlorize plug-in.
Fig. 8. A map depicting an overlap of two human protein interaction sub-networks targeted by MV-V in the presence or absence of viral infection. The same as on Fig. 7 but 142 cellular proteins found to interact with MV-V in the absence of viral background are shown in red and represent a large overlap between MV-V interactors recovered in the presence and absence of viral infection.

Fig. 9. Topological analysis of MV-V infection network. A, Degree and B, betweenness distributions of human proteins (black) and human proteins co-purified by MV-V (red) in the human interactome. P(k) is the probability of a node to connect k other nodes in the network. P(b) is the probability of a node to have a betweenness equal to b in the network. Solid lines represent the linear regressions. Vertical dashed lines correspond to mean degree and betweenness values. C, Human proteins co-purified with MV-V are highly interconnected. Protein sets were randomly drawn from the human proteome, and protein-protein interactions in the corresponding subgraph were numbered. This was repeated 10,000 times before plotting the observed distribution. The value observed with the set of proteins that co-purified with MV-V (168 interactions) is indicated by a vertical arrow, and demonstrates a statistically significant enrichment for highly interconnected cellular proteins. D, Cellular proteins interacting with MV-V are preferential targets of viruses. Protein sets were randomly drawn from the human proteome, and the number of proteins targeted by at least a viral protein was determined. This was repeated 10,000 times, and the corresponding distribution was plotted. The value observed with the set of proteins that co-purified with MV-V (110 cellular proteins targeted by at least one virus) is indicated by a vertical arrow, and demonstrates a statistically
significant enrichment for cellular proteins that represent frequent targets for viruses in general.

Fig. 10. **MV-V interactors are well-known targets of viruses in general.** VirHostNet database was queried to identify cellular proteins from MV-V interaction map that were previously identified as binding partners of other viruses. A yellow-to-red colour range was applied to identify host proteins targeted by 1, 2, 3, 4 or more than 4 distinct viruses.

Fig. 11. **Down-regulation of components of histone deacetylase complex (HDAC) up-regulates MV replication.** A specific HDAC inhibitor TSA up-regulates virus replication. HEK-293T cells were infected with rMV2/Luc at an MOI of 0.2 and treated with TSA at different concentration, mock-treated or mock-treated in the presence of 0.02% DMSO. Luciferase assay was performed 48h post-infection. Each bar is the mean ± standard error of the mean of luciferase activity measured in duplicate from a representative experiment repeated three times. B and C siRNA-mediated down-regulation of endogenous components of HDAC enhances MV replication in 293T cells. HEK-293T cells were transfected with siRNAs targeting six different components of HDAC complex (RBBP7, GATAD2A, MTA2, MBD3, CHD4 and MTA1), transfected with two control RNAs (CT1 and CT2) or mock-transfected. 24h post-transfection cells were infected with rMV2/Luc at an MOI of 0.5 and luciferase assay was performed 48h post-infection. One and two red asterisks demonstrate statistically significant enrichments with P-values <0.05 and <0.01, respectively in Luc production in comparison with mock-transfected cells. Each bar is the mean ± standard error of the mean of luciferase activity measured in triplicate
from a representative experiment repeated three times. C, Light and fluorescence microscopy analyse of Cherry protein expression by rMV2/STrEP-CH 48 h post-infection on HEK-293T cells transfected or not with siRNAs against specific components of HDAC complex. Black arrows indicate syncytia formed by infected cells. Fluorescence images were taken with a 5x objective. Light images were taken with the same objective but magnified 1.5 times with Bertrand lens. White and black scale bars correspond to 500 µm.
Komarova et al. Figure 1
Cell infection with tagged recombinant virus

Virus replication

Purification of the tagged proteins together with interacting cellular proteins

2X wash

TEV cleavage

SDS-PAGE analysis or direct LC-MS/MS analysis

Komarova et al. Figure 3
Komarova et al. Figure 4

STrEP-tag purification.
Blot: STAT1 and STAT2

Exp 1
Exp 2
Exp 3

STrEP-V
STrEP-CH
Marker, kDa
rMV2/STrEP-CH
rMV2/STrEP-V

19
28
39
51
13
66
20
19
146
63
66
129
Composite figure showing experimental data and schematic diagrams.

**A**
Schematic diagram of the TEV cleavage sites in MV-STrEP and CH-STrEP.

**B**
Western blot analysis of cell lysates showing N-STrEP and CH-STrEP bands.

**C**
Graph showing the titer (TCID50/mL) over time post-infection for different strains.

**D**
Western blot analysis of marker bands showing N-STrEP and CH-STrEP.

**E**
Diagram illustrating the molecular weights in kilodaltons (kDa) for different proteins.

Komarova et al. Figure 5
Komarova et al. Figure 7
Komarova et al. Figure 8
Figure 9

A. Probability distribution (log-log scale)

B. Probability distribution (log-log scale)

C. Observed interconnectivity
P-value < 0.0001

D. Observed nb of proteins
P-value < 0.0001
Number of viruses binding the same cellular protein:

1  2  3  4  >4

Komarova et al. Figure 10
Figure 11

Panel A: Bar graph showing luminescence (RLU) as a function of concentration (µM) for various concentrations ranging from 0 to 1 µM with DMSO as a control.

Panel B: Bar graph depicting luminescence (RLU) for different siRNAs: Mock, CT1, CT2, RBBP7, GATAD2A, MTA2, MBD3, CHD4, MTA1.

Panel C: Microscope images showing fluorescence for mock, CT1, RBBP7, and GATAD2A, with arrows indicating specific areas of interest.