

ARTICLE

Received 1 May 2014 | Accepted 28 Jul 2014 | Published 16 Sep 2014

DOI: 10.1038/ncomms5816

Conserved and host-specific features of influenza virion architecture

Edward C. Hutchinson¹, Philip D. Charles¹, Svenja S. Hester¹, Benjamin Thomas¹, David Trudgian^{1,†}, Mónica Martínez-Alonso¹ & Ervin Fodor¹

Viruses use virions to spread between hosts, and virion composition is therefore the primary determinant of viral transmissibility and immunogenicity. However, the virions of many viruses are complex and pleomorphic, making them difficult to analyse in detail. Here we address this by identifying and quantifying virion proteins with mass spectrometry, producing a complete and quantified model of the hundreds of host-encoded and viral proteins that make up the pleomorphic virions of influenza viruses. We show that a conserved influenza virion architecture is maintained across diverse combinations of virus and host. This 'core' architecture, which includes substantial quantities of host proteins as well as the viral protein NS1, is elaborated with abundant host-dependent features. As a result, influenza virions produced by mammalian and avian hosts have distinct protein compositions. Finally, we note that influenza virions share an underlying protein composition with exosomes, suggesting that influenza virions form by subverting microvesicle production.

¹Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK. † Present address: Biochemistry Department and Proteomics Core, UT Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9038, USA. Correspondence and requests for materials should be addressed to E.C.H. (email: edward.hutchinson@path.ox.ac.uk) or to E.F. (email: ervin.fodor@path.ox.ac.uk).

Viruses, unlike other parasitic nucleic acids, can cause infected cells to produce virions. These extracellular structures allow copies of the viral genome to be transferred to new hosts, and their composition is therefore the primary determinant of the stability, transmissibility, tropism and immunogenicity of a virus. Virions are assembled from virally encoded proteins and may also include proteins encoded by the host, particularly if the virion incorporates an envelope of host membrane. While some virions have a regular architecture whose structure can be clearly determined, many important viral pathogens have pleomorphic virions whose structure can only be determined with difficulty and through the combination of different techniques. In this study, we have used mass spectrometry as an alternative approach for the detailed characterization of pleomorphic virions.

Influenza viruses are serious clinical and veterinary pathogens which cause epithelial cells to produce enveloped, pleomorphic virions^{1–6}. These virions are constructed from a mixture of viral and host proteins⁷, but we only know the quantities within them of a small number of highly abundant viral proteins with distinctive biophysical properties or electrophoretic mobilities^{1–3,6,8–22}.

Here, to characterize influenza virions in greater detail, we analyse their protein composition using a sensitive liquid chromatography and tandem mass spectrometry (LC–MS/MS) approach, and then determine protein abundance by using the intensities of fragment ion spectra to perform label-free absolute quantitation. This allows us to establish a complete and quantified model for the protein composition of influenza virions. We show that the core architecture of spherical influenza virions is maintained across diverse combinations of viruses and hosts, but that it is elaborated with additional architectural features which depend on the host species. Host proteins, including those whose incorporation is species dependent, make a substantial contribution to influenza virion architecture. Furthermore, the host-encoded proteins in influenza virions are strikingly similar to the protein profile of virion-sized microvesicles shed by uninfected cells, suggesting that the virus subverts cellular pathways normally used for microvesicle formation to produce enveloped virions.

Results

Label-free protein ratios determined from mass spectra. Mass spectrometry provides a large amount of information which can be used to estimate protein abundance. This can be conveniently performed by spectral index normalized quantitation (SINQ)²³, a method for label-free absolute protein quantitation of data from LC–MS/MS analysis which is available as part of the Central Proteomics Facilities Pipeline software²⁴, and which is based on the intensities of fragment ion spectra. The method considers the summed intensity of fragment ions assigned to each protein, weighted to correct for ambiguous protein assignments, sample complexity and the tendency of longer proteins to produce more peptides²³. By analysing mixtures of proteins of known concentration, we determined that SINQ could simultaneously quantify the abundance of a large number of specific proteins in a mixture. As expected, the accuracy with which protein ratios were determined depended on both the proteins' abundance and the particular proteins being compared. However, we found that SINQ could determine most equimolar and tenfold protein ratios to within fourfold of the correct value and that even over a 1,000-fold difference in abundance it could correctly place ratios within eightfold (Methods and Supplementary Fig. 1a–c).

The composition of an influenza virion. We used SINQ to determine the ratio of viral proteins in virions produced by

bovine epithelial (MDBK) cells infected with influenza A/WSN/33 virus (WSN). After 2 days of infection, the cells' growth medium was harvested and subjected to low-speed centrifugation to remove cellular debris. Virions were then purified from the supernatant by ultracentrifugation through a sucrose gradient (Fig. 1a,b). By comparing data collected in three separate biological repeats (Table 1), we obtained reproducible estimates of viral protein abundance that varied over three orders of magnitude (Fig. 1c, Table 2). As in a previous study²⁵, we detected all nine of the recognized viral structural proteins, as well as the NS1 protein. We did not detect peptides unambiguously identifying the truncated proteins PB1-N40, PA-N155 or PA-N182, and we searched for, but did not detect, PB1-F2, PA-X, M3, M4, M42, NS3 and NSP/NEG8, which have been proposed or positively identified as viral gene products but which have not been shown to be structural proteins²⁵. In addition, we identified over 300 host-encoded proteins. No more than eight copies of the viral polymerase are thought to be incorporated per virion^{2,26,27}, and we reasoned that proteins found at less than a tenth of this abundance must be present at less than one copy per virion. We could reproducibly identify proteins even at abundances hundreds of times below this threshold, indicating that we had assessed the complete consistent protein composition of an influenza virion without the need for further fractionation of the sample.

Our estimates of protein abundance were in agreement with the limited data already available on influenza virion composition (Supplementary Data 1). For example, we correctly determined an equimolar ratio for the subunits of the trimeric viral polymerase (PA, PB1 and PB2) despite their low abundance in the virion. Accurate comparisons could also be made between proteins of high and low abundance. The ratio of the polymerase to the highly abundant nucleoprotein (NP; mean 66, s.d. 14), combined with estimates that each NP binds at least 24 nt of RNA²⁸ and that the complete influenza genome is bound to eight polymerase trimers²⁹, suggests a minimum genome size of 12.7 kb (s.d. 2.6 kb); the actual size is 13.6 kb. Some caution is needed for glycoproteins and transmembrane proteins whose features suppress the efficiency of detection by mass spectrometry. However, even for these problematic proteins, we were able to make consistent measurements that were comparable with those in the literature (Supplementary Data 1), allowing us for the first time to assess the abundance of all viral structural proteins simultaneously.

We were surprised by our consistent identification of NS1 in purified virions (Fig. 1c), as this protein had previously been thought to be non-structural. The protein was unambiguously identified by LC–MS/MS (Fig. 1d), and could also be detected in purified virions using a panel of specific antibodies (Supplementary Fig. 2a–c). When we examined the behaviour of NS1-bearing material during ultracentrifugation, we found that NS1 migrated through density gradients at the position of a visible band of virions. It comigrated with the viral polymerase and with actin, one of the more abundant host proteins in virions (Fig. 1e and Supplementary Fig. 2d). In samples of lysed virions, NS1 did not migrate with NP on glycerol gradients, suggesting it does not form stable associations with viral ribonucleoprotein complexes (RNPs; Supplementary Fig. 2e), but in samples of unlysed virions it was at least as resistant to protease treatment as NP, suggesting that it is an internal component of the virion (Supplementary Fig. 2f–h). We therefore concluded that NS1 is found at low levels within influenza virions.

A conserved architecture of viral proteins. We next compared viral proteins in virions shed by a selection of different hosts

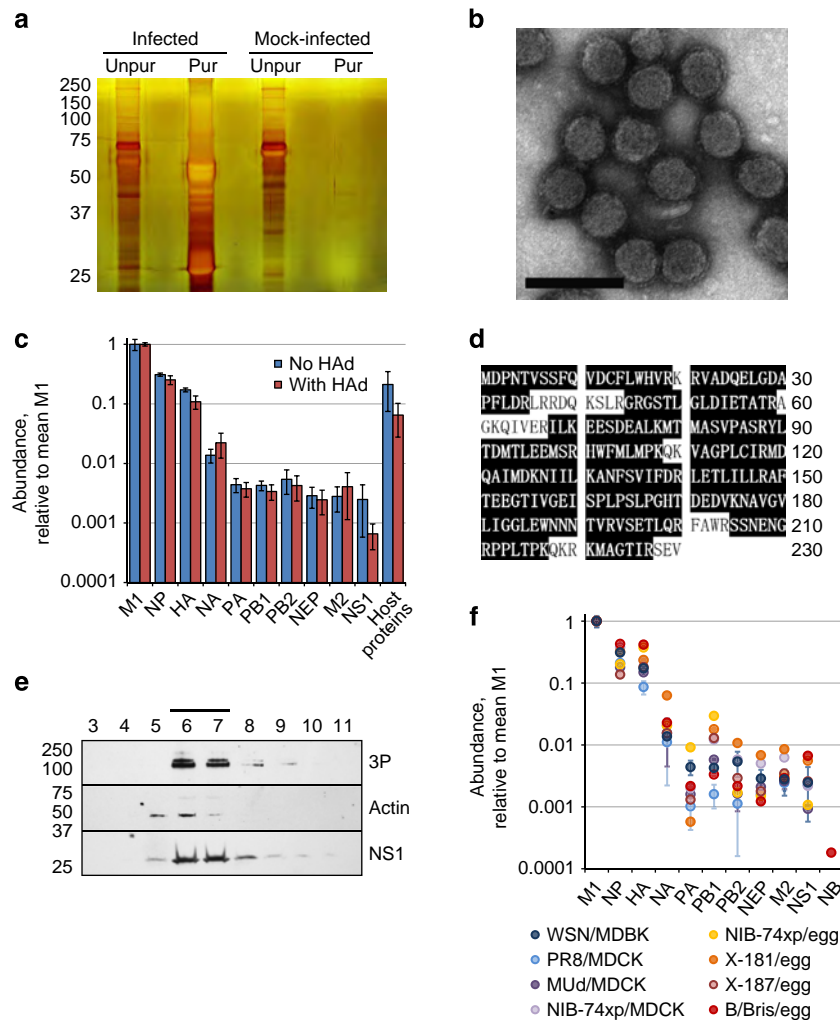


Figure 1 | The abundance of viral proteins in influenza virions. Bovine epithelial (MDBK) cells were infected with influenza A/WSN/33 virus (WSN) or mock infected, and media harvested after 48 h. **(a)** Samples of media were separated by SDS-PAGE either before (diluted 1/1,000) or after purification by gradient ultracentrifugation, and silver stained. **(b)** Purified influenza A/WSN/33 virions, negatively stained and visualized by EM; scale bar, 200 nm. **(c)** Abundance of viral proteins in purified virions, calculated by SINQ and normalized by the mean level of M1. The total abundance of all host proteins detected is also shown. The mean and s.d. is shown of three separate experiments, performed either with or without a haemadsorption/elution step (HAd). **(d)** The NS1 sequence of WSN shaded to show peptides identified by LC-MS/MS (from three separate experiments, purified without HAd; 87% coverage). **(e)** Virions were layered onto a 30–60% sucrose gradient and subjected to ultracentrifugation. The gradient was then harvested in 15 fractions from the top; the fractions shown were analysed by 14% SDS-PAGE and western blotting for the indicated proteins. The position of molecular weight markers is indicated in kDa; concentrated virions were visible as a band in the gradient between fractions 6 and 7 (line). **(f)** Abundance of viral proteins in virions of six different influenza A viruses and an influenza B virus, grown in a variety of hosts (bovine epithelial (MDBK) cells, canine epithelial (MDCK) cells and embryonated chicken eggs; see Table 1 for details). For WSN, the mean and s.d. of three separate experiments is shown, and for PR8 and MUd the mean and range of two separate experiments.

(mammalian epithelial tissue cultures and embryonated chicken eggs) infected with a selection of influenza viruses (Table 1). In clinical isolates, a mixture of spherical and filamentous influenza virions is typically found, but most influenza viruses adapted to growth in tissue culture or in eggs form only spherical virions³⁰. Filamentous influenza virions are highly diverse in their morphology and presumably have highly variable protein compositions as a result⁴. As we were analysing the bulk properties of samples of virions, we concentrated on spherical virions which have a more uniform morphology.

Viral proteins were incorporated at consistent levels for a selection of influenza A virions, regardless of the strain or host (Fig. 1f). The same ratio was observed for orthologous proteins in virions of an influenza B virus (B/Bris), despite the differences in protein sequence and regulation of gene expression that have

emerged since this genus diverged from the influenza A viruses³¹. All virions clearly contained NS1; B/Bris also incorporates low levels of the NB protein, for which there is no influenza A virus ortholog (although as this is a glycosylated transmembrane protein, we expect its abundance to be somewhat higher than estimated by SINQ^{32,33}). Despite the lack of a fixed virion structure, we show that spherical influenza virions incorporate their viral proteins in a consistent ratio, an optimal architecture that has been maintained throughout the separate evolution of influenza A and B viruses.

Influenza virions contain abundant host proteins. Host proteins have been identified in influenza virions, but their abundance has not previously been assessed. While most of the

Table 1 | Viruses and hosts used in the study.

Name	Virus	Reassortment*	Host	Experiments [†]	Replicates [‡]	
					Q Exactive	Orbitrap
WSN	A/WSN/33 (H1N1)		MDBK cells	3 (virus; no HAd), 3 (mock; no HAd) 3 (virus; with HAd) 3 (mock; with HAd)	1 1 1 1	
PR8	A/Puerto Rico/8/ 1934 (H1N1)		MDCK cells	2	1	
MUd	A/Puerto Rico/8/ 1934 (H1N1)	M from A/Udorn/307/1972 (H3N2)	MDCK cells	2	1	
NIB-74xp	NIB-74xp (PR8)	HA and NA from A/Christchurch/16/2010 (pdmH1N1)	MDCK cells, chicken eggs	1 (MDCKs), 1 (eggs)		2
X-181	NYMC X-181 (PR8)	HA, NA and PB1 from A/California/7/2009 (pdmH1N1)	Chicken eggs	1		2
X-187	NYMC X-187 (PR8)	HA and NA from A/Victoria/ 210/2009 (H3N2)	Chicken eggs	1		2
B/Bris	B/Brisbane/60/2008		Chicken eggs	1		2

*Donor segments, named after major gene product.
[†]Number of purified samples analysed.
[‡]Number of replicate analyses of each sample, performed on the instrument indicated.

Table 2 | The abundance of viral proteins in influenza A/WSN/33 virions.

Protein	Relative to M1*		Relative to Pol [†]	
	Mean	s.d.	Mean	s.d.
M1	1.0	0.2	1,704	360
NP	0.31	0.02	530	32
HA	0.172	0.013	294	23
NA	0.014	0.004	23	6
M2	0.0028	0.0013	4.7	2.2
NEP	0.0029	0.0011	4.9	1.9
PA	0.0044	0.0012	7.5	2.0
PB1	0.0043	0.0008	7.3	1.4
PB2	0.0054	0.0024	9.2	4.1
NS1	0.0025	0.0019	4.2	3.2

Absolute quantification of proteins from virions produced by WSN-infected bovine epithelial (MDBK) cells, purified by gradient ultracentrifugation without haemadsorption (HAd). Mean and s.d. of three experiments
*Scaled so that the mean abundance of M1 is 1.
[†]Scaled so that the mean abundance of the three polymerase subunits (PA, PB1 and PB2) is 8.

host proteins we identified were present at very low abundance and could not have been consistently incorporated into virions (Supplementary Data 2), several hundred proteins were abundant enough to be virion components (Supplementary Data 3–5), and together they made a substantial contribution to virion structure (Fig. 1c). More than a dozen of these proteins were more abundant than the viral polymerase, including some proteins whose abundance appeared roughly equal to that of the viral NA protein. These include proteins such as tetraspanins, which due to their substantial transmembrane domains are likely to be present in even higher abundance than our measurements suggest. Thus, purified virions contain a large number of host proteins at low abundance and a smaller set of host proteins that are major components of the virion.

Virion architecture is shaped by the host. To compare proteins between different hosts, we matched them to their human

orthologs (Supplementary Data 6). Differences in protein abundance were observed between all samples analysed, presumably due to a combination of genuine variation between virions and variations in the efficiency of detection (Supplementary Data 2). However, principal coordinates analysis captured the majority of the differences in log(abundance) in two axes (Fig. 2a). Visual inspection of this plot showed that proteins formed three clusters: proteins present only in virions from mammalian hosts (either bovine or canine epithelial tissue cultures), proteins present only in virions from avian hosts (embryonated chicken eggs) and proteins present in virions from any host. Each category contained proteins abundant enough to be consistent components of virions (Fig. 2b). Thus, influenza virions have a consistent core architecture which includes proteins encoded by the host as well as by the virus, as well as a variable complement of proteins whose incorporation depends on whether the virion was constructed in a mammalian or avian host (Fig. 2c).

Consistent with previous studies⁷, all virions incorporated abundant ubiquitin, annexins, cytoskeletal proteins and glycolytic enzymes, as well as peptidyl-prolyl *cis-trans* isomerase A (also known as cyclophilin A), a restriction factor for influenza replication^{7,34}. We also detected abundant membrane proteins, small GTPases and other regulators of signalling (including several 14-3-3 isoforms) as part of the core architecture. Many of the mammalian- and avian-specific proteins also fell into these categories, suggesting that the same functional niches in the virion are being filled (Supplementary Data 4 and 5). Notably, the tetraspanin CD9 is among the most abundant of the mammalian-specific proteins; in egg-grown virions, its place appears to be taken by similar amounts of uroplakin-1B (UPK-1B), a member of a separate tetraspanin family. Other proteins are unique to one host type, for example, the ubiquitin-like protein ISG15 was only detected in virions grown in canine epithelial (MDCK) cells. A full list of 'core' and host-dependent proteins is given in Supplementary Data 3–5.

Uninfected cells shed material which resembles virions. We purified virions using standard techniques based on sucrose gradient ultracentrifugation. To examine the stringency of these

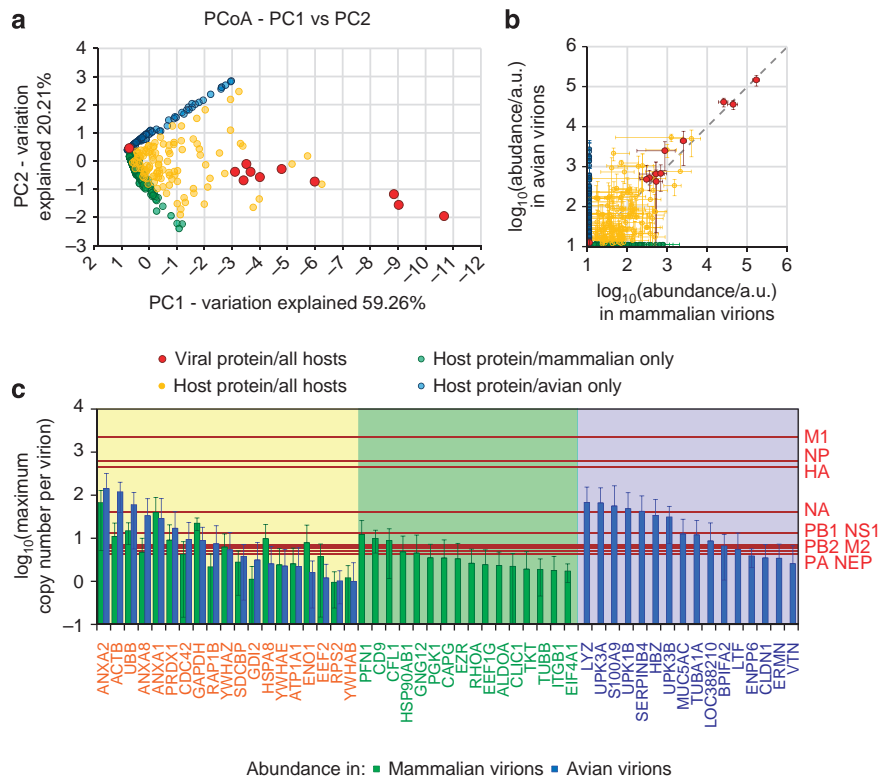


Figure 2 | The abundance of host proteins in influenza virions. Protein abundance in purified virions was determined by SINQ. A threshold was set at one-tenth the abundance of the least abundant polymerase subunit, and any protein of lower abundance was assigned this value. Proteins were matched to their human orthologs to allow comparisons between virions from different host species. **(a)** Principal coordinates analysis of $\log_{10}(\text{protein abundance/a.u.})$ for 548 protein orthologs (goodness of fit = 0.795). **(b)** $\log_{10}(\text{abundance/a.u.})$ of proteins in virions from mammalian cells and in virions from avian cells; mean and s.d. of four combinations of viruses and mammalian hosts, and four combinations of viruses and avian hosts (Table 1). **(c)** The core and host-dependent features of influenza virions. The maximum copy number of proteins in virions is shown, calculated by assuming that an average virion incorporates no more than eight viral polymerases. Viral proteins are shown as red lines, and host proteins with a mean copy number greater than one are shown as bars. Host proteins are only shown if detected in all four combinations of virus and host for mammalian hosts (green bars) or for avian hosts (blue bars); the mean and s.d. is indicated. The core architecture of a spherical influenza virion consists of the viral proteins (red lines) and those host proteins found in all hosts (yellow background). Additional host proteins are found only in virions from mammalian hosts (green background) or from avian hosts (blue background).

methods, we purified virions from the media of WSN-infected MDBK cells and performed parallel purifications on the media of mock-infected cells (Table 1). Using SDS-polyacrylamide gel electrophoresis (PAGE) and silver staining, we could detect abundant proteins in samples prepared from the media of infected cells but we could barely detect material purified from the media of mock-infected cells (Fig. 1a). Using LC-MS/MS, we could detect proteins in both infected and mock-infected samples, though the diversity and abundance of proteins in the mock-infected samples were much less (Fig. 3a,b). The conditions used to purify influenza virions were designed to exclude large fragments of cellular debris, and are very similar to procedures used to purify small extracellular microvesicles such as exosomes^{7,35–38}. We therefore concluded that uninfected cells shed low levels of microvesicles of a similar size to influenza virions.

Abundant host proteins are stably associated with virions. To determine whether host proteins were present in virions or in co-purifying microvesicles, we increased the stringency of our purification method by introducing an initial step in which haemadsorption (HAD) to and elution from red blood cells selected for material with both receptor-binding and -cleaving activities (as provided by the viral haemagglutinin and

neuraminidase, HA and NA, respectively). This step greatly increased the stringency of purification—other than protein standards and common contaminants, only two proteins were consistently found in the mock-infected sample, and these at very low levels (Fig. 3a,c).

In the infected samples, HAD also increased stringency and removed a portion of the host proteins, mainly those of low abundance (Fig. 1c). These included common cytoplasmic proteins that are plausible contaminants. For example, 57 of the 61 ribosomal proteins identified in the infected samples were removed by HAD (a full list of the proteins detected is given in Supplementary Data 2 and their correspondence with other experiments is given in Supplementary Data 3 and 4). HAD should not remove components of the virion, and consistent with this it did not affect the abundance of viral proteins (Fig. 1c; the mean abundance of NS1 was slightly reduced but not significantly so: $P = 0.2$, two-tailed unpaired t -test). The detection of abundant host proteins after more stringent purification demonstrates that these too are components of the virion (Figs 1c and 3a,c).

Influenza virions and exosomes share architectural features. Almost all of the proteins found in the uninfected material are known markers of exosomes^{39,40}, suggesting that this material did indeed consist of microvesicles. Although these microvesicles

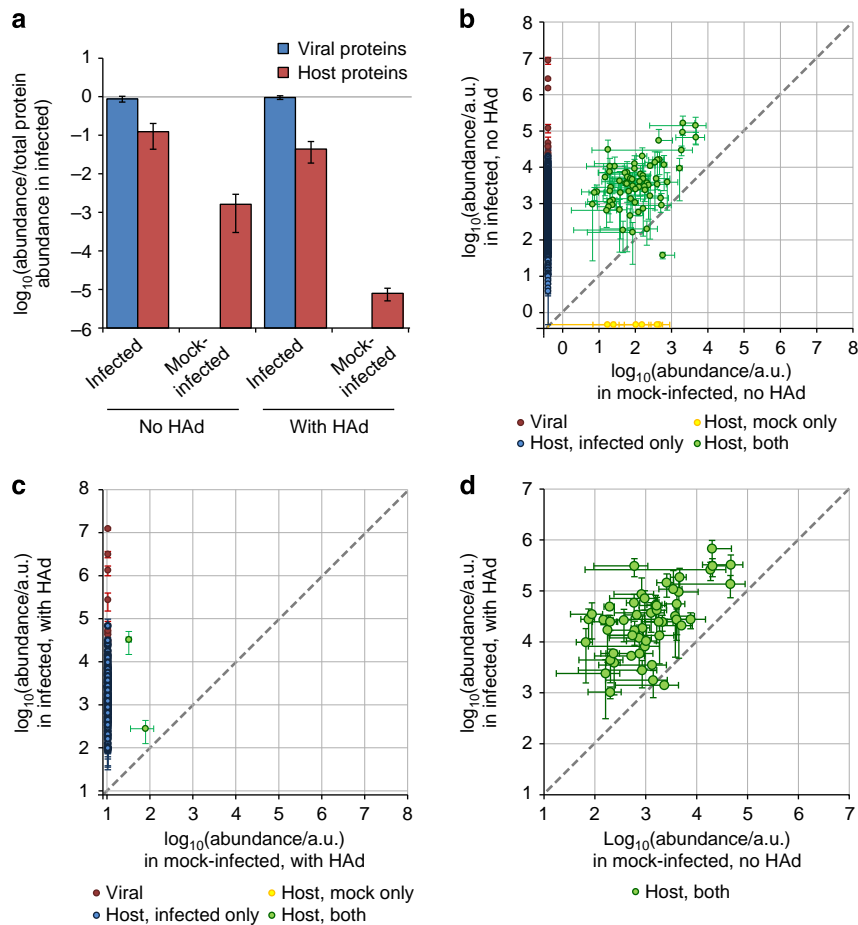


Figure 3 | Comparison of proteins shed by infected and uninfected cells. To compare shedding from infected and mock-infected cells, gradient ultracentrifugation was used to purify virion-sized material from the growth media of bovine epithelial (MDBK) cells 48 h after either infection with WSN or mock infection. Purifications were performed with or without a prior haemadsorption/elution step (HAd), and protein standards were added after purification to allow comparison of samples with different total protein concentrations. Purified proteins were quantified by SINQ. **(a)** Total abundance of purified viral and host proteins, relative to the mean total protein abundance in the infected samples (mean and s.d. of three separate experiments, not including protein standards). **(b–d)** The abundance of individual proteins shed by infected and mock-infected cells, and detected in at least two of three separate experiments (mean and s.d. if $N=3$, mean and range if $N=2$; not including protein standards). Proteins not detected in one condition (infected or mock infected) were assigned the lowest abundance of any protein detected in that condition. Panels show the abundance of proteins **(b)** purified without HAd, **(c)** purified with HAd and **(d)** common to the mock-infected sample without HAd and to the infected sample with HAd.

could be removed by HAd (Fig. 3a,c), this did not remove exosomal markers from purified virions. Indeed, the proteins detected in microvesicles are present in the same ratio in stringently purified virions shed by the same cell type (Fig. 3d), and the greater quantity of material purified in virions allowed the detection of many additional exosomal proteins that we could not detect in the low-abundance mock-infected samples. For all combinations of influenza virus and host, most of the host proteins identified in virions have also been identified in exosomes (Supplementary Data 3–5; the probability of this degree of enrichment occurring by chance, calculated by a hypergeometric test, is less than 0.001)^{39,40}. Influenza virions therefore resemble exosomes both in their hydrodynamic properties and in their protein composition.

Discussion

Influenza viruses, like many other medically important viruses, produce virions with a variable structure which has made it challenging to determine their composition in detail. To address this, we used mass spectrometry to gain detailed quantitative

information about the protein composition of influenza virions, as well as that of the microvesicles shed by uninfected cells. The same approach could be used to characterize many other pleomorphic virions which, for similar reasons, are not amenable to high-resolution structural analysis. More generally, its ability to determine protein ratios in complex mixtures across a large dynamic range, without the need for labelling, could be used in the study of any protein-containing structure as long as it can be purified and concentrated.

Mass spectrometry allowed us to analyse the protein composition of influenza virions with a sensitivity far exceeding that of other approaches^{1–3,6,8–22}. Furthermore, recent increases in instrument sensitivity meant that we could detect around ten times the number of host proteins previously identified in influenza virions by mass spectrometry⁷. By measuring protein abundance, we found that host proteins make a substantial contribution to the structure of influenza virions (Fig. 4). We could also detect proteins whose abundance was orders of magnitude less than the viral polymerase, indicating that we have obtained the first complete description of the protein composition of influenza virions.

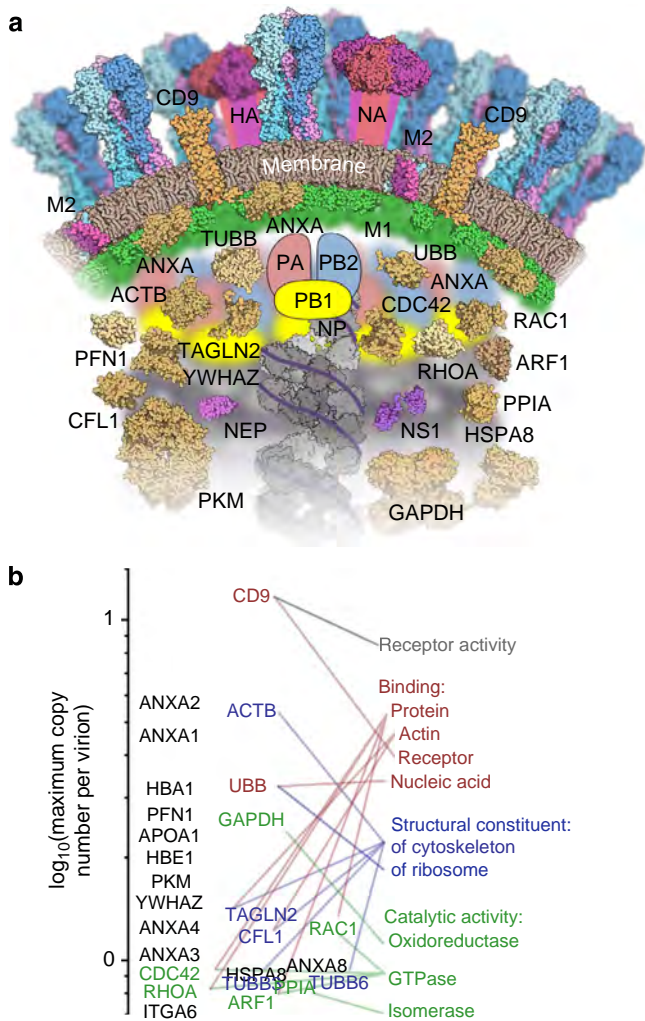


Figure 4 | The architecture of an influenza virion. Proteins in virions produced by WSN-infected bovine epithelial (MDBK) cells, present at more than one-tenth the abundance of the viral polymerase and found in material purified both with and without HA in at least two separate experiments. **(a)** Schematic cross-section of an influenza virion, showing proteins whose localization in the virion is known or can be inferred from other studies. Host proteins and membrane are brown and viral proteins are brightly coloured. **(b)** Host proteins, ranked by maximum copy number in virions (calculated as in Fig. 2) and linked to their gene ontology molecular function terms.

We were surprised to identify the viral NS1 protein as a consistent component of influenza virions. Previous reports that NS1 was non-structural are likely due to its low abundance, and possibly also to its position on polyacrylamide gels being obscured by CD9 and the highly abundant M1 protein, which have a similar electrophoretic mobility to the NS1 protein of some influenza strains^{17,41,42}. NS1 was not detected in a previous LC-MS/MS study⁷ but comparison with our results suggests that previously NS1 was near the lower limit of instrument sensitivity. In the present study, we were able to unambiguously detect NS1 even after stringent purification, and we concluded that it is a low-abundance component of influenza virions. As was previously the case for NS2 (now called NEP^{18,41}), NS1 can no longer formally be considered non-structural. NS1 is a multifunctional protein⁴³, and could potentially play a role in virion assembly or enhance infection when introduced into a new

host cell. Alternatively, as NS1 is present at high levels in the cytoplasm of infected cells^{17,42} it is possible that it is incorporated into virions merely as a passive bystander. It is clear that NS1 is not non-structural, but considering what is currently known about the protein we note that its primary role is as a nucleic acid-binding suppressor of immunity (NSI).

More generally, we identified a consistent underlying architecture for spherical virions of influenza A viruses and an evolutionarily diverged influenza B virus. This evolutionarily conserved architecture includes the viral proteins and a set of host proteins, some of which are highly enriched in virions. Other features of the virion are not fixed, with the greatest variation due to the host. Mammalian tissue cultures and embryonated chicken eggs incorporated distinct sets of proteins into virions, some of them in substantial amounts. In some cases, different proteins were recruited from related families (Fig. 2c); for example, among the tetraspanin superfamily mammalian cells incorporate large amounts of CD9 and smaller amounts of CD81 into virions, whereas eggs incorporate UPK-1B (normally, but not exclusively, found on urothelial cells⁴⁴).

These differences demonstrate that influenza virions derived from different hosts cannot be assumed to be the same. As virions produced by different hosts have typically been treated as equivalent when studying influenza infections, the phenotypic effects of these differences have not yet been explored in detail. The influence of the host on virion formation has been observed in influenza viruses with mutations interfering with correct genome packaging, which suppress virion budding in MDCK cells but do not have the same effect in chicken eggs⁴⁵. Additionally, a tetraspanin whose incorporation we show here to be specific to mammalian hosts has recently been shown to be involved in virion entry and budding⁴⁶. Host-dependent features of virions are of clinical interest as egg-grown virions are used in the preparation of influenza vaccines, and although allergies to egg proteins are not in themselves a contraindication to the administration of influenza vaccines, they do increase the need for precautions⁴⁷. Finally, host-specific differences suggest that virion composition may adapt when moving from avian to mammalian hosts, and hence that the structure of influenza virions may adapt during pandemic emergence.

Spherical influenza virions are of a similar size to exosomes, membrane-bound structures which also transfer protein and RNA between cells^{35,36}. By comparing separately purified exosomes and virions, we show here that they also have a strikingly similar protein profile—by many measures, an influenza virion is simply an exosome that has been enriched with additional components. Similarities have been noted between exosomes and a number of other enveloped viruses^{35,38}, most notably human immunodeficiency virus, for which the ‘Trojan exosome hypothesis’ was proposed to explain virion budding as a subversion of cellular pathways for exosome biogenesis⁴⁸. The correspondence between exosomes and influenza virions is not absolute: influenza uses its M2 protein for budding rather than the endosomal sorting complex required for transport (ESCRT) machinery used by many, though not all, microvesicles^{35,38,49,50}, and while exosomes bud into multivesicular bodies, influenza virions bud from the apical plasma membrane³⁵. On the other hand, influenza is capable of redirecting membrane traffic, for example, by translocating markers of autophagosomes to the plasma membrane⁵¹, and segments of the viral genome accumulate on unidentified structures adjacent to the plasma membrane before virion assembly⁵². The exosome-like features of influenza virions therefore cast light on the viral assembly process, suggesting that the virus commandeers elements of cellular pathways to construct the virions it needs to infect new hosts.

Methods

Cells, antisera and viruses. Mammalian epithelial tissue cultures were maintained in a humidified incubator at 37 °C and 5% CO₂. MDBK cells were obtained from the European Collection of Cell Cultures and grown in minimum essential medium (MEM) (Sigma) supplemented with 2 mM L-glutamine and 10% fetal calf serum (FCS); MDCK cells were obtained from American Type Culture Collection and grown in DMEM (Sigma) supplemented with 10% FCS.

Western blotting was performed using rabbit anti-NS1 at 1/1,000 (a kind gift of Juan Ortin, Centro Nacional de Biotecnología⁵³), rabbit anti-actin at 1/200 (A2066; Sigma) and an antibody against the viral polymerase produced by immunizing rabbits with an influenza A/NT/60/68 virus polymerase, purified as previously described and used at 1/500 (ref. 54). Additional western blotting in Supplementary Fig. 2 used a sheep anti-NS1 at 1/500 (a kind gift of Richard Randall, University of St Andrews⁵⁵), a second rabbit anti-NS1 at 1/500 (a kind gift of Adolfo García-Sastre, Icahn School of Medicine at Mount Sinai) and a rabbit anti-NP at 1/3,000 (a kind gift of Paul Digard, University of Edinburgh⁵⁶). Uncropped images of western blots are shown in Supplementary Fig. 3.

A list of the viruses used is given in Table 1. WSN was grown in MDBK cells in MEM supplemented with 2 mM L-glutamine and 0.5% FCS. PR8 and MUD (a kind gift of Paul Digard, University of Edinburgh) were grown in MDCK cells in DMEM supplemented with 0.14% bovine serum albumin and 1 µg ml⁻¹ bovine pancreatic trypsin (Sigma). Candidate vaccine viruses (CVVs) were a kind gift of Othmar Engelhardt (National Institute of Biological Standards and Controls, UK). NIB-74xp, X-181, X-187 and B/Bris were grown in embryonated chicken eggs; in addition, NIB-74xp was grown in MDCK cells in MEM supplemented with 2 mM L-glutamine, 0.14% bovine serum albumin (Sigma) and 0.75 µg ml⁻¹ bovine pancreatic trypsin (Sigma). Infections were carried out at a low multiplicity and virions harvested around 48 h.p.i.; for tissue culture studies between three and eight near-confluent T175 flasks of cells were infected for each experiment.

WSN and PR8 have undergone extensive laboratory passage and only produce spherical virions (Fig. 1b and Supplementary Fig. 1d). We assumed that the CVVs NIB-74xp, X-181, X-187 and B/Bris would also produce spherical virions as they had been adapted to egg culture³⁰, as the influenza A CVVs are reassortants with PR8 and carry key determinants of spherical morphology on segment 7 (refs 57,58), and as a recent structural study of another egg-grown influenza B virus, B/Lee/40, showed this to be spherical⁵⁹. We also considered influenza PR8 with segment 7 from the A/Udorn/307/1972 H3N2 strain (MUD). MUD is capable of producing filamentous virions, but electron microscopy showed that although our samples contained a small proportion of filamentous virions, the great majority of purified virions were spherical (Supplementary Fig. 1e). Consistent with this, when we compared purified PR8 and MUD virions, the ratios of protein abundance were similar for the majority of proteins (Supplementary Fig. 1f; when ratios of abundance were calculated for the 446 host proteins detected in both PR8 and MUD virions, normalized by the abundance of viral polymerase, the median value was 3.8). Our data therefore describe virions with a spherical morphology.

Purification of virions. Some of the samples analysed here have been described previously²⁵ and new material was purified using the same procedures. Briefly, the growth medium of infected cells was clarified (2,000 g at 30 min then 18,000 g at 30 min, at 4 °C), and material in the supernatant was then concentrated through a cushion of 30% sucrose in NTC (100 mM NaCl, 20 mM Tris-HCl pH 7.4, 5 mM CaCl₂; at 112,000 g for 90 min at 4 °C in an SW 28 rotor (Beckman Coulter)). The pellets were resuspended in NTC and sedimented through a 30–60% sucrose gradient in NTC (209,000 g for 150 min at 4 °C in an SW 41 Ti rotor (Beckman Coulter)) to produce a visible band of virions which was harvested with a needle (except where the entire gradient was harvested in fractions from the top). Mock-infected samples were purified in parallel with infected samples, and an equivalent region of the gradient was harvested. Purified material was pelleted through NTC (154,000 g for 60 min at 4 °C in an SW 41 Ti rotor) and resuspended in a small volume of NTC. For purification using OptiPrep (Sigma) essentially the same procedure was used, but with a cushion of 10% OptiPrep and a gradient of 10–40% OptiPrep. Material was harvested from infected eggs using a similar method which emulates that used to purify starting material for vaccine production²⁵, though without the additional processing steps that vaccine formulations require. Briefly, allantoic fluid was harvested, filtered and mixed with sodium azide. Virions were concentrated by ultracentrifugation, resuspended and sedimented on a 10–40% sucrose gradient to produce a visible band of virions which was harvested and pelleted by ultracentrifugation. For more stringent purification by HAd, cell media were clarified by low-speed centrifugation at 4 °C and then mixed with adult chicken blood (TCS Biosciences) at a final packed cell volume of approximately 0.2%. The suspensions were kept at 4 °C for 30 min, inverting regularly, after which the blood was pelleted (1,370 g for 5 min at 4 °C) and washed twice in 10 ml chilled phosphate-buffered saline. The cell pellet was then resuspended in 37 °C phosphate-buffered saline and incubated at 37 °C for 15 min, inverting regularly. Finally, the cells were pelleted (1,370 g for 5 min at 10 °C) and the supernatant layered onto a 30% sucrose cushion, after which purification proceeded as above.

Digests of virions were performed in NTC supplemented, when bovine pancreatic trypsin was used, with 5 mM MgCl₂ and 1 mM TCEP. To analyse viral contents, virions were purified using essentially the same procedures as above, though with solutions buffered in NTE (100 mM NaCl, 10 mM Tris-HCl pH 7.5,

1 mM EDTA) rather than NTC. Purified virions were lysed at room temperature for 30 min in 100 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 3% Triton X-100, 5% glycerol, 10 mg ml⁻¹ lysolectin and 1.5 mM dithiothreitol, and then layered onto a step gradient of 33, 40, 50 and 70% glycerol in 150 mM NaCl, 50 mM Tris-HCl pH 7.5. The gradient was centrifuged at 245,000 g for 240 min at 4 °C in an SW55 Ti rotor (Beckman Coulter) and then harvested from the bottom of the tube.

Purified virions were analysed by SDS-PAGE and silver staining or western blotting, using standard techniques. In some cases, virions were also visualized by transmission electron microscopy at the Bioimaging Facility of the Sir William Dunn School of Pathology. Virions were fixed in NTC with 2.5% glutaraldehyde and 2% paraformaldehyde, or 0.05% glutaraldehyde and 4% paraformaldehyde, adsorbed onto glow-discharged carbon pioloform or carbon formvar grids, negatively stained with 2% aqueous uranyl acetate and imaged in a Tecnai 12 transmission electron microscope (FEI, Eindhoven) operated at 120 kV.

A full description of the influenza virion purification protocol has been uploaded to *Protocol Exchange* ('Purification of influenza virions by haemadsorption and ultracentrifugation', <http://dx.doi.org/10.1038/protex.2014.027>).

Mass spectrometry. Purified virions were analysed by mass spectrometry as described previously²⁵. Briefly, samples were boiled in Laemmli buffer, purified by running a short distance into polyacrylamide, cut out, reduced, alkylated and digested with trypsin. Peptides were extracted, desalted on a C18 tip and then dissolved in 0.1% formic acid. LC-MS/MS was performed using an Ultimate 3000 RSLCnano HPLC system (Dionex, Camberley, UK) run in direct injection mode and typically coupled to a Q Exactive mass spectrometer (Thermo Electron, Hemel Hempstead, UK) in 'Top 10' data-dependent acquisition mode; an LTQ XL Orbitrap (Thermo Electron, Hemel Hempstead, UK) in 'Top 5' mode was also used for some samples (Table 1). Charge state + 1 ions were rejected from selection and fragmentation and dynamic exclusion with 40 s was enabled; fragmentation was by HCD (Q Exactive) or CID (Orbitrap).

Mass spectra were analysed as previously described²⁵ using the Central Proteomic Facilities Pipeline (CPFP), which uses iProphet to combine searches made with Mascot, OMSSA and X!TANDEM; peptide identifications are validated within CPFP using PeptideProphet. Protein identifications were inferred using ProteinProphet and protein identifications were filtered to a 1% false discovery rate by counting target and decoy sequences²⁴. Peptide spectral matches were made to custom databases that concatenated the predicted proteome of the virus with that of the host as well as to the proteins of the Universal Protein Standards 2 (UPS2) protein dynamic range standard set (Sigma), common contaminants and decoy sequences. The contaminants list was from www.maxquant.org and the host proteomes were the UniProt Reference Proteomes for *Bos taurus*, *Canis familiaris* and *Gallus gallus*; all were downloaded on 12 December 2013. When *Bos taurus* was the host species, bovine proteins were manually removed from the contaminants list. Tryptic peptides with up to two missed cleavages were searched for using a precursor mass tolerance of 20 p.p.m. and a fragment ion tolerance of 0.1 Da (0.5 Da for the LTQ XL Orbitrap), with carbamidomethylation of cysteine as a fixed modification and oxidation of methionine and deamidation of asparagine and glutamine as variable modifications; proteins were identified with a 1% false discovery rate. Normalized spectral abundance factor and SINQ were performed in CPFP, and intensity-based absolute quantitation (iBAQ) in MaxQuant, as previously described²²; for SINQ at least two peptide sequences were required for quantitation. Before comparing data sets, decoy sequences and common contaminants (keratins, trypsin, serum albumin, serum albumin precursor, filaggrin, hornerin, PRSS1 Trypsin-1 precursor and dermokine) were manually removed. Human orthologs of proteins were searched for using the Panther classification system (www.pantherdb.org^{60,61}) and by searching for sequence identity using UniProt Knowledgebase and BLAST (www.uniprot.org⁶²); a list of assigned orthologs is given in Supplementary Data 6. Gene ontology terms were assigned by the Panther classification system^{60,61}.

Testing ratio determination by SINQ. SINQ has previously been tested using LC-MS/MS data obtained with instruments including an LTQ XL Orbitrap²³. We expected that a Q Exactive mass spectrometer would provide more accurate quantitation due to its increased sensitivity and scan speed; to test the ability of SINQ to determine protein ratios when using this instrument, we analysed UPS2 dynamic protein standards (Sigma), digested with trypsin (Supplementary Fig. 1a). UPS2 is a mixture of 48 proteins spanning six orders of magnitude of abundance; at the concentrations used, we were able to detect 24 proteins over four orders of magnitude of abundance. Pairwise comparisons were made of protein abundances determined by SINQ (based on the intensity of fragment ion spectra), intensity-based absolute quantitation (iBAQ; a computationally demanding method based on the intensity of peptide spectra, reported to provide relatively accurate estimates of abundance^{23,63}), and normalized spectral abundance factor (NSAF; a less accurate but computationally less demanding method based on spectral counts⁶⁴). All three methods were effective when comparing proteins in relatively high equimolar amounts, with SINQ determining most ratios within twofold of the correct value, and the other methods determining most ratios within fourfold. Accuracy of all methods diminished for proteins of lower abundance or when

comparing proteins whose amounts differed by orders of magnitude, but SINQ and iBAQ were able to place most ratios with eightfold of the correct value over a 1,000-fold difference in abundance (Supplementary Fig. 1a). We decided to use SINQ for further studies.

To test the effectiveness of SINQ in complex mixtures, we mixed digested protein standards with either a digest of purified virions or of a mock sample purified from the media of uninfected cells (Supplementary Fig. 1b). The additional material was purified using HAd, resulting in a complex mixture of proteins in the virion-containing sample and very few detectable proteins in the mock sample (Fig. 3c). The most abundant protein standards were introduced at levels comparable with the viral polymerase subunits, which are low-abundance components of the virion. Ten standards were detected in the high-complexity (virion-containing) sample and 17 in the low-complexity (mock) sample, over three orders of magnitude of abundance. In both the high-complexity (infected) and low-complexity (uninfected) samples, most equimolar and tenfold ratios were determined within fourfold of the correct value.

Finally, to compare SINQ with an existing method of protein quantification, we measured the relative abundance of the two most highly expressed proteins in purified virions, NP and M1, by SINQ and by densitometry of protein bands separated by SDS-PAGE and stained with Coomassie Brilliant Blue (Supplementary Fig. 1c). The NP/M1 ratios determined by the two methods are within 15% of each other. We also compared ratios of viral proteins determined by SINQ with previously published estimates made using a variety of methods^{1,6,8–22}. In general, our estimates were consistent with previously published values (Supplementary Data 1). The largest discrepancies between SINQ and other methods could be seen when comparing proteins of high and low abundance, which we attribute to the limited dynamic range of the gel electrophoresis and autoradiography methods used in these studies compared with LC-MS/MS. Consistent with this, when the ratio of polymerase to NP was used to predict viral genome size (see Results for details), SINQ gave a much more accurate estimate than previous methods. We would expect glycosylation and transmembrane domains to reduce the efficiency of protein detection by LC-MS/MS. Consistent with this, SINQ appears to underestimate glycoprotein numbers by around fourfold compared with other methods (Supplementary Data 1). There is less data available on the abundance of the M2 ion channel, but estimates of M2 levels by SINQ were around ninefold less than in a previous study. Allowing for these caveats, we determined that SINQ was, within reasonable limits, an effective method for simultaneously quantifying a large number of specific proteins in purified virions across a wide range of abundance.

Estimates of abundance by SINQ are affected by the overall protein concentration of the sample. To control for this when comparing infected and uninfected samples, equal amounts of digested UPS2 dynamic protein standards were mixed with the digested samples immediately before LC-MS/MS. This was done for all HAd samples and for a replicate analysis of one of the non-HAd WSN samples. The protein standards included ubiquitin and peroxiredoxin 1, both of which were already present in virions and could not be used for calibration. The remaining protein standards were used to normalize total protein levels and were then excluded from further analysis.

Data analysis. Principal coordinates analysis was performed using the cmdscale function in R⁶⁵. A list of proteins identified in exosomes was downloaded from www.exocarta.org on 3 December 2014 (refs 39,40). A hypergeometric test was used to calculate the probability that our samples were enriched with these exosomal markers by chance: after removing redundancies, the number of occurrences of markers in the data in Supplementary Data 3–5 (705/907 proteins) was compared with the total number of markers (3,442) and the number of proteins annotated by EBI on 6 December 2014 in the genomes of humans, cows and chickens (20,649, 19,634 and 15,621, respectively). Protein structures were visualized using QuteMol⁶⁶ and Python Molecular Viewer⁶⁷ and composited using Inkscape (www.inkscape.org). Membrane coordinates were from Tieleman *et al.*⁶⁸, the RNP structure was adapted from Hutchinson *et al.*⁶⁹ and the structure of NEP was modelled using Quark⁷⁰; other protein coordinates used were from the Protein Data Bank, with accession codes 1RU7 (HA), 3BEQ (NA), 3LBW (M2), 1EA3 (M1), 2GX9 and 2ZKO (NS1), 1J6Z (ACTB), 1UBI (UBB), 4HNA (TUBB), 2BTF (PFN1), 1Q8G (CFL1), 4HKC (YWHAZ), 1ZJH (PKM), 1WYM (TAGLN2), 3TH5 (RAC1), 4HMY (ARF1), 3T06 (RHOA), 2KB0 (CDC42), 3KOM (PPIA), 4H5N (HSP41) and 3GPD (GAPDH). 1AIN (ANXA1) and 1W7B (ANXA2) were used to represent all annexins, and the modelled structure of CD81 (2AVZ) was used to represent CD9.

References

- Harris, A. *et al.* Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl Acad. Sci. USA* **103**, 19123–19127 (2006).
- Noda, T. *et al.* Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **439**, 490–492 (2006).
- Wasilewski, S., Calder, L. J., Grant, T. & Rosenthal, P. B. Distribution of surface glycoproteins on influenza A virus determined by electron cryotomography. *Vaccine* **30**, 7368–7373 (2012).
- Vijayakrishnan, S. *et al.* Cryotomography of budding influenza A virus reveals filaments with diverse morphologies that mostly do not bear a genome at their distal end. *PLoS Pathog.* **9**, e1003413 (2013).
- Nermt, M. V. Further investigation on the fine structure of influenza virus. *J. Gen. Virol.* **17**, 317–331 (1972).
- Booy, F. P., Ruigrok, R. W. & van Bruggen, E. F. Electron microscopy of influenza virus. A comparison of negatively stained and ice-embedded particles. *J. Mol. Biol.* **184**, 667–676 (1985).
- Shaw, M. L., Stone, K. L., Colangelo, C. M., Gulcicek, E. E. & Palese, P. Cellular proteins in influenza virus particles. *PLoS Pathog.* **4**, e1000085 (2008).
- Compans, R. W., Klenk, H. D., Caligiuri, L. A. & Choppin, P. W. Influenza virus proteins. I. Analysis of polypeptides of the virion and identification of spike glycoproteins. *Virology* **42**, 880–889 (1970).
- Skehel, J. J. & Schild, G. C. The polypeptide composition of influenza A viruses. *Virology* **44**, 396–408 (1971).
- Haslam, E. A., Hampson, A. W., Radishevics, I. & White, D. O. The polypeptides of influenza virus. 3. Identification of the hemagglutinin, neuraminidase and nucleocapsid proteins. *Virology* **42**, 566–575 (1970).
- Fujiyoshi, Y., Kume, N. P., Sakata, K. & Sato, S. B. Fine structure of influenza A virus observed by electron cryo-microscopy. *EMBO J.* **13**, 318–326 (1994).
- Ruigrok, R. W., Calder, L. J. & Wharton, S. A. Electron microscopy of the influenza virus submembranal structure. *Virology* **173**, 311–316 (1989).
- Schulze, I. T. The structure of influenza virus. I. The polypeptides of the virion. *Virology* **42**, 890–904 (1970).
- Schulze, I. T. The structure of influenza virus. II. A model based on the morphology and composition of subviral particles. *Virology* **47**, 181–196 (1972).
- Tiffany, J. M. & Blough, H. A. Estimation of the number of surface projections on myxo- and paramyxoviruses. *Virology* **41**, 392–394 (1970).
- Zebedee, S. L. & Lamb, R. A. Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions. *J. Virol.* **62**, 2762–2772 (1988).
- Inglis, S. C., Carroll, A. R., Lamb, R. A. & Mahy, B. W. Polypeptides specified by the influenza virus genome I. Evidence for eight distinct gene products specified by fowl plague virus. *Virology* **74**, 489–503 (1976).
- Yasuda, J., Nakada, S., Kato, A., Toyoda, T. & Ishihama, A. Molecular assembly of influenza virus: association of the NS2 protein with virion matrix. *Virology* **196**, 249–255 (1993).
- Ruigrok, R. W., Andree, P. J., Hoof van Huysduynen, R. A. & Mellema, J. E. Characterization of three highly purified influenza virus strains by electron microscopy. *J. Gen. Virol.* **65**(Pt 4): 799–802 (1984).
- Cusack, S., Ruigrok, R. W., Krygsmann, P. C. & Mellema, J. E. Structure and composition of influenza virus. A small-angle neutron scattering study. *J. Mol. Biol.* **186**, 565–582 (1985).
- Oxford, J. S., Corcoran, T. & Hugentobler, A. L. Quantitative analysis of the protein composition of influenza A and B viruses using high resolution SDS polyacrylamide gels. *J. Biol. Stand.* **9**, 483–491 (1981).
- Mellema, J. E. *et al.* Structural investigations of influenza B virus. *J. Mol. Biol.* **151**, 329–336 (1981).
- Trudgian, D. C. *et al.* Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. *Proteomics* **11**, 2790–2797 (2011).
- Trudgian, D. C. *et al.* CFP: a central proteomics facilities pipeline. *Bioinformatics* **26**, 1131–1132 (2010).
- Hutchinson, E. C. *et al.* Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. *PLoS Pathog.* **8**, e1002993 (2012).
- Hutchinson, E. C., von Kirchbach, J. C., Gog, J. R. & Digard, P. Genome packaging in influenza A virus. *J. Gen. Virol.* **91**, 313–328 (2010).
- Chou, Y. Y. *et al.* One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc. Natl Acad. Sci. USA* **109**, 9101–9106 (2012).
- Ortega, J. *et al.* Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J. Virol.* **74**, 156–163 (2000).
- Sugita, Y., Sagara, H., Noda, T. & Kawoaka, Y. Configuration of viral ribonucleoprotein complexes within the influenza A virion. *J. Virol.* **87**, 12879–12884 (2013).
- Seladi-Schulman, J., Steel, J. & Lowen, A. C. Spherical influenza viruses have a fitness advantage in embryonated eggs, while filament-producing strains are selected *in vivo*. *J. Virol.* **87**, 13343–13353 (2013).
- Powell, M. L., Napthine, S., Jackson, R. J., Brierley, I. & Brown, T. D. Characterization of the termination-reinitiation strategy employed in the expression of influenza B virus BM2 protein. *RNA* **14**, 2394–2406 (2008).
- Betakova, T., Nermt, M. V. & Hay, A. J. The NB protein is an integral component of the membrane of influenza B virus. *J. Gen. Virol.* **77**(Pt 11): 2689–2694 (1996).
- Brassard, D. L., Leser, G. P. & Lamb, R. A. Influenza B virus NB glycoprotein is a component of the virion. *Virology* **220**, 350–360 (1996).

34. Liu, X. *et al.* Cyclophilin A interacts with influenza A virus M1 protein and impairs the early stage of the viral replication. *Cell. Microbiol.* **11**, 730–741 (2009).
35. Meckes, Jr D. G. & Raab-Traub, N. Microvesicles and viral infection. *J. Virol.* **85**, 12844–12854 (2011).
36. Raposo, G. & Stoorvogel, W. Extracellular vesicles: exosomes, microvesicles, and friends. *J. Cell Biol.* **200**, 373–383 (2013).
37. Vlassov, A. V., Magdaleno, S., Setterquist, R. & Conrad, R. Exosomes: current knowledge of their composition, biological functions, and diagnostic and therapeutic potentials. *Biochim. Biophys. Acta* **1820**, 940–948 (2012).
38. Wurdinger, T. *et al.* Extracellular vesicles and their convergence with viral pathways. *Adv. Virol.* **2012**, 767694 (2012).
39. Mathivanan, S., Fahner, C. J., Reid, G. E. & Simpson, R. J. ExoCarta 2012: database of exosomal proteins, RNA and lipids. *Nucleic Acids Res.* **40**, D1241–D1244 (2012).
40. Simpson, R. J., Kalra, H. & Mathivanan, S. ExoCarta as a resource for exosomal research. *J. Extracell. Vesicles* **1**, 18374 (2012).
41. Richardson, J. C. & Akkina, R. K. NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells. *Arch. Virol.* **116**, 69–80 (1991).
42. Lamb, R. A. & Choppin, P. W. Synthesis of influenza virus proteins in infected cells: translation of viral polypeptides, including three P polypeptides, from RNA produced by primary transcription. *Virology* **74**, 504–519 (1976).
43. Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* **89**, 2359–2376 (2008).
44. Sun, T.-T., Kreibich, G., Pellicer, A., Kong, X.-P. & Wu, X.-R. in *Tetraspanins Proteins and Cell Regulation* (eds Berditchevski, F. & Rubinstein, E.) 299–320 (Springer, 2013).
45. Hutchinson, E. C., Curran, M. D., Read, E. K., Gog, J. R. & Digard, P. Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. *J. Virol.* **82**, 11869–11879 (2008).
46. He, J. *et al.* Dual function of CD81 in influenza virus uncoating and budding. *PLoS Pathog.* **9**, e1003701 (2013).
47. Chung, E. H. Vaccine allergies. *Clin. Exp. Vaccine Res.* **3**, 50–57 (2014).
48. Gould, S. J., Booth, A. M. & Hildreth, J. E. The Trojan exosome hypothesis. *Proc. Natl Acad. Sci. USA* **100**, 10592–10597 (2003).
49. Babst, M. MVB vesicle formation: ESCRT-dependent, ESCRT-independent and everything in between. *Curr. Opin. Cell Biol.* **23**, 452–457 (2011).
50. Rossman, J. S., Jing, X., Leser, G. P. & Lamb, R. A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* **142**, 902–913 (2010).
51. Beale, R. *et al.* A LC3-interacting motif in the influenza A virus M2 protein is required to subvert autophagy and maintain virion stability. *Cell Host Microbe* **15**, 239–247 (2014).
52. Eisfeld, A. J., Kawakami, E., Watanabe, T., Neumann, G. & Kawaoka, Y. RAB11A is essential for transport of the influenza virus genome to the plasma membrane. *J. Virol.* **85**, 6117–6126 (2011).
53. Marion, R. M., Aragon, T., Beloso, A., Nieto, A. & Ortin, J. The N-terminal half of the influenza virus NS1 protein is sufficient for nuclear retention of mRNA and enhancement of viral mRNA translation. *Nucleic Acids Res.* **25**, 4271–4277 (1997).
54. York, A., Hengrung, N., Vreede, F. T., Huisken, J. T. & Fodor, E. Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc. Natl Acad. Sci. USA* **110**, E4238–E4245 (2013).
55. Hale, B. G. *et al.* Structural insights into phosphoinositide 3-kinase activation by the influenza A virus NS1 protein. *Proc. Natl Acad. Sci. USA* **107**, 1954–1959 (2010).
56. Noton, S. L. *et al.* Identification of the domains of the influenza A virus M1 matrix protein required for NP binding, oligomerization and incorporation into virions. *J. Gen. Virol.* **88**, 2280–2290 (2007).
57. Campbell, P. J. *et al.* The M segment of the 2009 pandemic influenza virus confers increased NA activity, filamentous morphology and efficient contact transmissibility to A/Puerto Rico/8/1934-based reassortant viruses. *J. Virol.* **88**, 3802–3814 (2014).
58. Elleman, C. J. & Barclay, W. S. The M1 matrix protein controls the filamentous phenotype of influenza A virus. *Virology* **321**, 144–153 (2004).
59. Katz, G. *et al.* Morphology of influenza b/lee/40 determined by cryo-electron microscopy. *PLoS ONE* **9**, e88288 (2014).
60. Mi, H., Muruganujan, A. & Thomas, P. D. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* **41**, D377–D386 (2013).
61. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* **8**, 1551–1566 (2013).
62. UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **42**, D191–D198 (2014).
63. Schwanhausser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
64. Zybailov, B. *et al.* Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J. Proteome Res.* **5**, 2339–2347 (2006).
65. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).
66. Tarini, M., Cignoni, P. & Montani, C. Ambient occlusion and edge cueing to enhance real time molecular visualization. *IEEE Trans. Vis. Comput. Graph.* **12**, 1237–1244 (2006).
67. Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graph. Model.* **17**, 57–61 (1999).
68. Tieleman, D. & Berendsen, J. Molecular dynamics simulations of fully hydrated DPPC with different macroscopic boundary conditions and parameters. *J. Chem. Phys.* **105**, 4871–4880 (1996).
69. Hutchinson, E. C. & Fodor, E. Transport of the influenza virus genome from nucleus to nucleus. *Viruses* **5**, 2424–2446 (2013).
70. Xu, D. & Zhang, Y. *Ab initio* protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **80**, 1715–1735 (2012).

Acknowledgements

We thank David Brown, George Brownlee, Daniel Lunn and Omer Dushek for helpful discussions, Narin Hengrung for providing reagents and Errin Johnson for technical assistance. The work was supported by the MRC (programme grant MR/K000241/1 to E.F.).

Author contributions

E.C.H. and E.F. designed the experiments and wrote the paper. E.C.H. and M.M.-A. performed experiments. S.S.H. and B.T. analysed samples by mass spectrometry. E.C.H., P.D.C. and B.T. analysed data. D.T. developed the SINQ quantitation method and advised on its use.

Additional information

Accession codes: Raw files for all mass spectra used in this analysis have been deposited at the Mass spectrometry Interactive Virtual Environment (MassIVE; Center for Computational Mass Spectrometry at University of California, San Diego) and can be accessed at <http://massive.ucsd.edu/ProteoSAFe/datasets.jsp> using the MassIVE ID MSV000078740.

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Hutchinson, E. C. *et al.* Conserved and host-specific features of influenza virion architecture. *Nat. Commun.* **5**:4816 doi: 10.1038/ncomms5816 (2014).