

The structure of the influenza A virus genome

Bernadeta Dadonaite^{1,5}, Brad Gilbertson^{2,5}, Michael L. Knight¹, Sanja Trifkovic², Steven Rockman^{2,3}, Alain Laederach⁴, Lorena E. Brown^{2*}, Ervin Fodor^{1*} and David L. V. Bauer^{1*}

Influenza A viruses (IAVs) constitute a major threat to human health. The IAV genome consists of eight single-stranded viral RNA segments contained in separate viral ribonucleoprotein (vRNP) complexes that are packaged together into a single virus particle. The structure of viral RNA is believed to play a role in assembling the different vRNPs into budding virions^{1–8} and in directing reassortment between IAVs⁹. Reassortment between established human IAVs and IAVs harboured in the animal reservoir can lead to the emergence of pandemic influenza strains to which there is little pre-existing immunity in the human population^{10,11}. While previous studies have revealed the overall organization of the proteins within vRNPs, characterization of viral RNA structure using conventional structural methods is hampered by limited resolution and an inability to resolve dynamic components^{12,13}. Here, we employ multiple high-throughput sequencing approaches to generate a global high-resolution structure of the IAV genome. We show that different IAV genome segments acquire distinct RNA conformations and form both intra- and intersegment RNA interactions inside influenza virions. We use our detailed map of IAV genome structure to provide direct evidence for how intersegment RNA interactions drive vRNP cosegregation during reassortment between different IAV strains. The work presented here is a roadmap both for the development of improved vaccine strains and for the creation of a framework to ‘risk assess’ reassortment potential to better predict the emergence of new pandemic influenza strains.

In influenza virions, as well as in infected cells, RNA genome segments are assembled into vRNP complexes where the termini of viral RNA associate with viral RNA-dependent RNA polymerase while the rest of the viral RNA is bound by an oligomeric nucleoprotein^{14,15}. To examine IAV RNA structure at single-nucleotide resolution, we used selective 2'-hydroxyl acylation analysed by primer extension and mutational profiling (SHAPE-MaP)^{16–18}, which probes the conformational flexibility (that is, base pairing) of each nucleotide. We carried out our analysis both in virio and ex virio (Fig. 1a and Supplementary Fig. 1). For the ex virio experiments, the eight viral RNA segments of influenza A/WSN/1933 (H1N1) (WSN) were individually transcribed from PCR products using T7 RNA polymerase (in vitro transcribed RNA) or ‘naked’ viral RNA was purified from deproteinated virus particles (naked viral RNA). For the in virio experiments, viral RNA was probed in the context of vRNPs, directly inside purified WSN virions.

Examination of the in virio SHAPE-MaP profiles revealed that viral RNA in the context of vRNPs is capable of accommodating secondary RNA structures with considerable base pairing, as

evidenced by extensive regions of low SHAPE-MaP reactivities in virio (SHAPE-MaP values below the median; Fig. 1b and Supplementary Table 1). The eight different vRNPs in virio have unique viral RNA conformations, and these structures are consistent and reproducible across three replicates (Supplementary Fig. 2). Comparison with ex virio SHAPE-MaP profiles and SHAPE-MaP-guided predictions of secondary RNA structures show that the presence of nucleoprotein leads to less structurally constrained viral RNA, as evidenced by a global increase in SHAPE-MaP reactivity (Fig. 1c) and decrease in the number of high-probability base-paired structures in virio (Fig. 1d–f). However, some secondary structure remains that is resistant to nucleoprotein, and the 5' end of each segment in virio tends to be more structured overall than the rest of the segment (Supplementary Fig. 3). We used the SHAPE-MaP data to predict local constrained RNA structures (<150 nucleotides (nt); see Methods) and found a substantial number present throughout each segment (for example, Fig. 1f and Supplementary Figs. 4 and 5). In agreement, viral RNA within regions reported to be enriched in nucleoprotein^{19,20} has significantly higher SHAPE-MaP reactivity (Supplementary Fig. 6a). These observations are also consistent with early studies using enzymatic and chemical probing of naked and nucleoprotein-bound short RNAs²¹ and recapitulate RNA structures that have been suggested using computational methods^{22,23} (Supplementary Fig. 5).

As a comparison, we carried out SHAPE-MaP on a closely related IAV strain, A/Puerto Rico/8/34 (H1N1) (PR8), and the more distant A/Udorn/72 (H3N2) (Udorn) strain. We found that the SHAPE-MaP profiles of segments with high nucleotide sequence identity (>90%) preserve a similar RNA conformation to that of WSN; however, segments with lower sequence identity (for example, Udorn PB1, HA and NA) tend to have different RNA conformations (Supplementary Fig. 6b–i), suggesting that the RNA sequence itself is the primary determinant of viral RNA structure, even in the presence of nucleoprotein.

Overall, our results suggest that parts of viral RNA could be exposed and accessible to form intermolecular RNA–RNA interactions between segments. Therefore, we proceeded to analyse such interactions occurring in virio using sequencing of psoralen-cross-linked, ligated and selected hybrids (SPLASH)²⁴, which cross-links base-paired RNAs and maps them using high-throughput sequencing (Supplementary Fig. 7). We performed two biological replicates of SPLASH using purified WSN virions to identify intersegment RNA interactions (Fig. 2a) and could unambiguously identify discrete loci of interaction between segments (Supplementary Fig. 8). The method was highly reproducible between replicates with respect to both sequencing read coverage (Supplementary Fig. 7d,e) and the loci identified (Supplementary Fig. 8c). We also validated our

¹Sir William Dunn School of Pathology, University of Oxford, Oxford, UK. ²Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne, Victoria, Australia. ³Seqirus Ltd, Parkville, Victoria, Australia. ⁴Department of Biology, University of North Carolina, Chapel Hill, NC, USA. ⁵These authors contributed equally: Bernadeta Dadonaite, Brad Gilbertson. *e-mail: lorena@unimelb.edu.au; ervin.fodor@path.ox.ac.uk; david.bauer@path.ox.ac.uk

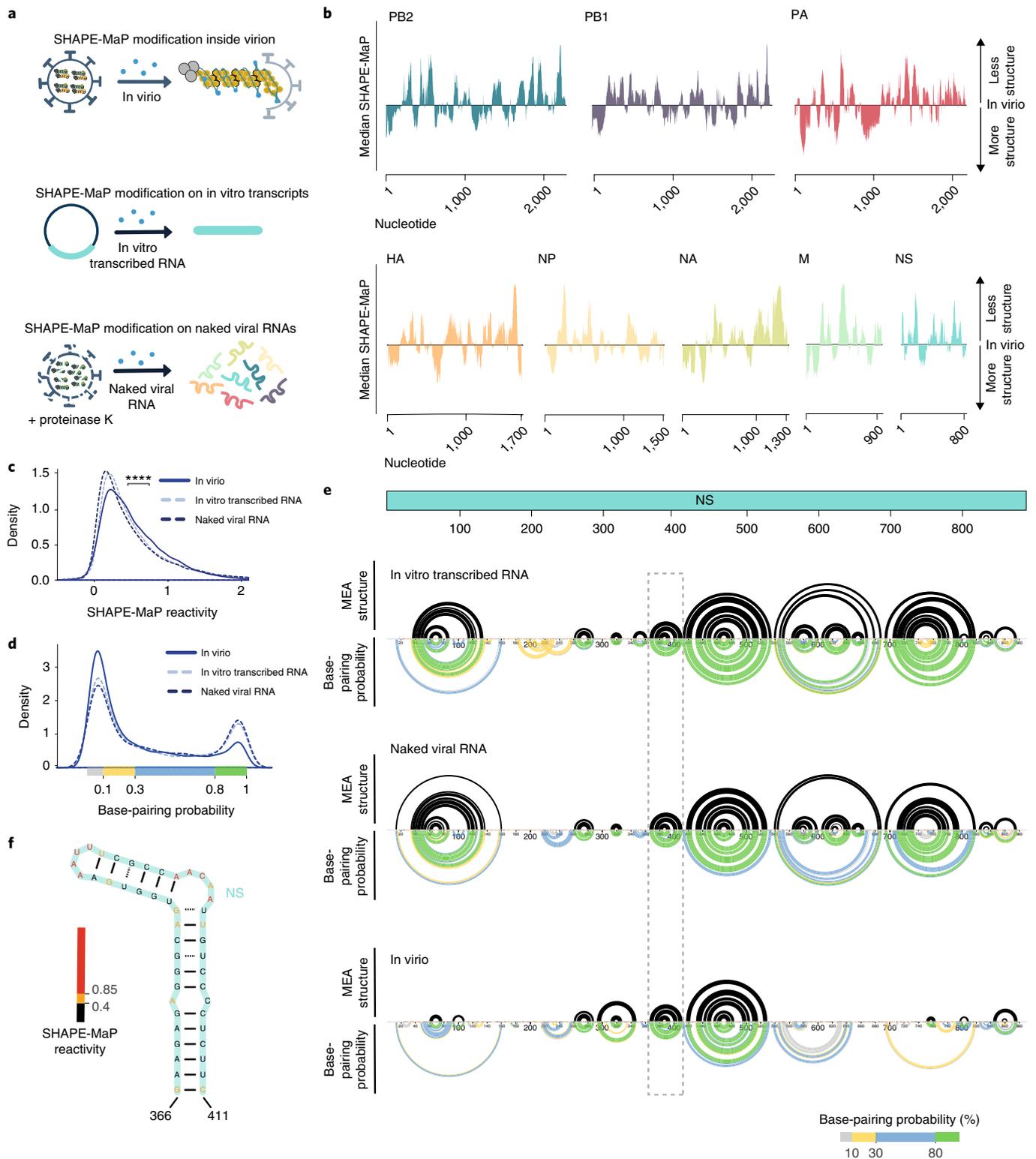


Fig. 1 | Analysis of the IAV genome structure using SHAPE-MaP. **a**, Schematic showing different samples used for SHAPE-MaP analysis. **b**, Median SHAPE-MaP reactivities of different WSN viral RNA segments in virio. Medians were calculated over 50 nucleotide windows and plotted relative to the global median of a given segment. **c**, SHAPE-MaP reactivity distributions in different samples; **** $P \leq 2.2 \times 10^{-16}$, two-sided Wilcoxon rank-sum test, $n = 13,581$ nucleotides per sample. In virio data is an average of three biologically independent samples, in vitro-transcribed RNA data is an average of two biologically independent samples and naked viral RNA data is from a single biological sample. **d**, Base-pairing probability distributions in different viral RNA samples, calculated using the SHAPE-MaP-informed partition function. **e**, Secondary RNA structure of the NS segment. The upper black arcs indicate the maximum expected accuracy (MEA) RNA structure; only interactions associated with >80% base-pairing probabilities are shown. The lower coloured arcs indicate base-pairing probabilities. **f**, RNA hairpin in the non-structural protein (NS) segment resistant to nucleoprotein (NP) indicated by the dashed rectangle in **e**. All sequence positions are annotated as 5'-3' in viral RNA sense. PA, PB1 and PB2 represent three different polymerase (P) subunits. M, matrix protein; NA, neuraminidase; HA, hemagglutinin.

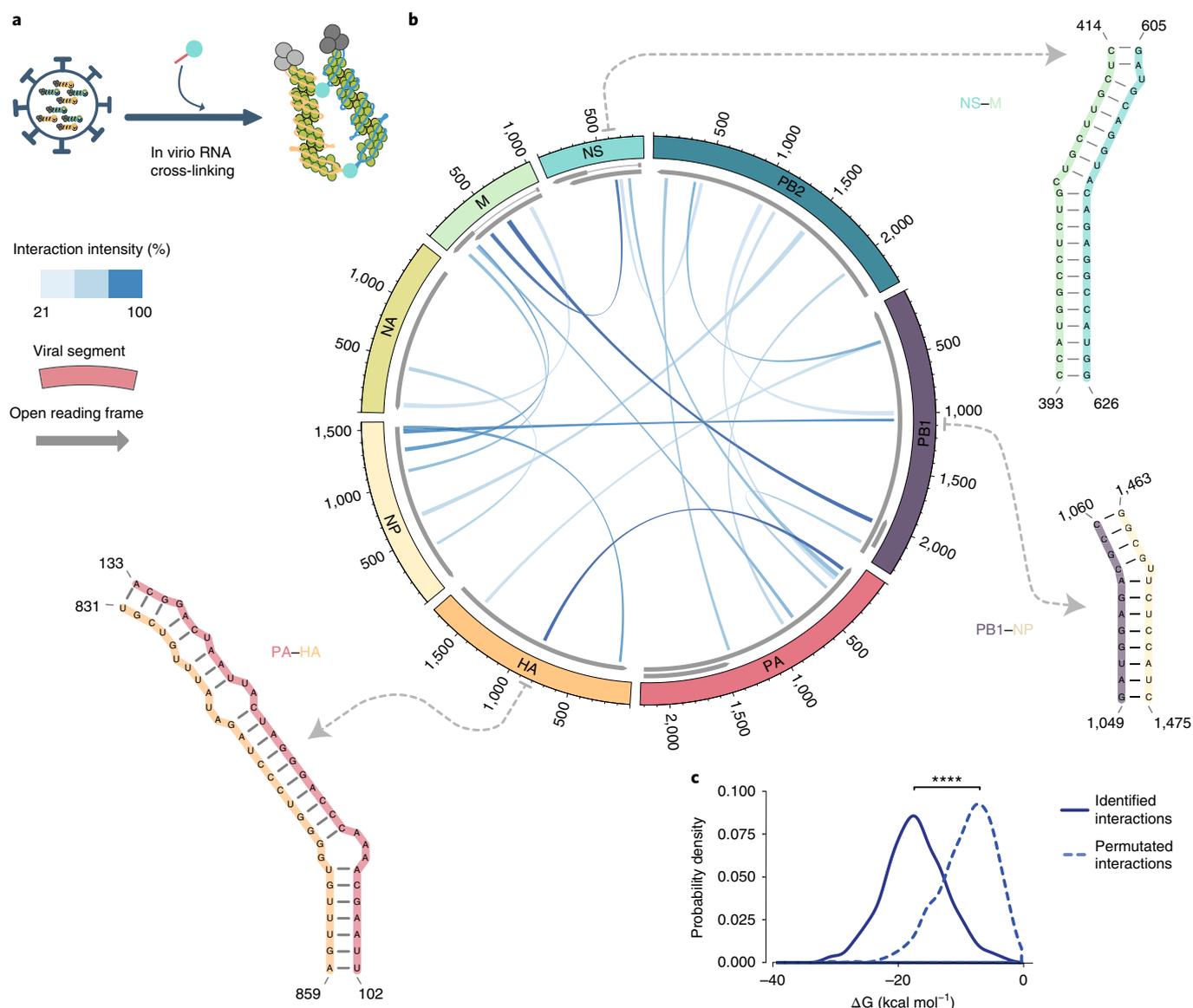


Fig. 2 | Intersegment RNA interactions in the IAV genome. **a**, Schematic summarizing the SPLASH method. **b**, The most prevalent intersegment RNA interactions identified using SPLASH for the WSN (H1N1) IAV strain. The eight viral RNA segments are shown on the perimeter of the circle and the links indicate the regions involved in intersegment base pairing. The links are shaded by interaction frequencies based on the chimeric read contact matrix (see Methods). The dotted arrows point to SHAPE-MaP-guided RNA structure predictions of three representative interactions. **c**, Distribution of the ΔG energies associated with the interactions identified by SPLASH versus a permutated interaction dataset; **** $P \leq 1 \times 10^{-16}$, two-sided Wilcoxon rank-sum test, $n = 611$ interactions in common between two biologically independent experiments.

results using quantitative PCR (qPCR) (Supplementary Fig. 8d–f) and confirmed the presence of intrasegment 5′–3′ promoter interactions (Supplementary Fig. 8g).

Together, the data reveal that an extensive, redundant, complex network of RNA–RNA interactions exists between segments within influenza virions. Instead of a finite set of discrete interactions (for example, what might be imagined from a single set of ‘packaging signals’), there are a large number of possible interactions within a virion population, with some interactions occurring much more frequently than the others. For WSN, we identified 611 interaction loci in common between the 2 replicates, with the top 3% of these loci (18 out of 611 interactions; Fig. 2b) accounting for 25% of the sequencing reads in the dataset.

Importantly, we found that the distribution of the most prevalent interaction loci varies between the eight different viral RNA

segments and that interaction sites are not restricted to certain regions (Fig. 2b). Most segments can interact with multiple other segments; in some cases, the same region can mediate interactions with multiple segments, suggesting that there is a level of redundancy in the network of intermolecular interactions and that multiple RNA conformations exist even within a genetically identical population of virions.

We then predicted the single-nucleotide resolution structures of intersegment interaction loci and also incorporated our empirical SHAPE-MaP reactivity values to constrain these predictions. After benchmarking this approach against structured host RNAs (Supplementary Fig. 9), we calculated the intermolecular free energies (ΔG) of formation of predicted influenza intersegment RNA–RNA structures (for example, in Fig. 2b and Supplementary Table 2) and found they tended to be energetically highly favourable

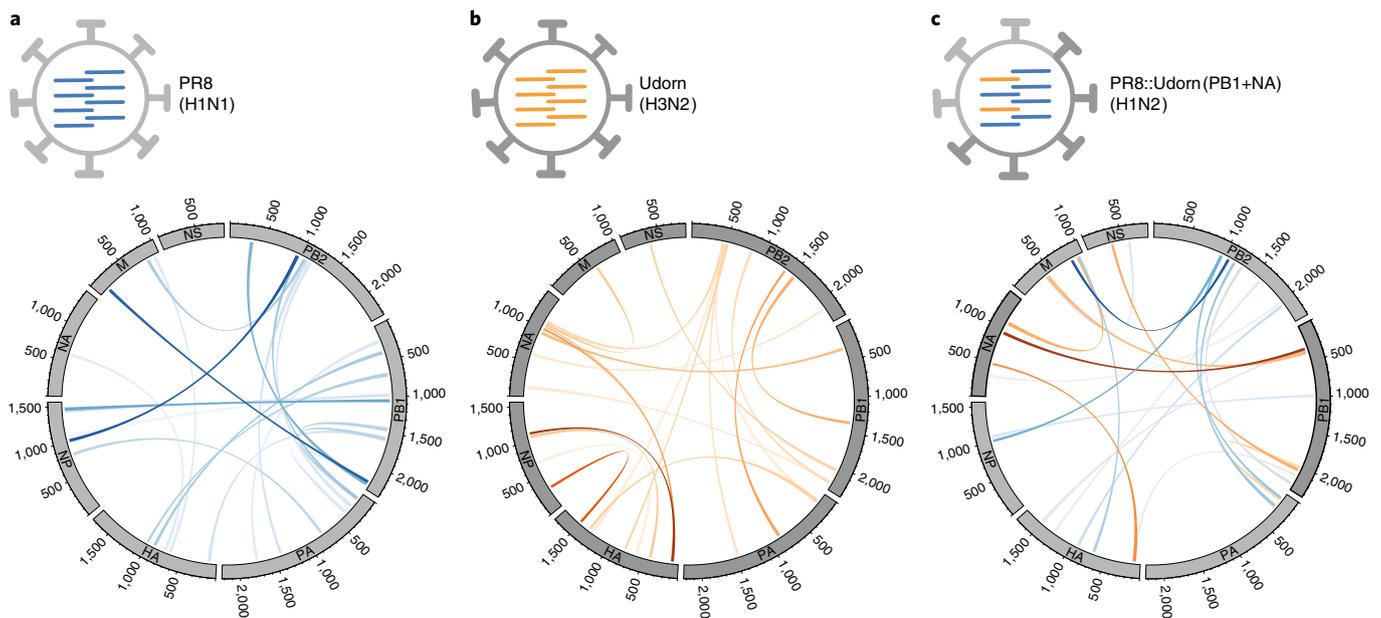


Fig. 3 | RNA interactions form a redundant, plastic network to accommodate variation and reassortment. **a–c**, Intersegment RNA interaction maps for two parent strains, PR8 (H1N1) (**a**), Udorn (H3N2) (**b**), and a reassortant of PR8 that bears the Udorn PB1 and NA gene segments, PR8::Udorn(PB1+NA) (H1N2) (**c**). Interactions are coloured according to the parent strain that donated the interaction and shaded according to their interaction frequency, as indicated earlier (see Fig. 2). Udorn, A/Udorn/307/72.

(median $\Delta G = -18 \text{ kcal mol}^{-1}$; solid line in Fig. 2c). To verify if the interaction loci were specific (versus forming randomly due to confinement within a virion or higher guanine-cytosine content), we permuted the dataset (that is, shuffled the partners of each interaction) and compared the intermolecular free energies (ΔG) of our real dataset to the permuted one: the permuted dataset was predicted to form much weaker interactions (median $\Delta G = -8 \text{ kcal mol}^{-1}$; dashed line in Fig. 2c). This difference in the distributions was highly significant ($P < 1 \times 10^{-16}$, Wilcoxon rank-sum test), confirming that the regions we identified by SPLASH were indeed specific. We also found that RNA within the most prominent interaction loci identified by SPLASH was overall more structured than RNA in the rest of the genome, according to our SHAPE-MaP data ($P = 5.93 \times 10^{-5}$; Supplementary Fig. 8h), although it is important to emphasize that substantial RNA structure is still present outside of SPLASH regions (for example, hairpins and pseudoknots; Supplementary Figs. 4 and 5).

We then carried out SPLASH analysis on the closely related PR8 strain (96% sequence identity to WSN) and compared the RNA–RNA interaction network of PR8 to that of WSN (Supplementary Fig. 10a). We found that the core of the network is broadly similar to that of WSN, but the prevalence of the interactions is prone to change, suggesting the structure is highly plastic in response to changes in nucleotide sequence. One-third of the top 2% of interaction loci in PR8 were also highly prevalent in WSN (in the top 10%) and another third of the top PR8 loci were present in WSN as minor interactions (below the top 10%), while the remaining third of interactions were unique to PR8. We also carried out SPLASH analysis on the more distant Udorn strain and found that it forms a much more distinct network of interactions (Supplementary Fig. 10a), with only 3 out of 20 of the top interaction loci in common with WSN or PR8. To examine the effects of reassortment on the RNA–RNA interaction network, we then examined a reassortant virus containing the PB1 and NA gene segments from Udorn and the remaining 6 segments from PR8, PR8::Udorn(PB1+NA), as well as its specific parent Udorn and PR8 strains, which had >99%

sequence identity to the Udorn and PR8 strains we examined earlier (Fig. 3 and Supplementary Table 2). We found that the reassortant inherited its interaction network from both parent strains (Supplementary Fig. 10b), with some previously minor interaction loci rising substantially in prevalence to accommodate the reassorted segments. This suggests that the plasticity of the RNA interaction network allows the influenza virus to assemble new gene constellations (as observed during antigenic shift) and accommodate small sequence variations (as observed in antigenic drift, for example, between WSN and PR8).

While analysing the networks of the Udorn virus and its PR8 reassortant, we noticed that one of the most prominent interaction loci forms between the H3N2-origin PB1 and NA segments (Fig. 4a). Interestingly, twice in the past century, an avian PB1 segment has reassorted with an N2-NA segment during the generation of a pandemic influenza virus: first in 1957, leading to the Asian influenza pandemic (which arose from the reassortment of PB1, NA and HA segments from an avian H2N2 strain with other segments from the then-seasonal H1N1 virus); then again in 1968, leading to the H3N2 Hong Kong pandemic (when the seasonal H2N2 virus again acquired PB1 and HA segments from an avian source, perhaps as a result of an interaction between the seasonal N2-NA and the avian PB1)¹¹. In addition, our previous studies examining seasonal H3N2 vaccine seed viruses produced using classical reassortment methods between egg-adapted PR8 and seasonal H3N2 viruses have noted that the PB1 and NA segments from H3N2 strains tend to cosegregate^{9,25,26}. We showed that the region in the PB1 segment responsible for cosegregation was within 272–566 nt (1,776–2,070 nt in 3′–5′ viral RNA coordinates⁹), precisely encompassing the prominent RNA–RNA interaction identified using SPLASH in this study (305–338 nt of PB1; Fig. 4a and Supplementary Fig. 11a,b) and suggesting this intersegment interaction could drive the observed cosegregation patterns during reassortment.

To test this hypothesis, we performed competitive reverse-engineering of viruses, where plasmids encoding six viral genome segments of the PR8 virus were transfected together with a plasmid

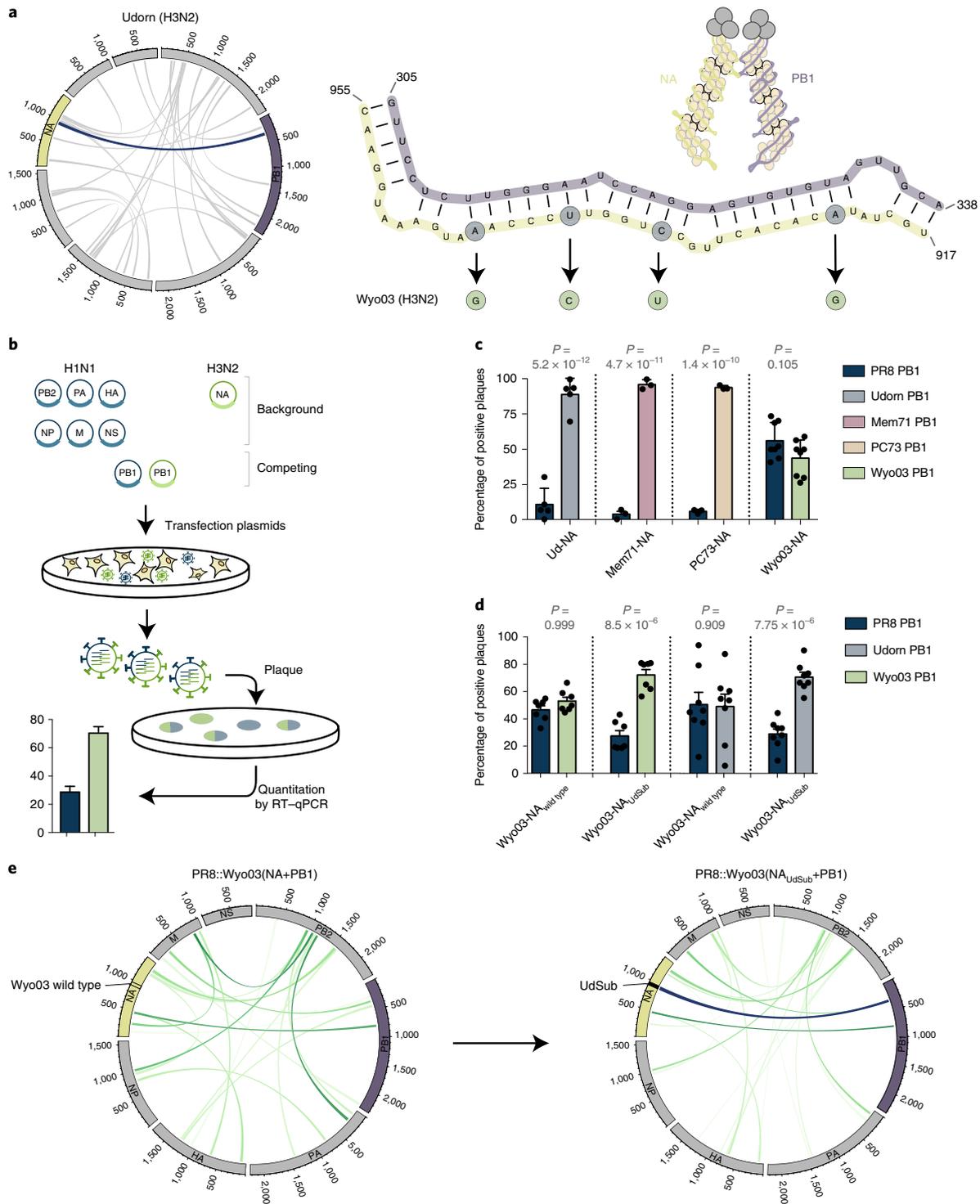


Fig. 4 | Intersegment RNA interactions drive IAV segment cosegregation during reassortment. a, Intersegment RNA interaction map for the Udorn virus, with the interaction between the PB1 and NA segments highlighted in dark blue. Structure prediction for the highlighted PB1-NA interaction is shown on the right. The circled nucleotides highlight the bases that differ between the Udorn and Wyo03 strains. **b**, Competitive reverse-engineering of influenza viruses. Six plasmids encoding H1N1 background segments are transfected together with an H3N2 NA segment-encoding plasmid and PB1 segment-encoding plasmids from both the H1N1 and H3N2 strains. The origin of the PB1 segment in the progeny viruses is determined using RT-qPCR. **c**, Cosegregation between H1N1 or H3N2 PB1 segments and H3N2 NA segments. P values as indicated; analysis of variance (ANOVA) with Sidak correction for multiple testing; $n = 5$ (Ud-NA), $n = 3$ (Mem71-NA and PC73-NA) and $n = 8$ (Wyo03-NA) biologically independent experiments. The centre of the bar plot represents the mean; the error bars indicate the s.e.m. **d**, Preferential Wyo03 PB1 and NA segment cosegregation is recovered by substituting the four nucleotides that differ in the PB1-interacting region of Udorn-NA (highlighted in **a**) for an Udorn-like sequence. P values as indicated; ANOVA with Sidak correction for multiple testing; $n = 7$ (PR8 PB1 versus Wyo03 PB1 competitions) and $n = 8$ (PR8 PB1 versus Udorn PB1 competitions) biologically independent experiments. The centre of the bar plot represents the mean; the error bars indicate the s.e.m. **e**, Substitution of the four Udorn-like nucleotides regenerates a strong intersegment RNA interaction between the H3N2-origin PB1 and NA segments in reassortant viruses.

encoding the Udorn-NA segment and plasmids encoding competing PB1 segments from the PR8 and Udorn viruses (Fig. 4b). The origin of the incorporated competing segment in the resulting reassortant virus was determined using qPCR. Potential detrimental effects of protein incompatibility in the possible reassortant progeny were ruled out in experiments demonstrating their equivalent replication kinetics (Supplementary Fig. 11c). As shown previously^{25,26}, a virus possessing the N2-NA gene segment originating from the Udorn virus preferentially incorporates the Udorn PB1 segment; we observed the same cosegregation pattern with other seasonal H3N2 strains (A/Memphis/1/71 (Mem71) and A/Port Chalmers/1/73 (PC73)) (Fig. 4c). In contrast, the NA segment from another H3N2 virus, A/Wyoming/3/03 (Wyo03), did not drive incorporation of the Wyo03 PB1 segment in the progeny viruses over the incorporation of the PR8 PB1 segment (Fig. 4c), in agreement with the gene constellations of vaccine seed viruses derived from this strain²⁵. Intriguingly, the Udorn-NA segment could drive coselection of the Wyo03-PB1 segment, implying that the Wyo03-NA segment (versus the Wyo03-PB1 segment) is responsible for the lack of NA-PB1 cosegregation in the Wyo03 strain (Supplementary Fig. 11d,e). We noted that the sequence of the PB1-interacting region of NA identified in our SPLASH data (917–955 nt) is conserved in Udorn, Mem71 and PC73, whereas there are 4 single nucleotide changes in Wyo03 (Fig. 4a and Supplementary Fig. 11f). Therefore, we generated a Udorn-like Wyo03-NA gene (Wyo03-NA_{UdSub}) to restore complementarity and found that introducing these 4 nucleotide mutations into the Wyo03-NA segment was sufficient to bias the NA-PB1 interaction towards generation of PR8 reassortants with both Wyo03-derived NA and PB1 segments (Fig. 4d). SPLASH analysis of the reassortants confirmed the regeneration of the prominent ‘Udorn-like’ PB1-NA interaction at the site of mutation (Fig. 4e). Together, these results confirm the importance of RNA–RNA interactions in driving segment cosegregation during reassortment and show that information on RNA structure can predict the ability of segments from different viral strains to reassort.

Overall, our study presents a global high-resolution structure of the IAV genome. We show that different IAV genome segments in virions have distinct RNA conformations, despite the presence of nucleoprotein, and form both intra- and intersegment RNA interactions. This suggests that both nucleotide sequence changes of the RNA itself, as well as amino acid changes to the nucleoprotein that alter its affinity to bind and restructure RNA (for example, a ‘nucleoprotein code’²⁷) could affect this network, either directly at the interaction site, or indirectly by affecting which regions of viral RNA along a vRNP are rendered accessible. It has been suggested that intersegment interactions are mediated by the termini, supported by work showing that the minimal sequence for efficient incorporation of a given segment into virions includes approximately 50–150 nt of each end², and that defective interfering RNAs maintain on average approximately 200 nt of each end²⁸. Nevertheless, studies examining intersegment interactions directly have shown that they can occur outside the termini^{3,9,29}. Our work supports the latter model and also reveals that IAV has evolved a degree of flexibility in its genome structure, with multiple redundant loci involved in the assembly of eight vRNPs into virions. Such a redundancy would explain the sufficiency of terminal sequences for segment incorporation and might act as an evolutionary strategy to accommodate changes in genome sequence caused by genetic drift or changing evolutionary pressures. Crucially, the redundancy in vRNP–vRNP interactions would accommodate the need for selective packaging of the eight vRNPs during infection while also allowing reassortment to occur in a co-infection event, thus providing the influenza virus with a mechanism to escape established immunity in a particular host. Further exploration to generate a comprehensive understanding of the formation and dynamics of intersegment RNA interactions in

influenza viruses will enable us to better understand the gene constellations that may result from reassortment between a given set of strains, guiding vaccine design and risk assessment of potential pandemic influenza viruses.

Methods

Cell culture, virus amplification and purification for SHAPE-MaP and SPLASH experiments. Madin–Darby Bovine Kidney (MDBK) and Madin–Darby canine kidney (MDCK) cells were grown in Minimum Essential Medium (Merck), supplemented with 2 mM L-glutamine and 10% FCS. Stocks of WSN (H1N1) virus were produced by infecting MDBK cells with influenza virus at a multiplicity of infection (MOI) of 0.01. Viral stocks of Udorn (H3N2) and PR8 (H1N1) (PR8) viruses were produced by infecting MDCK cells at an MOI of 0.001 in the presence of 0.8 µg ml⁻¹ of *N*-Tosyl-L-phenylalanine chloromethyl ketone (TPCK)-treated trypsin (Merck). Viruses were collected 2 d post-infection. Viruses were purified by ultracentrifugation: first, the infected cell culture medium was clarified by centrifugation at 4,000 r.p.m. for 10 min at 4 °C followed by centrifugation at 10,000 r.p.m. for 15 min at 4 °C. The virus was then purified by centrifugation through a 30% sucrose cushion at 25,000 r.p.m. for 90 min at 4 °C in a SW 32 rotor (Beckman Coulter). The purified virus pellet was resuspended in a resuspension buffer (0.01 M Tris-HCl, pH 7.4, 0.1 M NaCl, 0.0001 M EDTA). We note that neither virion nor RNA tertiary structures are disrupted by ultracentrifugation^{30–32}.

SHAPE-MaP. SHAPE-MaP was performed according to published protocols¹⁷; 1-methyl-7-nitroisatoic anhydride (1M7) was synthesized from 4-nitroisatoic anhydride as described previously³³. For the in vitro transcribed RNA experiments, each viral RNA segment was synthesized from a linear DNA template using the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs). The products were checked for size and purity on a 3.5% polyacrylamide gel electrophoresis (PAGE)-urea gel. Naked viral RNA samples were prepared by purifying the WSN particles over sucrose cushion as described earlier. Purified viruses were treated with 250 µg ml⁻¹ of Proteinase K (Roche) in proteinase K buffer (10 mM Tris-HCl, pH 7.0, 100 mM NaCl, 1 mM EDTA, 0.5% SDS) for 40 min at 37 °C. Before the modification, in vitro transcribed RNA and naked viral RNA samples were folded at 37 °C for 30 min in folding buffer (100 mM HEPES-NaOH, pH 8.0, 100 mM NaCl, 10 mM MgCl₂). 1M7 (dissolved in anhydrous dimethylsulfoxide (DMSO; Merck)) was added to a final concentration of 10 mM to the folded RNA and the samples were incubated for 75 s at 37 °C. The in vitro modifications were performed by adding 1M7 directly to purified virus. The ability of SHAPE-MaP reagents to penetrate viral particles was initially tested as described previously³⁴ by performing ³²P-labelled primer extensions on RNA extracted from SHAPE-MaP reagent-treated virus using an NA segment-targeting primer (5'-AATTGGTCCAAAGGAGACG-3'). In parallel to the 1M7-treated samples, control samples were treated with DMSO. *n*-methylisatoic anhydride (NMIA; Thermo Fisher Scientific) SHAPE-MaP reagent was also tested in virio. Experiments with NMIA were performed as described for 1M7, except the purified virions were treated with NMIA for 45 min.

Sequencing library preparation was carried out as described previously¹⁷ according to the randomer workflow. In brief, after 1M7 or control treatment, RNA was purified using the RNA Clean & Concentrator-5 Kit (Zymo Research). The RNA was reverse-transcribed using the Random Primer Mix (New England Biolabs) with SuperScript II (Invitrogen) in MaP buffer (50 mM Tris-HCl (pH 8.0), 75 mM KCl, 6 mM MnCl₂, 10 mM dithiothreitol and 0.5 mM deoxynucleoside triphosphate). The Nextera XT DNA Library Prep Kit (Illumina) was used to prepare the DNA libraries. Final PCR amplification products were size-selected using Agencourt AMPure XP Beads (Beckman Coulter) and quality-assessed with the Agilent DNA 1000 kit (Agilent Technologies) on a Bioanalyzer 2100 System (Agilent Technologies). For WSN, naked viral RNA and in vitro transcribed RNA, the libraries were sequenced (2 × 150 base pairs (bp)) on a HiSeq 4000 System (Illumina); for PR8 and Udorn viruses, the libraries were sequenced (1 × 150 bp) on a NextSeq 500 System (Illumina).

SPLASH. SPLASH samples were prepared as published previously^{24,35}, with some modifications, for two replicates each of WSN, PR8 and Udorn viruses and a single replicate for each of the H3N2 reassortant viruses. Purified virus was incubated with 200 µM of EZ-Link Psoralen-PEG3-Biotin (Thermo Fisher Scientific) and 0.01% digitonin (Merck) for 5 min at 37 °C. Virus was spread on a 6-well dish, covered with a glass plate, placed on ice and irradiated for 45 min using a UVP Ultra Violet Product Handheld UV Lamp (Thermo Fisher Scientific). Crosslinked virus was treated with proteinase K and viral RNA was extracted using TRIzol (Invitrogen). An aliquot of extracted viral RNA was used to detect biotin incorporation using a Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Fisher Scientific) on a Hybond-N Nylon Membrane (GE Healthcare Life Sciences). The rest of the extracted viral RNA was fragmented using the NEBNext Magnesium RNA Fragmentation Module (New England Biolabs) and size-selected for fragments shorter than 200 nt using the RNA Clean & Concentrator-5 Kit. The samples were enriched for biotinylated viral RNA using Dynabeads MyOne Streptavidin C1 beads (Thermo Fisher Scientific); on-bead proximity ligation

and psoralen-cross-link reversal were carried out as published previously^{24,35}. Sequencing libraries were prepared using the commercial SMARTer smRNA-Seq Kit (Clontech Laboratories). Final size selection was done by running the PCR-amplified sequencing libraries on a 6% PAGE gel (Thermo Fisher Scientific) in Tris/boric acid/EDTA (TBE), selecting for 200–300 bp DNA. Libraries were sequenced 1 × 150 bp on a NextSeq 500 System.

Processing of SHAPE-MaP sequencing reads. Sequencing reads were trimmed to remove adaptors using Skewer v0.2.2 (ref. ³⁶). The SHAPE-MaP reactivity profiles were generated using the published ShapeMapper2 pipeline³⁷, which aligns the reads to the reference genome and calculates mutation rates at each nucleotide position. The mutation rates are then converted to the SHAPE-MaP reactivity values defined as:

$$R = \text{mutr}_{1M7} - \text{mutr}_{\text{DMSO}}$$

where mutr_{1M7} is the nucleotide mutation rate in the 1M7-treated sample and $\text{mutr}_{\text{DMSO}}$ is the mutation rate in the DMSO-treated sample. All SHAPE-MaP reactivities were normalized to an approximate 0–2 scale by dividing the SHAPE-MaP reactivity values by the average reactivity of the 10% most highly reactive nucleotides after excluding outliers (defined as nucleotides with reactivity values that are >1.5 the interquartile range). High SHAPE-MaP reactivities indicate more flexible (that is, single-stranded) regions of RNA and low SHAPE-MaP reactivities indicate more structurally constrained (that is, base-paired) regions of RNA.

Processing of SPLASH sequencing reads. Sequencing reads were trimmed to remove adaptors using Skewer v0.2.2 (ref. ³⁶). STAR v2.5.3 (ref. ³⁸) was used to align the reads to the appropriate virus reference genome (Supplementary Table 2). Only the chimeric reads where at least 20 nt aligned to the reference segments were used in further processing (STAR parameter: `-chimSegmentMin 20`). Chimeric reads were deduplicated using CIGAR strings and alignment positions. CIGAR strings in each read alignment were processed to find the read start and end coordinates. Chimeric read coordinates were used to produce a matrix of sequence interactions in the R software. Discrete loci in the matrix were selected and individually fitted with a Gaussian curve based on read overlap intensity to define an interaction window; interaction windows of complex overlapping loci were separated into individual windows. The width of the interaction window was used to determine the start and end coordinates of each interaction; the number of reads that were within (or partially within) this region was used as a measure of the interaction frequency. For the generation of figures, the top 20 interactions in each virus were visualized using the circlize package v0.4.5 (ref. ³⁹) in R v3.5.1. The full set of interaction loci is provided in Supplementary Table 2. For qPCR validation of interacting loci, psoralen-cross-linked samples were prepared and enriched as described earlier, but with a shortened fragmentation time (3 versus 4 min) to generate longer RNA fragments. RNA was polyadenylated with Poly(A) Polymerase (Takara Bio) and complementary DNA was generated using the smRNA dT Primer (Takara Bio) and PrimeScript Reverse Transcriptase (Takara Bio) according to the manufacturer's instructions. A 50-cycle qPCR was carried out according to the manufacturer's instructions on a StepOnePlus instrument (Applied Biosystems) using the Brilliant II SYBR Green qPCR Master Mix with ROX (Agilent Technologies) and primer pairs to test for intersgment interactions. Primer sequences are given in Supplementary Table 3. Following 50 cycles of qPCR amplification, products were resolved by 8% PAGE (29:1 acrylamide/bis-acrylamide, 1× TBE buffer) and visualized with blue light transillumination.

RNA structure predictions. The IntaRNA algorithm v2.3.1 (ref. ⁴⁰) was used to predict the ability of RNA–RNA interactions to occur in the regions identified during the SPLASH analysis, using the Exact mode (–mode E) and no seed constraint (–noSeed) options. SHAPE-MaP reactivities were included in the modelling of RNA–RNA interactions (–tShape and –qShape). Permutated datasets were generated by randomly shuffling the specific interaction partners identified by SPLASH and assessing the interaction ΔG energies using IntaRNA. The significance of the difference between the probability distributions of the ΔG energies associated with the SPLASH-identified intermolecular RNA interactions versus the permutated datasets was calculated using the Wilcoxon rank-sum test in the R software. The IntaRNA structure predictions were then used to trim the interaction regions to the nucleotides involved in the base pairing. Where SHAPE-MaP data were not available (PR8 reassortants with H3N2 viruses), pre-folding of each RNA strand ('accessibility') was disabled in IntaRNA (–qAcc=N –tAcc=N). For validation against known structures, RNA secondary structure data were extracted from the RNA3Dhub database⁴¹, based on the cryogenic electron microscopy structures of the 80S ribosome⁴² (PDB: 6EK0) and from the U4/U6.U5 triple small nuclear ribonucleoprotein spliceosomal complex⁴³ (EMDB: EMD-2966). Reference RNA sequences were corrected to match bovine (MDBK) sequences for ribosomal RNA and U4/U6 small nuclear RNAs as described previously⁴⁴. For intramolecular RNA structure predictions, the ViennaRNA package v2.0 (ref. ⁴⁵) was used. The RNAfold command was used to predict secondary RNA structures and partition functions for each segment. SHAPE-MaP reactivities were included as pseudoenergy restraints. A 50 nt sliding median window correlation analysis between the ex virio and in virio SHAPE-MaP reactivity profiles was used to

determine the extent of SHAPE-MaP correlation between T7-transcribed and vRNP-associated RNA. We found that no correlation existed >150 nt; therefore, we set the maximum pairing distance constraint for structure and partition function predictions to 150 nt. For intramolecular structure predictions, we set the nucleotides within the promoter region to be single-stranded.

Cell culture for reverse-engineering of influenza viruses. MDCK cells and human embryonic kidney (HEK 293T) cells were sourced from an existing collection in the Department of Microbiology and Immunology, University of Melbourne. MDCK cells were grown in Roswell Park Memorial Institute (RPMI) 1640 medium (Sigma-Aldrich) and HEK 293T cells were grown in DMEM (Thermo Fisher Scientific). Both media were supplemented with 10% heat-inactivated FCS, 2 mM L-glutamine, 2 mM sodium pyruvate, 24 $\mu\text{g ml}^{-1}$ gentamicin, 50 $\mu\text{g ml}^{-1}$ streptomycin and 50 IU ml^{-1} penicillin. Cocultures of MDCK cells and HEK 293T cells for transfection were established in Opti-MEM (Thermo Fisher Scientific) with 50 $\mu\text{g ml}^{-1}$ streptomycin and 50 IU ml^{-1} penicillin.

Construction of reverse-engineered influenza viruses. Individual gene segments from PR8, Udorn, Mem71 (H3N2), PC73 (H3N2) and Wyo03 (H3N2) viruses were reverse-transcribed and cloned into pHW2000 plasmids⁴⁶. The reverse genetics-derived viruses contained either a PR8, Udorn, Mem71, PC73 or Wyo03 PB1 gene with either a wild-type or modified NA gene in a genetic background comprising six segments from the PR8 virus (PB2, PA, HA, NP, M and NS). The modified NA gene (Wyo03_{Udasub}) was derived from Wyo03 and carried 4 substitutions towards the Udorn-NA sequence: nucleotides G943A; C938U; U933C; and G923A. Three of these four nucleotide changes were silent, with the fourth (A534G) resulting in a conservative lysine to arginine change (K172R). Complementary fragments incorporating these changes were generated by PCR and joined together by another round of PCR using segment-specific primers containing BsmBI restriction sites⁴⁷; the product was cloned into the pHW2000 expression vector for virus rescue. Primer sequences are given in Supplementary Table 3. Rescued viruses were amplified in 10-d embryonated hen's eggs. Infectious allantoic fluid was titrated for virus content by plaque formation in MDCK cells⁴⁸ and stored at -80°C .

Determination of viral replication kinetics of reverse-engineered viruses. The replication characteristics of the reverse-engineered viruses were determined by infecting MDCK cells at an MOI of 0.01. Following 1 h adsorption (at $t=0$ h) the inoculum was removed and cells were washed and incubated at 37°C , 5% CO_2 in RPMI 1640 supplemented with 2 mM L-glutamine, 2 mM sodium pyruvate, 24 $\mu\text{g ml}^{-1}$ gentamicin, 50 $\mu\text{g ml}^{-1}$ streptomycin, 50 IU ml^{-1} penicillin and 1 $\mu\text{g ml}^{-1}$ of TPCK-treated trypsin (Worthington Biochemical Corporation). Cell culture supernatants were collected at various time points post-infection and stored at -80°C for analysis. Viral titres were determined by plaque formation on confluent monolayers of MDCK cells.

Nine-plasmid competitive reverse-engineering of influenza viruses. Competitive reverse-engineering of influenza viruses was undertaken using a modified version of the eight-plasmid reverse genetics system²⁶ as described previously²⁵. Briefly, plasmids encoding the PB2, PB1, PA, hemagglutinin (HA), nucleoprotein (NP), matrix protein (M) and non-structural protein (NS) viral RNA segments of PR8 were cotransfected with plasmids encoding neuraminidase (NA) viral RNA and competing PB1 viral RNA of either Udorn, Mem71, PC73 or wild-type or modified Wyo03. Each plasmid (1 μg) was mixed with FuGENE 6 transfection reagent (Promega) in Opti-MEM and added to a coculture of HEK 293T and MDCK cells. Six hours post-transfection, media was replaced with Opti-MEM supplemented with 50 $\mu\text{g ml}^{-1}$ streptomycin and 50 IU ml^{-1} penicillin. Twenty-four hours later, TPCK-treated trypsin (1 $\mu\text{g ml}^{-1}$) was added and the supernatant collected after a further 42 h and stored at -80°C . To determine the incorporation frequencies of the competing PB1 genes, the progeny viruses in the transfection supernatant were subjected to a plaque assay in MDCK cells. Randomly chosen plaques (approximately 36 per experiment) were picked by sampling through the agarose and resuspended in 0.05% Triton-X100. The source of the competing gene segments was identified by gene-specific quantitative RT-PCR using the SensiFast probe no-ROX one-step RT-PCR Kit (Bioline). Each 20 μl reaction was performed using 5 μl of plaque-picked virus suspension, 10 μl of 2× SensiFast SYBR No-ROX One-Step mix, 0.2 μl of reverse transcriptase, 0.4 μl of Ribosafe RNase inhibitor, 0.8 μl of each 10 μM gene-specific forward and reverse primer and 0.08 μl of each 25 μM gene-specific probe. Primer and probe sequences are given in Supplementary Table 3. The RT-PCR reaction was incubated for 10 min at 45°C , before proceeding with qPCR amplification, according to the manufacturer's instructions. The reported values are combined data from 3–8 independent transfection experiments for each competition.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All sequencing data have been deposited with the Sequence Read Archive (accession numbers: SRP127020, SRP126994, SRP150669, SRP150677 and SRP212811); the processed SHAPE-MaP reactivities are available in SNRNASM format as Supplementary Table 1 and processed SPLASH data as Supplementary Table 2.

Received: 13 February 2019; Accepted: 12 June 2019;
Published online: 22 July 2019

References

- Gerber, M., Isel, C., Moules, V. & Marquet, R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol.* **22**, 446–455 (2014).
- 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).
- Gavazzi, C. et al. A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc. Natl Acad. Sci. USA* **110**, 16604–16609 (2013).
- Fournier, E. et al. A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res.* **40**, 2197–2209 (2012).
- Marsh, G. A., Rabadán, R., Levine, A. J. & Palese, P. Highly conserved regions of influenza A virus polymerase gene segments are critical for efficient viral RNA packaging. *J. Virol.* **82**, 2295–2304 (2008).
- Gog, J. R. et al. Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res.* **35**, 1897–1907 (2007).
- 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).
- Hutchinson, E. C., Wise, H. M., Kudryavtseva, K., Curran, M. D. & Digard, P. Characterisation of influenza A viruses with mutations in segment 5 packaging signals. *Vaccine* **27**, 6270–6275 (2009).
- Gilbertson, B. et al. Influenza NA and PB1 gene segments interact during the formation of viral progeny: localization of the binding region within the PB1 gene. *Viruses* **8**, E238 (2016).
- Lowen, A. C. Constraints, drivers, and implications of influenza A virus reassortment. *Annu. Rev. Virol.* **4**, 105–121 (2017).
- Taubenberger, J. K. & Kash, J. C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **7**, 440–451 (2010).
- Arranz, R. et al. The structure of native influenza virion ribonucleoproteins. *Science* **338**, 1634–1637 (2012).
- Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B. & Wilson, I. A. Organization of the influenza virus replication machinery. *Science* **338**, 1631–1634 (2012).
- Te Velthuis, A. J. & Fodor, E. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. *Nat. Rev. Microbiol.* **14**, 479–493 (2016).
- Pflug, A., Lukarska, M., Resa-Infante, P., Reich, S. & Cusack, S. Structural insights into RNA synthesis by the influenza virus transcription-replication machine. *Virus Res.* **234**, 103–117 (2017).
- Wilkinson, K. A. et al. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol.* **6**, e96 (2008).
- Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A. & Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nat. Protoc.* **10**, 1643–1669 (2015).
- Watts, J. M. et al. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**, 711–716 (2009).
- Lee, N. et al. Genome-wide analysis of influenza viral RNA and nucleoprotein association. *Nucleic Acids Res.* **45**, 8968–8977 (2017).
- Williams, G. D. et al. Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat. Commun.* **9**, 465 (2018).
- Baudin, F., Bach, C., Cusack, S. & Ruigrok, R. W. Structure of influenza virus RNP. I. influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J.* **13**, 3158–3165 (1994).
- Kobayashi, Y. et al. Computational and molecular analysis of conserved influenza A virus RNA secondary structures involved in infectious virion production. *RNA Biol.* **13**, 883–894 (2016).
- Gulyaev, A. P. et al. RNA structural constraints in the evolution of the influenza A virus genome NP segment. *RNA Biol.* **11**, 942–952 (2014).
- Aw, J. G. A. et al. In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell* **62**, 603–617 (2016).
- Cobbin, J. C., Verity, E. E., Gilbertson, B. P., Rockman, S. P. & Brown, L. E. The source of the PB1 gene in influenza vaccine reassortants selectively alters the hemagglutinin content of the resulting seed virus. *J. Virol.* **87**, 5577–5585 (2013).
- Cobbin, J. C. et al. Influenza virus PB1 and neuraminidase gene segments can cosegregate during vaccine reassortment driven by interactions in the PB1 coding region. *J. Virol.* **88**, 8971–8980 (2014).
- Moreira, E. A. et al. A conserved influenza A virus nucleoprotein code controls specific viral genome packaging. *Nat. Commun.* **7**, 12861 (2016).
- Saira, K. et al. Sequence analysis of in vivo defective interfering-like RNA of influenza A H1N1 pandemic virus. *J. Virol.* **87**, 8064–8074 (2013).
- Noda, T. et al. Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat. Commun.* **3**, 639 (2012).
- 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. et al. Conserved and host-specific features of influenza virion architecture. *Nat. Commun.* **5**, 4816 (2014).
- Mitra, S. Detecting RNA tertiary folding by sedimentation velocity analytical ultracentrifugation. *Methods Mol. Biol.* **1086**, 265–288 (2014).
- Turner, R., Shefer, K. & Ares, M. Jr. Safer one-pot synthesis of the 'SHAPE' reagent 1-methyl-7-nitroisatoic anhydride (1m7). *RNA* **19**, 1857–1863 (2013).
- Wilkinson, K. A., Merino, E. J. & Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* **1**, 1610–1616 (2006).
- Aw, J. G. A., Shen, Y., Nagarajan, N. & Wan, Y. Mapping RNA-RNA interactions globally using biotinylated psoralen. *J. Vis. Exp.* **123**, 55255 (2017).
- Jiang, H., Lei, R., Ding, S. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
- Busan, S. & Weeks, K. M. Accurate detection of chemical modifications in RNA by mutational profiling (MaP) with ShapeMapper 2. *RNA* **24**, 143–148 (2018).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. *circize* implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
- Busch, A., Richter, A. S. & Backofen, R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**, 2849–2856 (2008).
- Leontis, N. B. & Zirbel, C. L. in *RNA 3D Structure Analysis and Prediction* (eds Leontis, N. & Westhof, E.) 281–298 (Springer, 2012).
- Natchiar, S. K., Myasnikov, A. G., Kratzat, H., Hazemann, I. & Klaholz, B. P. Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **551**, 472–477 (2017).
- Nguyen, T. H. D. et al. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* **523**, 47–52 (2015).
- Te Velthuis, A. J. W. et al. Mini viral RNAs act as innate immune agonists during influenza virus infection. *Nat. Microbiol.* **3**, 1234–1242 (2018).
- Lorenz, R. et al. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).
- Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R. G. DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl Acad. Sci. USA* **97**, 6108–6113 (2000).
- Hoffmann, E., Stech, J., Guan, Y., Webster, R. & Perez, D. Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* **146**, 2275–2289 (2001).
- Tannock, G. A., Paul, J. A. & Barry, R. D. Relative immunogenicity of the cold-adapted influenza virus A/Ann Arbor/6/60 (A/AA/6/60-ca), recombinants of A/AA/6/60-ca, and parental strains with similar surface antigens. *Infect. Immun.* **43**, 457–462 (1984).

Acknowledgements

We thank G.G. Brownlee and J. Kenyon for helpful discussions, J. Kenyon, J.G. Aw, and Y. Wan for sharing protocols, J. Robertson for making the 1M7 reagent, J. Sharps for technical assistance and St Jude Children's Research Hospital for providing the pHW2000 plasmid. This work was supported by a Wellcome Trust studentship (no. 105399/Z/14/Z to B.D.), a UK Biotechnology and Biological Sciences Research Council studentship (no. BB/M011224/1 to M.L.K.), grants from the National Institutes of Health (nos. HL111527, GM101237 and HG008133 to A.L.), a National Health and Medical Research Council of Australia programme grant (no. ID1071916 to L.E.B.), UK Medical Research Council programme grants (nos. MR/K000241/1 and MR/R009945/1 to E.F.) and a Sir Edward Penley Abraham Cephalosporin Junior Research Fellowship to D.L.V.B.

Author contributions

B.D. designed, performed and analysed the SHAPE-MaP and SPLASH experiments and associated virus replication and purification work. M.L.K. performed and analysed the SPLASH experiments on reassortant viruses and associated virus replication and purification experiments. D.L.V.B. designed the SPLASH experiments and analysed the resulting data. B.G. designed and performed the competitive reassortant virus rescue experiments, created the reassortant viruses and analysed the data. S.T. and B.G. performed the viral replication experiments with the reassortant viruses. S.R., A.L., L.E.B., E.F. and D.L.V.B. supervised the work. B.D., B.G. and D.L.V.B. wrote the paper with contributions from the other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41564-019-0513-7>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to L.E.B., E.F. or D.L.V.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. [For final submission](#): please carefully check your responses for accuracy; you will not be able to make changes later.

▶ Experimental design

1. Sample size

Describe how sample size was determined.

Three or two replicates for each experiment were performed. Sample sizes and statistical tests used are indicated in figure legends and Methods section.

2. Data exclusions

Describe any data exclusions.

No data was excluded

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

We have performed three repeats for in virio SHAPE experiments on WSN. Each repeat was done on newly grown virus stocks and the SHAPE modification was performed on separate days between the repeats. For ex virio SHAPE experiments two repeats for in vitro transcribed RNA and one repeat for naked viral RNA SHAPE experiments and in viro experiments on PR8 and Udorn were done. Ex virio RNAs were also SHAPE modified on different days.

We have performed two repeats for SPLASH experiments on genetically-identical samples of WSN virus that were done months apart. We have also performed two repeats each for SPLASH on PR8 and Udorn viruses, one from the Fodor Lab and one from the Brown lab, again performed months apart. For reassortant viruses, we have performed SPLASH once per reassortant, with each segment covered multiple times in the different reassortants.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Not relevant to study

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Not relevant to study

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- Test values indicating whether an effect is present
Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

For SHAPE experiments the sequencing reads were trimmed using Skewer package. The SHAPE reactivities were calculated using ShapeMapper2 scripts (published previously Siegfried et al., 2014, Nature). Sequencing data from SPLASH experiments was trimmed using Skewer software, and aligned to reference genome using STAR software. An interaction matrix was constructed and visualised in R, and the top interactions between segments were visualized using the circlize package in R. RNA structure predictions were performed using the ViennaRNA Package (2.0). Intermolecular RNA interaction predictions were done using the IntaRNA (v2.3.1) package. Structures were visualized using the VARNA package. Probability distribution function plots were acquired using ggplot2 package in R. Virus growth curves were plotted and ANOVA performed using GraphPad Prism software

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

1-methyl-7-nitroisatoic anhydride (1M7) reagent was custom synthesized following a previously published protocol (Turner et al., 2013, RNA).

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used in this study

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

Madin-Darby Canine Kidney cells were acquired from American Type Culture Collection

b. Describe the method of cell line authentication used.

No authentication was done

c. Report whether the cell lines were tested for mycoplasma contamination.

Cell lines tested negative for mycoplasma

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

No animals were used

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Research did not involve human participants