MiLeSIM: combining super-resolution and machine learning permits high-throughput virus structure analysis

Romain F. Laine[1], Gemma Goodfellow[1], Laurence J. Young[1], Jon Travers[2], Danielle Carroll[3], Oliver Dibben[3], Helen Bright[3], Clemens F. Kaminski[1]

[1]Laser Analytics Group, Department of Chemical Engineering and Biotechnology, Cambridge University, UK
[2]MedImmune Ltd., Sir Aaron Klug Building, Granta Park, Cambridge, UK
[3]AstraZeneca, Speke, Liverpool L24 9JW, UK

Abstract

The development of super-resolution microscopy techniques has enabled unprecedented structural description of supramolecular assemblies such as viruses. Here, we developed a methodology based on Structured Illumination Microscopy (SIM) combined with machine learning classification and followed by class-specific image quantification in order to perform high-resolution structural analysis of large population of viruses. This allows us to fully quantify the structural content of virus populations with important applications in the biopharmaceutical industry. We demonstrate the approach on viruses produced for oncolytic viriotherapy (Newcastle Disease Virus) and vaccine development (Influenza). This unique tool enables the rapid assessment of the quality of viral production with high throughput and the molecular specificity of fluorescence microscopy, which allows the use of direct un-purified samples from pooled harvest fluids.

Introduction

The potential of super-resolution microscopy (SRM) to unravel details of the structure and replication of viruses was explored early on[1,2] since the birth of SRM and has provided unprecedented insights into viral structure[3–6]. Most studies use the SRM techniques providing the highest resolutions such as Stimulated Emission Depletion (STED)[7] microscopy and Stochastic Optical Reconstruction Microscopy (STORM)[8] or *d*STORM[9], in order to extract highest level of details in the virus structure, but often at a cost of long acquisition times, thus limiting the imaging to static samples and low throughput. Recently, the technique of Structured Illumination Microscopy (SIM)[10,11] was applied to study large viruses such as the prototypic poxvirus[5,12] with much higher imaging speeds. However, beyond understanding the fine structure of viruses, there is a need to identify and analyse classes of structures within large viral populations especially in the biotechnology industry where virus production is often hindered by the large heterogeneity obtained from batch production schemes. In particular, campaigns of influenza immunization heavily rely on the timely and efficient production of specific virus strains. Similarly, a deeper understanding of the structural heterogeneity of oncolytic viruses such as Newcastle Disease Virus (NDV)[13,14] would enable optimization of the production processes and in turn an increase in accessibility of viriotherapy. However, quantifying and understanding this structural heterogeneity and its relation to the efficacy require imaging large numbers of viruses at sufficient spatial resolution to reveal its details. Typically, virus structural studies are performed using electron microscopy (EM) but this has the drawbacks of low throughput and low specificity. Therefore, the recent advances in SRM are a great opportunity to circumvent these. On the other hand, machine learning (ML) is now becoming a ubiquitous and powerful tool for image analysis and in particular image classification and is already commonly used in microscopy, especially for image segmentation[15,16].

Here, we present MiLeSIM (**M**ach**i**ne **Le**arning **S**tructured **I**llumination **M**icroscopy) as an efficient combination of SRM, ML-based classification and advanced image analysis for the quantification of virus population heterogeneity. We use ML algorithms to classify super-resolved images of a heterogeneous viral population

into a number of viral particle classes with uniform structural components (e.g. spherical, filamentous). Once classified, these uniform populations can be analysed via image analysis pipelines specifically adapted to each structural class, such as image model fitting. We and others have shown that appropriate model fitting can lead to precision in structural parameters beyond the resolution of the images used[3,17]. This allows an in-depth exploration of virus populations unachievable by electron microscopy (EM) by allowing the measurement of large numbers of particles with high specificity. This approach has great potential in the industrial production of virus-based biotechnology such as oncolytic viriotherapy and vaccine development.

First, we identify total internal reflection fluorescence (TIRF) SIM (TIRF-SIM)[18–20] as a method of choice for high-throughput imaging at a relevant resolution and exploit the speed and resolution advantages of TIRF-SIM compared to conventional SIM. We show that virus structure can be imaged with high-throughput potential (~500 virus particles/s) and high resolution (~90 nm). For this, we compared various imaging modalities potentially suitable for virus imaging to highlight the optimal trade-offs of TIRF-SIM. Then, the large imaging datasets are fed into a ML routine allowing for the automated classification of Newcastle Disease Virus (NDV) and live attenuated influenza virus (LAIV) vaccines, enabling further shape-specific quantitative analyses revealing in turn a fully quantitative description of the viral population structural content. The purpose of our study is to validate the MiLeSIM approach as a powerful tool for the virus-based biotechnology industry where virus structures can inform on viral production and for research laboratories interested in high-throughput structural studies.

Results

Comparison of fluorescence modalities for virus structure imaging

First, we explored three common types of SRM modalities for investigating the structure of purified NDV: *d*STORM, STED and TIRF-SIM. Figure 1 shows representative images obtained with these different modalities and individual particles imaged using TIRF-SIM. A conventional TIRF wide-field image is also shown for comparison. This imaging highlighted the trade-offs offered by the different methods and the typical resolutions and acquisition times required here are also indicated in Figure S1(a). It is clear that in order to improve the resolution beyond the ~90 nm provided by TIRF-SIM, the increase in acquisition time is significant.
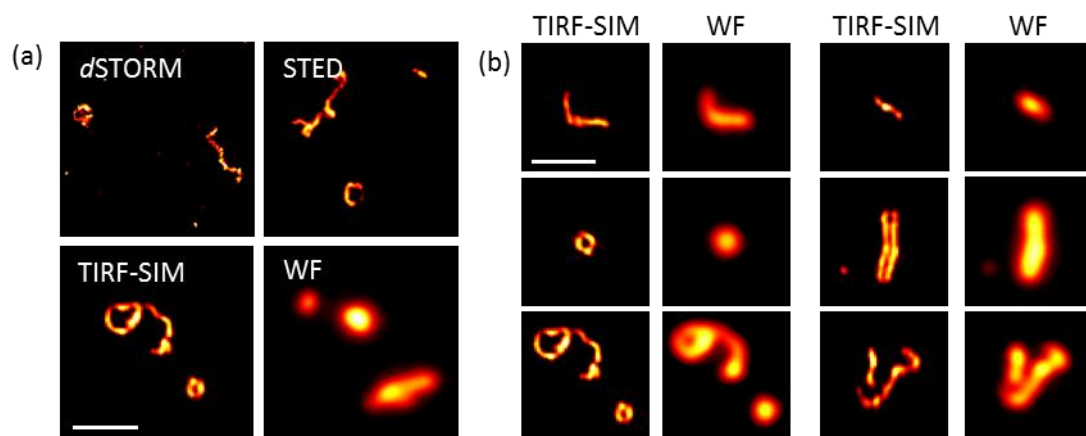
Figure 1: Super-resolution microscopy (SRM) for the study of NDV virus structure. (a) Representative images of purified NDV viruses with different imaging modalities. (b) Representative images of NDV virus population imaged with TIRF-SIM and their corresponding TIRF wide-field image. WF: wide-field TIRF microscopy. Scale bar: 1 μm.

Furthermore, we found that the increased resolution provided by STED and *d*STORM does not reveal additional structural details that were not already present in the TIRF-SIM images. We established that TIRF-

SIM is perfectly suited to this application as it provides sufficient structural detail to discern the multiple structural classes of viruses within the NDV population while providing the highest throughput. In particular, filamentous, spherical and rod-like structures were clearly identified. Figure 1 (c) shows representative images of NDV viruses obtained with the TIRF-SIM modality.

Workflow of MiLeSIM

The images obtained with TIRF-SIM (Figure 1(b)) revealed a number of stereotypical virus structures indicating a wide structural diversity in the virus population, occurring whether at the viral replication or purification levels. Understanding the origins and consequences of this heterogeneity not only informs on the life cycle of the virus, but also provides essential insight into the virus production process, which is of particular importance to manufacturers of virus-based therapeutics. Classification offers an insight into the extent of the heterogeneity and allows the application of automated image analysis adapted to each individual class independently. The workflow of the approach is shown in Figure 2. The individual virus particles are identified by automated segmentation and fed to the ML-based classification. Supervised ML using the random forest method[21] was used for robustness and ease of implementation. We identified 6 major structural classes: long and short filamentous, small and large spherical, rod-like shape and unknown. The unknown class is made of clumps of viral material with no identifiable and consistent shapes. Some combinations of structures were occasionally found such as double spherical and spherical with a tail but these constituted a very low fraction of the population.
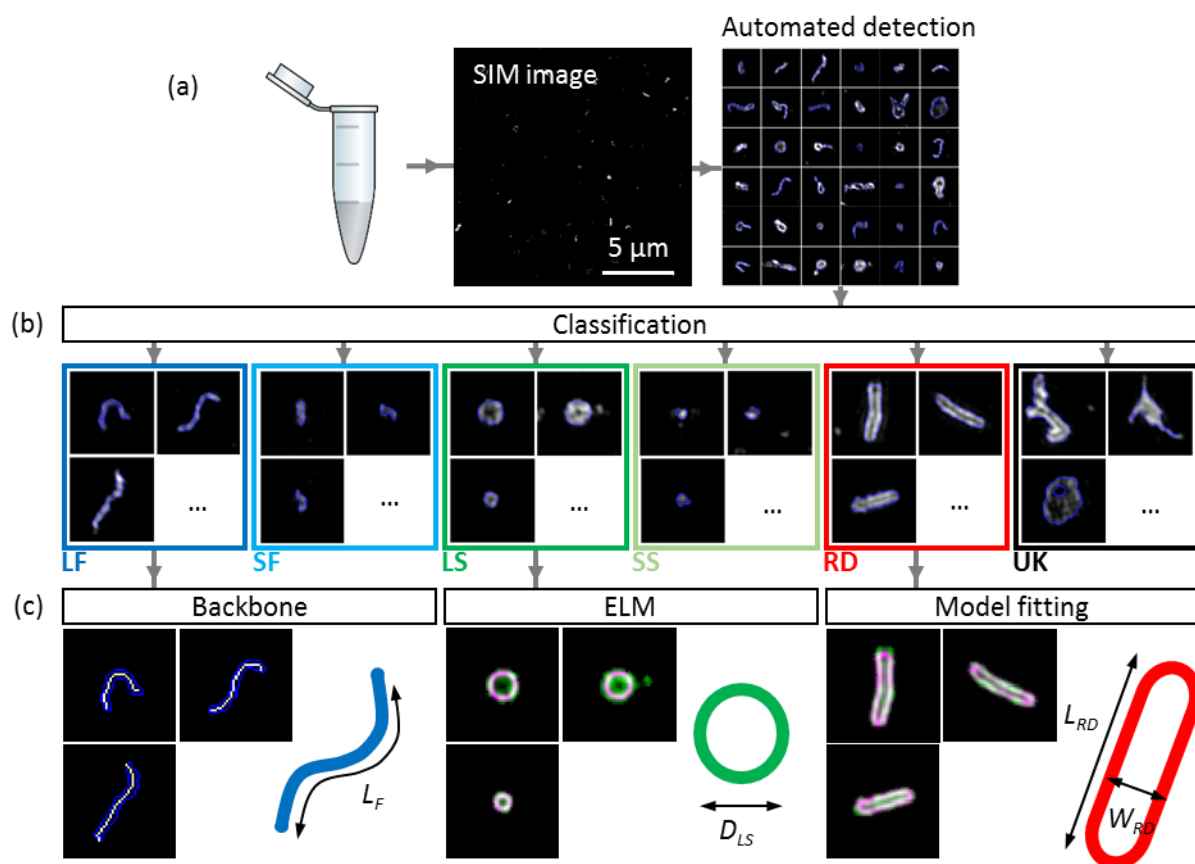

Figure 2: Workflow of automated detection, classification and analysis of NDV viral particles. SIM image (a) and segmented particles (b). The classified single-virus images (c) can be further analysed with a set of class-specific tools (d). For the backbone analysis the mask and backbone are showed in blue and white respectively. For the model-fitting approach (spherical and rod-like), the data and model are showed in green and magenta respectively. LF: long filamentous, SF: short filamentous, LS: large spherical, SS: small spherical, RD: rod-shape, UK: unknown. L_F, D_{LS}, L_{RD} and W_{RD}

represents the length of the filamentous particles, the diameter of the large spherical, and the length of the rod-shaped particles and the width of the rod-shaped particles respectively. Images of individual particles cover a field of view of 1.6 x 1.6 μm.

The filamentous particles (long and short) can in turn be analysed by extracting the backbone of the linear structure and measuring the length. The width of these structures appeared to be limited by the resolution of the imaging technique (~90 nm for TIRF-SIM, see Figure S1 but also observed in higher resolution approaches such as *d*STORM) and therefore, they were considered as 1D structures. The large spherical structures were analysed with the recently-developed technique, ellipsoid localization microscopy (ELM)[17]. This approach constitutes a form of model fitting to imaging data, which we showed allow the extraction of structural parameters beyond the image resolution[6,17]. However, we observed no significant ellipsoidicity and therefore we applied the spherical approach of ELM, extracting the radius of the particles. A similar approach to image model fitting was developed to analyse the rod-like viral particles by designing an appropriate model for this structural class (see materials and methods section and Figure S3 for details), which estimated both the length and width of the particles.

Machine learning-based virus structure classification

Structural classification was performed using a supervised ML algorithm in order to learn from a manually annotated dataset of particle images and their provided features, and then exploit this to identify the class of a new particle based on its features. This approach allows for rapid and automated classification of large datasets. Simple shape features include the area or length of the viral particle but more complex features are available (see Materials and Methods for details). The choice of algorithm and the set of features (also often called predictors) were optimised to maximise the overall predictive power based on the training dataset (comprised of 370 manually annotated particles). Figure 3 describes the list of chosen, individual predictors as well as the predictive power associated with the resulting classifier. A combination of predictors (Figure 3(a)) was selected from basic shapes features, Hu's image moments[22], features obtained from the pre-trained convolutional neural network (CNN) AlexNet[23] and from Speeded Up Robust Features (SURF)[24].

The predictors were selected based on the following criteria: basic structural features able to describe well the particles (e.g. area, eccentricity) and Hu's moments were chosen for their rotational and translational symmetry. For the features from AlexNet and SURF a feature selection approach was designed based on maximising the standard deviation across the different structural classes. This approach constitutes a more rational choice compared to simple principal component analysis (PCA) which does not typically take the information regarding the classes into account, therefore our method selects for predictors that have high potential for class discrimination. This data reduction narrowed down the number of predictors to 6 for AlexNet and 6 for SURF. A total of 24 descriptors was chosen: 6 based on basic shapes (area, ratio of axis lengths, eccentricity, solidity, perimeter-to-area, mean intensity, standard deviation of pixel intensities), 5 of Hu's image moments (Hu1, Hu4, Hu5, Hu6 and Phi4), 6 features obtained from the pre-trained convolutional neural network (CNN) AlexNet and 6 from a SURF bag of features. Figure 3(a) highlights the set of features used to train the random forest algorithm.
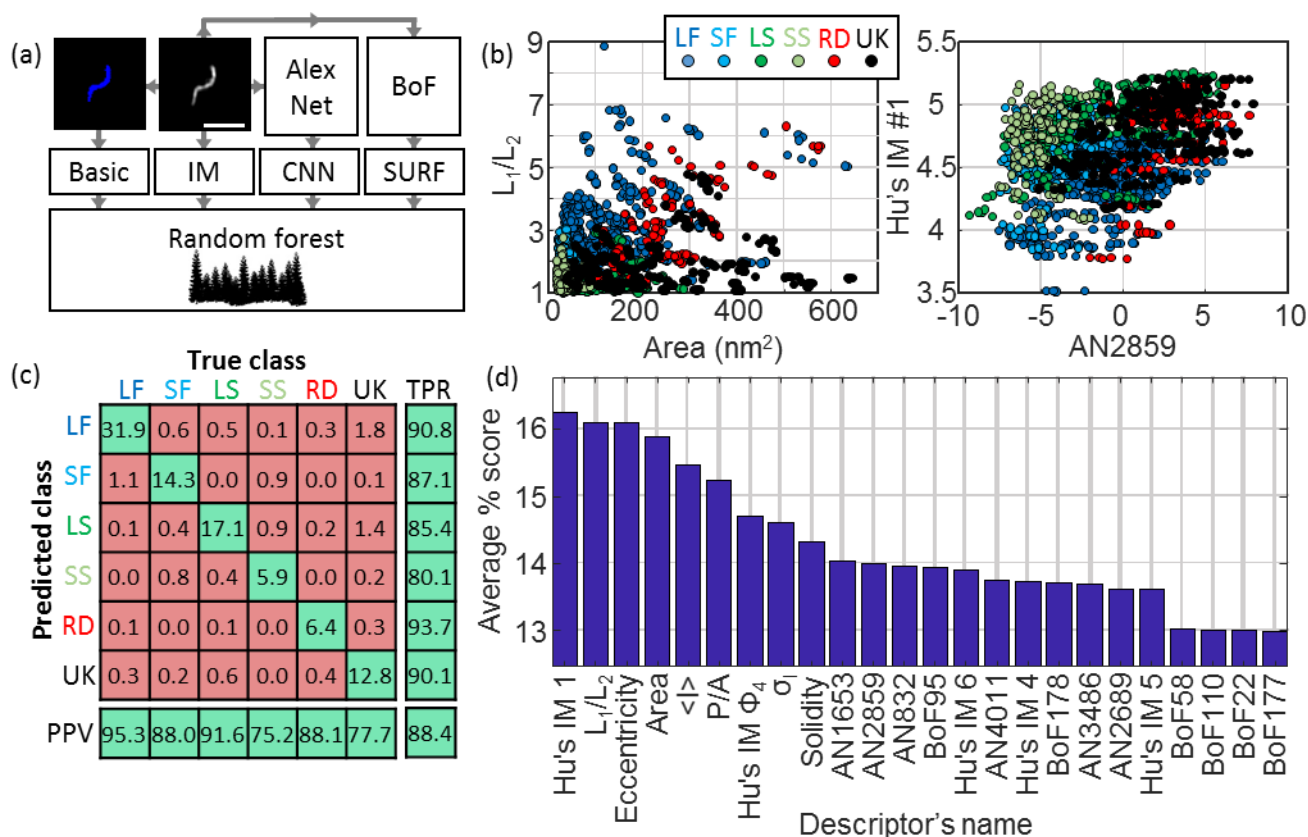
Figure 3: Machine learning-based classification. (a) Building the list of predictors from basic features, image moments, convolutional neural network (CNN) features and SURF bag of features (BoF). (b) 2D scatter plot of pairs of predictors showing the different predictive power of combination of predictors. (c) Confusion matrix obtained from the random forest showing the high true positive rate (TPR) and positive predictive values (PPV) of the classification. All numbers shown here are in percentage. (d) Scoring of the predictors sorted in descending order. IM: image moment. AN: AlexNet feature. BoF: SURF features. L_1/L_2: ratio of long axis over short axis. <I>: average intensity. P/A: perimeter to area ratio. σ_I: standard deviation of intensity. LF: long filamentous, SF: short filamentous, LS: large spherical, SS: small spherical, RD: rod-shape, UK: unknown.

Panels described in Figure 3(b) indicate the spread of pairs of descriptors highlighting that some descriptors support classification across specific classes better than other combinations. Here, the training dataset was used to build a scatter plot of pairs of descriptors with the knowledge of their true classification (see colour scheme). For instance, the pair of descriptors L_1/L_2 and Area show a good separation between long filamentous (dark blue labels) and unknown structures (black labels). The confusion matrix (Figure 3(c)) highlights the effective true positive rate (TPR) and positive predictive values (PPV) across the different classes with an overall predictive power of 88.4%. We note that some long filamentous viruses are misclassified as small filamentous, some small filamentous are misclassified as small spherical and that, on occasion, some unknown structures populate the predicted long and large spherical structures. Considering the simple shapes of these viruses, it is expected that a small fraction of particles are misclassified as structures with close resemblance.

The scoring of the predictors presented in Figure 3(d) indicates the average predictive power of each individual predictor. A high score indicates a great capacity to discriminate the different classes. This scoring was performed by measuring the predictive power of the classification for many combinations of predictors and distributing the predictive power score across the predictors tested (see Materials and Methods for details). In other words, if a combination of 2 descriptors alone give a predictive power of 60%, a score of 30% is awarded to both predictors. This method was repeated, and the scores were averages across > 13,000 different combinations of descriptors.

Structural details of an NDV virus population

We analysed a total of ~6,500 particles using MiLeSIM and established that 49.7% of NDV particles present a filamentous shape whereas the large spherical, small spherical and rods represent 18.6%, 7.8% and 7.3% respectively (Figure 4(a)). In addition to structural classification, our high-resolution images also allow dimensional analysis at the single particle level. Here, we applied spherical ELM analysis to the large spherical viruses to extract the particle radius; backbone extraction to the short and long filamentous particles, to estimate the particle length; measured the equivalent radius of the particles for small spherical; and designed a model fitting for the rod structures. Figure 4 shows the distribution of structural parameters for each class. We observe that both long filamentous and large spherical are well described by a Gamma distribution whereas the small filamentous and small spherical are well described by a Gaussian distribution. The model-fitting applied to the rod-shaped particles (see materials and methods and Figure S3 for details) allows the extraction of both the width and length of each particle. Therefore, it is possible to plot the distribution of structural parameters as a contour plot (Figure 4(c)).
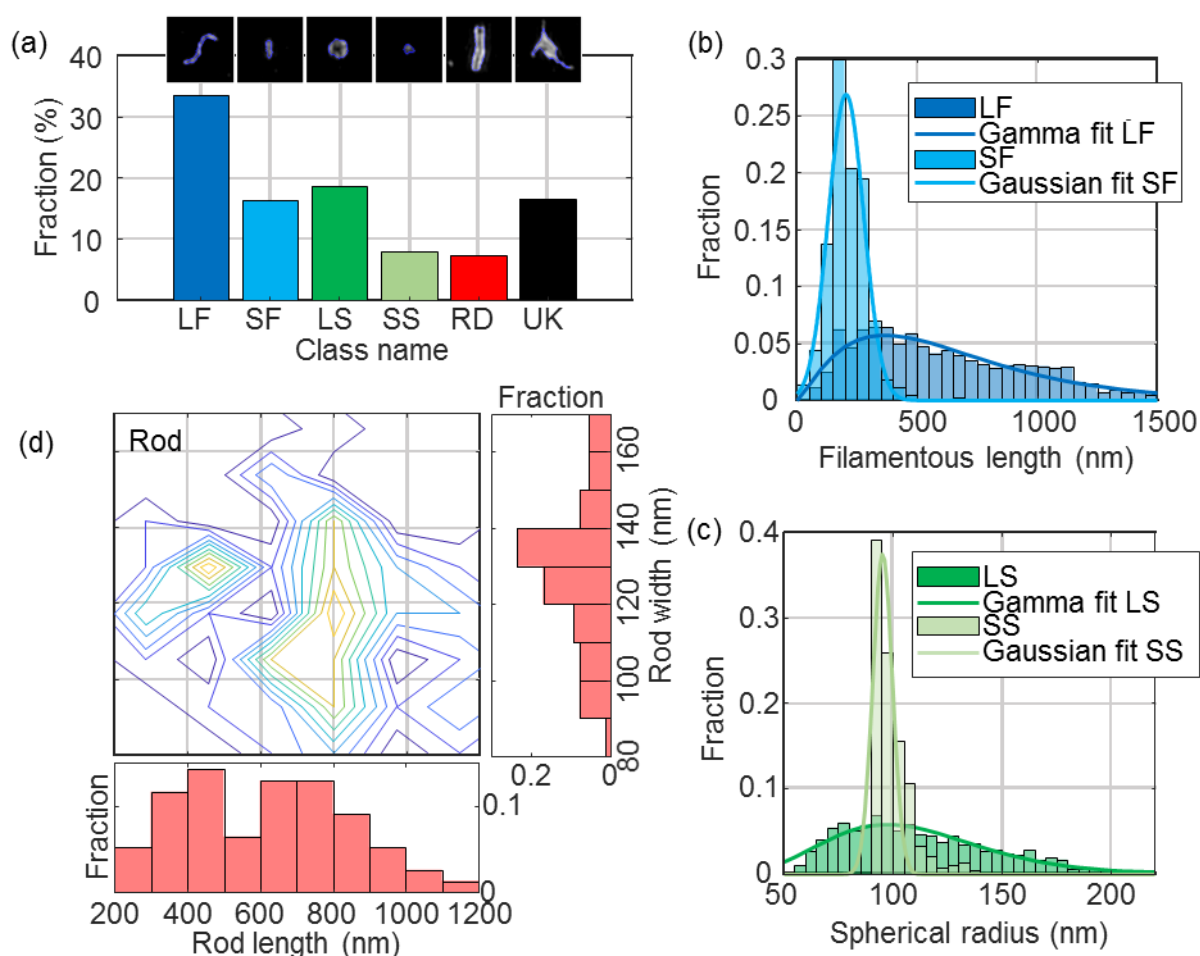
Figure 4: Quantitative analysis of NDV. The distribution of structural parameters for all classes was obtained from a total of ~6,500 virus particles. LF: long filamentous, SF: short filamentous, LS: large spherical, SS: small spherical, RD: rod-shape, UK: unknown.

We estimated the mean and standard deviation of the structural parameters from the distributions and obtained the following: $<L_{LF}> = 652 \pm 432$ nm, $<L_{SF}> = 203 \pm 104$ nm, $<D_{LS}> = 209 \pm 73$ nm, $<D_{SS}> = 190 \pm 12$ nm. For the rod-shaped particles, we observed that the width $<W_{RD}> = 135 \pm 28$ nm and the length $<D_{LS}> = 614 \pm 353$ nm are distributed around two populations as shown on the contour plot in Figure 4(d). It may first appear surprising that the distribution of radius of small spherical lies within that of the large spherical but in the ELM analysis the broadening of the image structure due to the finite optical resolution is effectively removed by

taking into account the point spread function (PSF). It should be noted that the small spherical distribution is centred on the value of optical resolution of our SIM microscope, which indicates that the small spherical structures are here diffraction-limited.

Comparison of the population structural details of different influenza strains used for vaccines

We extended the approach to a range of Live Attenuated Influenza Virus (LAIV) and classified the particles using the same classifier as for NDV. The LAIV structure were here dominated by spherical structures (> 60%). Figure 5 shows the distribution of particle sizes for four virus types: B Victoria, B Yamagata, A South Dakota and A Bolivia. The fractions of small and large spherical particles are shown as well as the equivalent radius and representative images of the viruses. Here, we clearly notice that B Victoria particles are dominated with large hollow particles with an equivalent radius of ~130 nm. This value is in good agreement with the ELM analysis leading to a particle diameter of ~200 nm and a resolution of ~90 nm (Figure S4). In contrast, the B Yamagata strain shows small and large particles of equal amount, indicating that the particles sizes are distributed around the region of overlap between small and large particles. This is confirmed by the nearly identical equivalent radius distributions.
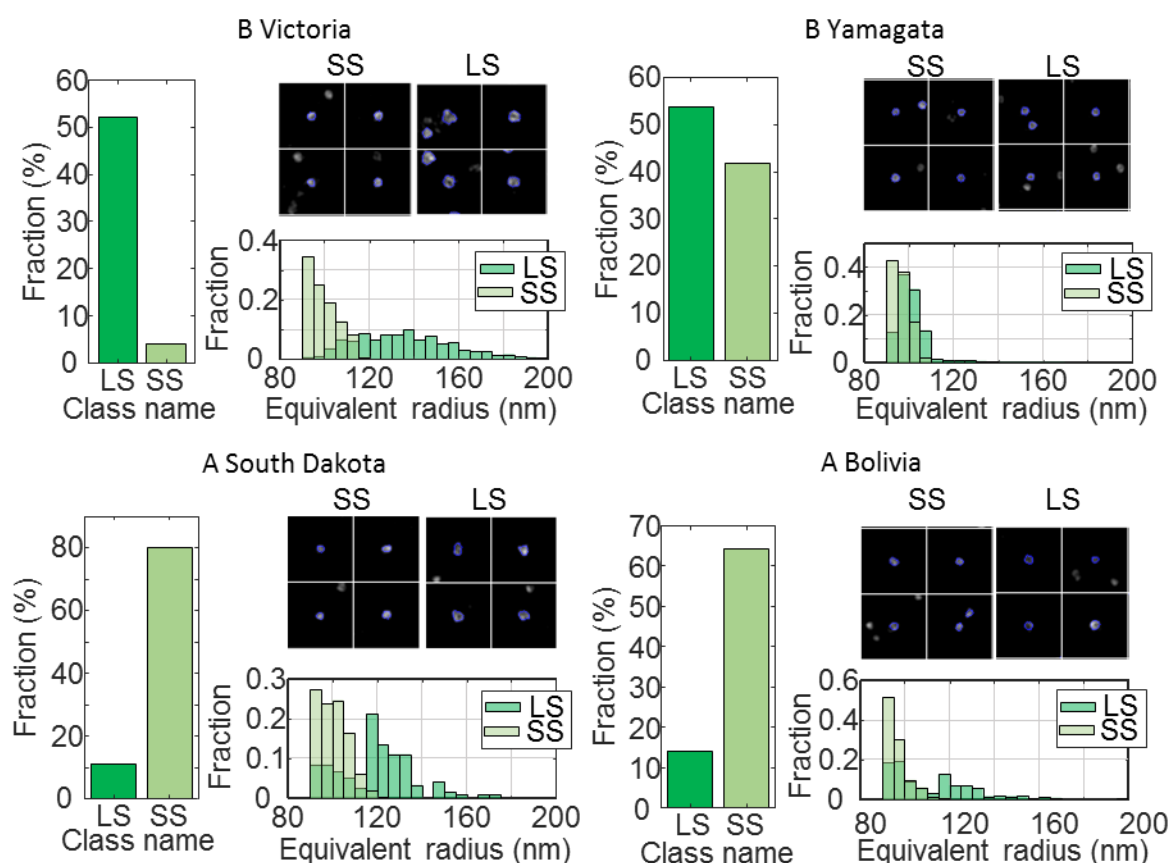

Figure 5: MiLeSIM approach applied to Live Attenuated Influenza Virus (LAIV). 2 types of B and A viruses were analysed here. The population was dominated by small and large spherical particles. The distributions of equivalent radius are shown here for both the large and small spherical for direct comparisons. The number of particles analysed were N=3,821, 4704, 1,062 and 1,756 for B Victoria, B Yamagata, A South Dakota and A Bolivia respectively.

Both A strains appeared clearly dominated by small spherical particles with sizes close to the resolution limit of our imaging. However, our high-throughput approach reveals subtle differences in the distribution of small

spherical structures where the A South Dakota viruses appear more heterogeneous (standard deviation ~ 10 nm), whereas the A Bolivia viruses are sharply distributed (standard deviation ~4 nm).

We also investigated the potential of directly imaging pool harvested fluid (PHF). LAIV are commonly propagated in embryonated hens' eggs where progeny viruses are released into the allantoic fluid of the egg. This fluid is harvested from numerous eggs and pooled. This constitutes a very basic and commonly used virus material. It is easy to produce and does not undergo any downstream purification. Consequently, PHF is impure, containing a variety of egg-derived impurities. The high molecular specificity of fluorescence microscopy allowed us to visualize the structure of the viruses with the same image quality despite the presence of large amount of impurities (Figure S5). Therefore our approach enables the study of unpurified samples and therefore allows probing the virus production at any intermediate levels of production and purification. This constitutes a strong advantage over electron microscopy (EM) techniques which typically require highly purified samples.

Discussion

We have demonstrated the potential of high-throughput imaging of virus structures, taking advantage of the trade-offs in TIRF-SIM between acquisition time and resolution. TIRF-SIM provided sufficient resolution to identify, discriminate and analyse individual viral structural classes. This approach combines ML, which was exploited to classify NDV viruses, with a model-based approach or direct quantification of the virus structure such as equivalent diameter. By comparing images for purified samples and samples from PHF, we showed that image quality was not degraded by using this low purification sample process. The high content imaging presented here allows fast imaging at up to 500 particles / second with 90 nm resolution, enabling the observation of structural variability in the NDV population and between different strains of LAIV. In this context, EM would not be able to provide the equivalent high throughput and high specificity for virus samples in their aqueous, unaltered forms.

Our particular classification uses random forest with a selection of descriptors from simple shape parameters, rotational and translational invariant image moments and features from AlexNet and common feature for image recognition such as SURF. The overall predictive power of the approach is ~88.5% and the mis-classification occur between classes that are similar (between small spherical and small filamentous for instance). The structural parameters that we extract from the model fitting are precise beyond the image resolution as they take into account the finite optical resolution. This therefore reveals subtle differences in populations such as the two sub-classes observed in the rod-shaped class. This approach will be beneficial especially when heterogeneous populations are present and need to be quantified. In our case, we see this being powerful to correlate structural information at the nanometre-scale and functional information that can be obtained through infectivity assays for instance. In turn, such combination of assays can allow process optimization for the production of virus-based therapeutics. However, the general applicability can be adapted to any type of fluorescence microscopy, SRM or not. This study demonstrates that MiLeSIM is a unique tool for high-throughput virus structural studies, applicable to industrially relevant bioprocessing and researchers interested in understanding the structural content of virus populations.

Materials and methods

Sample preparation. The NDV samples were prepared on cover slips as previously described[3]. Briefly, viruses were adhered on poly-L-lysine-coated Ibidi 8-well dishes, fixed, permeabilised and immuno-labelled with primary antibodies (mouse anti-hemagglutinin-neuraminidase HN, Abcam, UK) followed by secondary labelling (goat anti-mouse labelled with Alexa Fluor 647 for *d*STORM, with Alexa Fluor 488 for TIRF-SIM and with ATTO647-N for STED, Abcam, UK). The LAIV samples were prepared identically but using F16 mouse antibody for B victoria, Infa0121 mouse antibody for B Yamagata and FY1 human antibody for A South Dakota and A Bolivia. The corresponding secondary antibodies were used (donkey anti-mouse DyLight 488 labelled or rabbit anti-human DyLight 488 labelled antibodies, ThermoFisher). All virus samples originated from the master virus bank (MVB) and are therefore highly purified, unless indicated in the text, where the direct pool harvest fluid (PHF) was used.

TIRF-SIM, STED and *d*STORM imaging. Our custom-built TIRF-SIM system was described previously[20]. We used an Olympus UAPON 100x TIRF NA=1.49 and an Orca Flash 4.0 camera, with a sample pixel size of 64 nm. A total of 9 SIM images were acquired (3 phases, 3 orientations) with a camera exposure time of 200 ms and ~250 µW of 488 nm laser, measured at the back aperture of the objective. The SIM images were obtained using the reconstruction code provided by Dr Lin Shao[18], providing images with doubled resolution and 32 nm final pixel size using a Wiener filter of 0.01. The STED imaging was performed on our custom-built STED microscope as described previously[25,26]. The *d*STORM imaging was performed on a custom-built single-molecule microscope previously described[27,28] and with and mercaptoethylamine (MEA) buffer as previously described[3]. The *d*STORM image reconstruction was carried out using rapidSTORM 3[29].

The resolution achieved by the TIRF-SIM microscope was assessed by identifying the edge of the spatial frequency support using the SIMcheck plugin[30], as shown in Figure S1. For STED microscopy, the resolution was estimated from cross-sections of 20 nm beads and reporting the full width at half maximum (FWHM). The dSTORM resolution reported here was obtained from the FWHM of the localization precision, estimated by Thompson *et al.*[31].

Classification. All segmentations, features extractions and classifications were performed using MATLAB (Mathworks). A general diagram of the method is shown in Figure S2. The segmentation was obtained by an initial Otsu binarization and refined by active contour. This allowed a better outline of the particles and efficient separation of particles in close proximity. The particles that were judged too small or too dim to be particles were excluded from further analysis.

The basic shape features were extracted using the function *regionprops*. Hu's image moments were computed from the 71x71 pixels particle image centred on the centre of mass of the particle. The absolute values of the logarithm of the moments were used in the classification. For the features obtained from AlexNet[23], the individual 71x71 pixels images were resized to 227x227 pixels and used as all three color layers of the RGB images taken by AlexNet. Then, feature extraction was performed using AlexNet as a pre-trained network. 4096 features were obtained and data reduction was performed to limit the number of predictors used. For this, the features were averaged within each individual class and the standard deviation of every feature across the classes was computed. The 6 features with the highest standard deviation was selected. For the SURF features, first a bag of visual words was created from the training dataset, this bag was then used to check the presence of visual words in the 71x71 pixels images of individual particles. Similarly to AlexNet features, we selected only the 6 visual word features with the highest standard deviation across the different classes for classification. This allowed the computation of a total of 24 features for ML.

The classification was performed using a random forest algorithm. A training dataset made of 370 manually labelled individual particles was used to train the random forest across 60 epochs. The classification was

validated by 10-fold cross validation on the same dataset. The confusion matrix obtained from this cross-validation is show in Figure 3. For training the training data was augmented 5-fold by transforming the images with image translation and rotation randomly picked between 0 and 1 pixels and between 0 and 360 degrees respectively.

Feature scoring. The features were scored by computing the predictive power of the random forest trained on the training dataset but with only subsets of features. Out of the 24 features all combinations of 2, 3, 4, 24, 23 and 22 features were tested corresponding to a total of 13,227 combinations of features. For each combination of features, the predictive power obtained was split equally across the different features used, producing a "local" predictive power for each feature. This local score was average across all combinations using a specific feature to obtain the global score.

$$S_i = \frac{1}{N_c^i} \sum_{j=1}^{N_c} a_{ij} \frac{P_j}{n_j}$$

Where S_i is the global score of the feature i, N_c^i is the total number of combinations tested involving feature i, a_{ij} is a factor reflecting the presence of the feature i in the combination j. a_{ij} is equal to 1 if i is present in j, 0 otherwise. P_j is the predictive power of the combination j, N_c is the total number of combination tested and n_j is the number of features present in the combination j.

Quantitative analysis. All quantitative analyses were performed using MATLAB (Mathworks). The length of the filamentous structures were extracted by measuring the geodesic distance along the skeletonized image of the filament. The ELM analysis is freely available[17] and the code was adapted to insert within the workflow of our approach. The equivalent radius r of the small spherical particles were simply calculated from the area A of the segmented particle.

$$r = \sqrt{\frac{A}{\pi}}$$

The image model for the rod-shaped particles is presented in Figure S4. Briefly, the backbone of the particle was extracted by image thinning and then dilated by a disk-shaped kernel of radius equal to half of the width of the rod. The length of the rod could be adjusted by shortening the ends of the backbone or by extrapolating it outwards to lengthen it. The interior pixels of the image obtained were removed to leave the outline of the particle shape. This outline was then convolved with a Gaussian kernel in order to take into account the effect of the image resolution (here 90 nm). The intensity, the width and length of the model image were adjusted to minimize the sum of the square difference of intensity χ^2.

$$\chi^2 = \sum_{ij} [I_m(i,j) - I_d(i,j)]^2$$

Where i and j refer to the indices in the image, $I_m(i,j)$ is the image model, and $I_d(i,j)$ is the original image.

Acknowledgements

References

1. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313,** 1642–5 (2006).

2. Müller, B. & Heilemann, M. Shedding new light on viruses: Super-resolution microscopy for studying human immunodeficiency virus. *Trends Microbiol.* **21,** 522–533 (2013).

3. Laine, R. F. *et al.* Structural analysis of herpes simplex virus by optical super-resolution imaging. *Nat. Commun.* **6,** 5980 (2015).

4. Huang, F. *et al.* Ultra-High Resolution 3D Imaging of Whole Cells. *Cell* **166,** 1028–1040 (2016).

5. Gray, R. D. M. *et al.* VirusMapper: open-source nanoscale mapping of viral architecture through super-resolution microscopy. *Sci. Rep.* **6,** 29132 (2016).

6. Albecka, A., Laine, R. F., Janssen, A. F. J. J., Kaminski, C. F. & Crump, C. M. HSV-1 Glycoproteins Are Delivered to Virus Assembly Sites Through Dynamin-Dependent Endocytosis. *Traffic* **17,** 21–39 (2016).

7. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19,** 780 (1994).

8. Rust, M. J., Bates, M. & Zhuang, X. W. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* **3,** 793–795 (2006).

9. Heilemann, M. *et al.* Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes. *Angew. Chemie Int. Ed.* **47,** 6172–6176 (2008).

10. Gustafsson, M. G. L. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. SHORT COMMUNICATION. *J. Microsc.* **198,** 82–87 (2000).

11. Heintzmann, R. & Cremer, C. G. Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. in *Proc.SPIE 3568, Optical Biopsies and Microscopic Techniques III* (eds. Bigio, I. J., Schneckenburger, H., Slavik, J., Svanberg, K. & Viallet, P. M.) 185–196 (1999). doi:10.1117/12.336833

12. Horsington, J., Turnbull, L., Whitchurch, C. B. & Newsome, T. P. Sub-viral imaging of vaccinia virus using super-resolution microscopy. *J. Virol. Methods* **186,** 132–136 (2012).

13. Ganar, K., Das, M., Sinha, S. & Kumar, S. Newcastle disease virus: Current status and our understanding. *Virus Res.* **184,** 71–81 (2014).

14. Lichty, B. D., Breitbach, C. J., Stojdl, D. F. & Bell, J. C. Going viral with cancer immunotherapy. *Nat. Rev. Cancer* **14,** 559–567 (2014).

15. Van Valen, D. A. *et al.* Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments. *PLOS Comput. Biol.* **12,** e1005177 (2016).

16. Sommer, C., Straehle, C., Kothe, U. & Hamprecht, F. A. Ilastik: Interactive learning and segmentation toolkit. in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 230–233 (IEEE, 2011). doi:10.1109/ISBI.2011.5872394

17. Manetsberger, J. *et al.* Ellipsoid Localization Microscopy Infers the Size and Order of Protein Layers in Bacillus Spore Coats. *Biophys. J.* **109,** 2058–2066 (2015).

18. Shao, L., Kner, P., Rego, E. H. & Gustafsson, M. G. L. Super-resolution 3D microscopy of live whole cells using structured illumination. *Nat. Methods* **8,** 1044–1046 (2011).

19. Kner, P., Chhun, B. B., Griffis, E. R., Winoto, L. & Gustafsson, M. G. L. Super-resolution video

microscopy of live cells by structured illumination. *Nat. Methods* **6,** 339–342 (2009).

20. Young, L. J., Ströhl, F. & Kaminski, C. F. A Guide to Structured Illumination TIRF Microscopy at High Speed with Multiple Colors. *J. Vis. Exp.* (2016). doi:10.3791/53988

21. Breiman, L. Random Forests. *Mach. Learn.* **45,** 5–32 (2001).

22. Hu, M.-K. Visual pattern recognition by moment invariants. *Inf. Theory, IEEE Trans.* **8,** 179–187 (1962).

23. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* 1–9 (2012). doi:http://dx.doi.org/10.1016/j.protcy.2014.09.007

24. Bay, H., Ess, A., Tuytelaars, T. & Van Gool, L. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.* **110,** 346–359 (2008).

25. Mahou, P., Curry, N., Pinotsi, D., Kaminski Schierle, G. & Kaminski, C. Stimulated emission depletion microscopy to study amyloid fibril formation. in (eds. Enderlein, J., Gregor, I., Gryczynski, Z. K., Erdmann, R. & Koberling, F.) 93310U (2015). doi:10.1117/12.2079320

26. Curry, N., Ghézali, G., Kaminski Schierle, G. S., Rouach, N. & Kaminski, C. F. Correlative STED and Atomic Force Microscopy on Live Astrocytes Reveals Plasticity of Cytoskeletal Structure and Membrane Physical Properties during Polarized Migration. *Front. Cell. Neurosci.* **11,** (2017).

27. Ströhl, F. *et al.* Single Molecule Translation Imaging Visualizes the Dynamics of Local β-Actin Synthesis in Retinal Axons. *Sci. Rep.* **7,** 709 (2017).

28. Wong, H. H.-W. H. H.-W. *et al.* RNA Docking and Local Translation Regulate Site-Specific Axon Remodeling In Vivo. *Neuron* **95,** 852–868.e8 (2017).

29. Wolter, S. *et al.* rapidSTORM: accurate, fast open-source software for localization microscopy. *Nat. Methods* **9,** 1040–1 (2012).

30. Ball, G. *et al.* SIMcheck: a Toolbox for Successful Super-resolution Structured Illumination Microscopy. *Sci. Rep.* (2015). doi:10.1038/srep15915

31. Thompson, R. E., Larson, D. R. & Webb, W. W. Precise Nanometer Localization Analysis for Individual Fluorescent Probes. *Biophys. J.* **82,** 2775–2783 (2002).