Valproic Acid-Induced Changes of 4D Nuclear Morphology in Astrocyte Cells

Alexandr A. Kalinin[1,2,3], Xinhai Hou[1,2,4], Alex S. Ade[2], Gordon-Victor Fon[2], Walter Meixner[2], Gerald A. Higgins[2], Jonathan Z. Sexton[5,6,7], Xiang Wan[1], Ivo D. Dinov[2,3,8], Brian D. Athey[2,8,9,*]

[1]Shenzhen Research Institute of Big Data, Shenzhen 518172, Guangdong, China

[2]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA

[3]Statistics Online Computational Resource (SOCR), Health Behavior and Biological Sciences, University of Michigan, Ann Arbor, MI 48109, USA

[4]School of Science and Engineering, Chinese University of Hong Kong–Shenzhen, Shenzhen 518172, Guangdong, China

[5]Department of Internal Medicine, Gastroenterology, University of Michigan, Ann Arbor, MI 48109, USA

[6]Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, MI 48109, USA

[7]Center for Drug Repurposing, University of Michigan, Ann Arbor, MI 48109, USA

[8]Michigan Institute for Data Science (MIDAS), University of Michigan, Ann Arbor, MI 48109, USA

[9]Department of Psychiatry, University of Michigan, Ann Arbor, MI 48109, USA

[*]Corresponding author: bleu@umich.edu

Abstract

Histone deacetylase inhibitors, such as valproic acid (VPA), have important clinical implications as human therapeutics and cellular reprogramming agents. They induce chromatin re-organization associated with changes in cell nuclear morphology. Current approaches aiming to quantify these changes so far have been limited to basic 2D measures. Here, we quantified changes in 3D nuclear morphology of primary human astrocyte cells treated with VPA for 7 days (hence, 4D). We compared volumetric and surface-based 3D shape representations of cell nuclei and selected subset of features that jointly discriminated between normal and treated cells with 85% accuracy on day 7. Over time, VPA-treated nuclear morphologies progressed towards larger size and higher shape irregularity. On day 7, all 11 selected size and shape descriptors demonstrated significant difference between treated and untreated nuclear morphologies, including 22.5% increase in volume and 8.3% reduction in extent (shape regularity) for VPA-treated nuclei. Overall, we showed that 4D surface morphometry accurately characterizes the temporal changes in astrocyte nuclear form that are reflective of the underlying valproate-induced chromatin re-organization. These nuclear structural alterations may serve as a biomarker for histone (de-)acetylation events and provide insights into mechanisms of astrocytes-to-neurons reprogramming.

1 Introduction

Changes in nuclear morphology are reflective of chromatin reorganization that are regulated by complex biological mechanisms associated with cell differentiation, proliferation, disease, and mechanics[1–3]. Quantitative analyses of nuclear morphological changes become of major relevance as the studies of chromatin and DNA architecture emerge in the spatial and temporal framework, known as the 4D nucleome[4–7]. One of the most important mechanisms in chromatin remodeling is the post-translational modification of the N-terminal tails of histones by acetylation[8].

Histone acetylation is a significant chromatin event tightly controlled by the balance between the histone acetyl transferases and histone deacetylases (HDACs). Histone deacetylation results in chromatin condensation and subsequent transcriptional repression while acetylation has an antagonistic effect leading to increased de-compacted euchromatin and gene expression in cells[9–11]. Histone (de-)acetylation and other chromatin remodeling events are associated with quantifiable changes in nuclear morphology. For example, they play an important role in identifying cancer development and in predicting its progression[12,13]. This provides an opportunity to use nuclear morphometry to quantify complex phenotypes associated with the disease, to assess the efficacy of treatments at cellular and subcellular levels, and to link these to the underlying biological mechanisms[14,15].

To assess the role of chromatin in nuclear morphology, inhibitors of the enzymes that modulate histone modification state can be used to mimic chromatin alterations that are observed in human diseases[11,16,17]. Valproic acid (2-propylpentanoic acid, VPA) is an established drug in the long-term therapy of epilepsy, bipolar disorders, social phobias, and neuropathic pain[8,10]. VPA is known to relieve HDAC-dependent transcriptional repression and to cause hyperacetylation of histones in cultured cells and *in vivo*[8]. Inhibition of HDACs through small-molecule inhibitors has gained significant attention in clinical research[10,18,19]. Along with other HDAC inhibitors, VPA is now being tested in clinical studies as an anticancer drug and as an adjunctive therapeutic for trauma and sepsis in preclinical studies[19,20]. Another potential application of VPA is related to its ability to potentially promote neurogenesis and neuronal maturation[21–23] as it enhances the efficiency of cellular reprogramming mediated by HDAC inhibition[24,25]. Previous reports indicate that the neuronal regeneration capacity in adults may be insufficient for brain repair[26–28]. Thus, cell replacement therapy via neuronal reprogramming of terminally differentiated somatic cells can become a strategy to generate functional neurons[23,26]. Other studies performed in our laboratory

indicated that VPA activates neurogenic transcriptional programs in adult brain following traumatic brain injury[29].

Astrocytes, the most abundant cell types in the brain, play important roles in maintaining brain homeostasis and modulating neural circuit activity[30,31]. Astrocytes developmentally originate from the same precursor cells as neurons, are capable of proliferating in response to brain damages, and therefore are considered as ideal starting cells to regenerate neurons[32]. Recent studies have demonstrated the ability to directly reprogram human astrocytes into functional neurons with a set of small molecules, including VPA[22,23,33,34]. Compared to transcription-factor-based reprogramming, small molecules offer ease of use and a broader range of downstream applications[22]. However, the results reported by these studies are not always congruent. Some indicate that the removal of VPA from the chemical small molecule cocktail reduced reprogramming efficiency[22,33] or even report that VPA alone can induce astrocytes into neuroblasts[33]. Others argue that adding VPA to the reprogramming cocktail may slightly reduce reprogramming efficiency[35]. Treatment protocols, concentrations, and combinations of small molecules differ between these studies, indicating that underlying mechanisms driving the trans-differentiation process are not well understood.

There have been few studies that quantified nuclear size or shape changes associated with valproate-induced chromatin re-organization. For example, it has been shown that VPA treatment of prostate cancer cells after 7 days leads to an increase in the nuclear maximum diameter[36]. Similar results showing increase in cell or nuclear size in 2D have been demonstrated in other cell lines[16,37,38]. Some recent studies took basic measures of nuclear shape and reported an increase in 2D nuclear outline irregularity and an emergence of nuclear protrusions (“blebs”) that are reflective of altered histone modifications and chromatin decompaction under VPA treatment[3,11]. However, these approaches were limited to very few basic 2D surrogates of

geometric measures, such as cross-sectional nuclear area or maximum diameter. In our previous works, we have demonstrated that combining 3D size and shape characteristics yields a more representative characterization of the nuclear form that accurately reflects underlying changes in nuclear architecture and enables better discrimination of morphological phenotypes[7,39,40]. To our knowledge, there have been no studies that assessed effects of VPA-induced chromatin decompaction on astrocyte nuclear morphology.

To address these issues, we quantified VPA-induced changes in 4D nuclear morphology of primary human astrocyte cells. We first compared volumetric and surface-modeled representations of astrocyte nuclei, in order to assess their utility in characterizing changes in 3D nuclear organization over time. In order to do so, we extracted various size and shape descriptors from both representations. This allowed us to compare the predictive performance of machine learning models trained to distinguish between nuclear morphological profiles of treated cells vs. normal astrocytes over time. Machine learning-derived feature ranking was used to identify 11 surface-based features that were the most representative of changes in 3D structure. Further statistical analysis identified the increase of nuclear size in VPA-treated cells over the course of treatment, confirming and extending previously reported results. Valproate-induced changes in the 4D nuclear architecture were also captured in our analysis by shape features, showing higher nuclear shape irregularity and elongation in treated cells compared to the control population. These results allow to accurately characterize the effects of VPA on the astrocyte nuclear morphology as an effect of chromatin re-organization and provide insight towards describing the underlying biological mechanisms of astrocyte-to-neuron reprogramming.

2 Results

2.1 Experiment and data

In order to determine how valproate-induced alterations of astrocyte chromatin structure are reflected in 4D nuclear morphology, we treated human astrocyte cells with 1.5 mM of VPA at multiple time points (days 1, 3, and 5). We obtained volumetric images of DAPI-labeled nuclei using confocal microscopy at 3 time points (days 3, 5, and 7). This provided us with 4D images (3D + time) in following conditions: normal human astrocytes (NHA) and cells treated with VPA (VPA), as schematically shown in Figure 1a. In order to perform morphometric analysis, we first segmented individual nuclei of both normal and treated astrocytes. The automatic 3D segmentation of DAPI-stained nuclei using the Farsight Nuclear Segmentation algorithm[41] allowed avoiding a labor-intensive process of manual pixel-level annotation of large 3D volumetric imaging data. As the result, we extracted 3D nuclear binary segmentation masks for each cell condition on days 3, 5, and 7 (Figure 1b). We then performed post-segmentation processing that included 3D hole filling and semi-automated quality control of the extracted binary masks by removing the objects on edges of the volume, touching objects, or those with volume values outside of the empirically estimated interval. Numbers of nuclear masks after segmentation and quality control are listed in Table 1 for each day and treatment condition. Details of the segmentation and quality control protocols can be found in the Methods section.

Day / **Treatment**	**3**	**5**	**7**
NHA	164	150	93
VPA	123	55	70
Total	187	205	163

Table 1. Number of segmented 3D nuclear binary masks per day for each treatment condition.

Using segmented 3D binary masks, we performed the extraction of the volumetric and surface model-based nuclear shape representations, see Figure 1b. From raw voxels, we computed the total of 15 geometric characteristics of the individual nuclear masks: nuclear volume, bounding box and convex hull volumes, extent, solidity, lengths of main axes, inertia tensor principal components, and sphericity. As an alternative to a raw volumetric representation, we also used shape modeling to obtain nuclear surface reconstructions. Robust and smooth surface reconstruction was obtained from a 3D binary voxel mask via Laplace-Beltrami eigen-projection followed by topology-preserving boundary deformation to remove artifacts[42]. We obtained extracted surfaces that were smooth, accurately represented the shape of an object, and were well-suited for morphometric analysis of cell nuclei and subnuclear structures[40]. Surface reconstructions of nuclear morphologies were saved as discretized triangular meshes. From these surface representations, we computed a set of size and shape measures that were equivalent to that extracted from volumetric data, with an addition of bounding cylinder and sphere volumes, mean and Gaussian curvatures, curvedness, shape index, and fractal dimension. All extracted features were aggregated into per-nucleus feature vectors, from which we constructed a feature table per each day of treatment with the labels corresponding to phenotypic conditions (NHA and VPA). Combined, these feature tables formed the feature tensor containing nuclear morphological profiles (Figure 1b) that was used for model and feature selection, statistical and machine learning-based analysis, and interpretation of results (Figure 1c).

2.2 Classification and feature selection

We compared volumetric and surface-based feature sets as inputs for training a number of machine learning algorithms to discriminate between NHA and VPA nuclear morphological phenotypes (Figure 2a). Six different sets of features were used to train eight different classifiers at each time point. Area under the Receiving Operator Characteristic curve (AUC) was used as the main classification performance metric to evaluate models. To compare classifiers, we

averaged the performance of each model over days 3, 5, and 7, and over all feature sets. To compare feature sets, we averaged the AUCs over all models and days. AUCs (Figure 2a and Figure 2d), confusion matrices (Figure 2e), and feature ranking (Figure 2f) were averaged from 10 repetitions of the internal statistical 5-fold stratified cross-validation. Due to the ~3:1 sample imbalance for day 5 data, we used random subsampling of the prevalent class by the total number of per-fold training samples from an underrepresented class at each iteration of the cross-validation procedure. Although all classifiers demonstrated consistent performance with time-averaged AUCs between 74% and 80%, the random forest classifier (RFC)[43] robustly demonstrated highest average performance across all feature sets (mean AUC=80%) compared to alternatives. Therefore, we used it for further analysis.

We compared two subsets of equivalent features V_{sub} and S_{sub} extracted from voxel data and surface representations, respectively. Each of subsets consisted of 10 size and shape descriptors: volume, bounding box volume, convex hull volume, extent, solidity, inertia tensor eigenvalues, major axis length, and sphericity. As shown in Figure 2a, the results were similar, although voxel features resulted in slightly better performance in some models, including random forest. When we compared full sets of volumetric (*V*) and surface-based features (*S*), additional shape descriptors measured from surfaces alleviated that difference in performance. However, the best average performance was achieved when using the voxel and surface-based features together (*V*+*S*).

In order to identify the set of features that are the most predictive of the phenotype on these data, we first computed correlations between all features that identified 2 large clusters, roughly corresponding to size and shape descriptors (Figure 2b). To handle multicollinearity, we performed hierarchical clustering on the features' Spearman rank-order correlations. We computed AUCs for tested classifiers at different thresholds to choose one that produced the

minimal number of clusters without sacrificing average AUC. This provided us with 11 groups of features, from which we chose a single measure per cluster to obtain the unbiased estimation of feature relevance by calculating permutation importance and to perform univariate statistical analysis. As shown by the t-SNE[44] 2D projection in Figure 2c, the selected features provided a good representation of morphological profiles of nuclei at different time points. Specifically, there was a fairly good separation between NHA and VPA density clusters on days 3 and 7, with the exception for the day 5, when clusters densities were close. Figure 2c shows that S_{11} feature set provided an informative representation of nuclear morphologies in treated and control populations, although discriminating between conditions may be not trivial. It should be noted that cluster sizes and inter-cluster distances should be interpreted with care when using t-SNE[45].

These observations are further confirmed by the analysis of the random forest classifier performance, see Figure 2d. The model showed robust results with only 1% decrease in AUC on S_{11} features compared to V+S, and reached high classification performance on days 3 and 7 with mean AUCs of 82% and 85%, respectively. However, AUC was only 70% on the day 5, suggesting that morphological differences induced by VPA by day 3 become less prominent 24 hours later and then show large effect again by day 7. The confusion matrix in Figure 2e shows that classification errors were split almost equally between false positives and false negatives for each day, confirming the effectiveness of the prevalent class subsampling. This random forest model was then used to compute relative feature ranking using a permutation strategy.

The permutation feature importance reflects the decrease in a model performance when a single feature value is randomly shuffled[43]. This breaks the relationship between the feature and the outcome, while being model agnostic and can be calculated many times with different permutations of the feature. We computed permutation importance on a held-out set on each cross-validation cycle, highlighting which features contribute the most to the generalization power

of the trained model. As Figure 2e shows, bounding cylinder volume and average mean curvature were the most important features for day 3, followed by minimum axis length, compactness and volume. That indicates the relevance of both size and shape features for discriminating treated and untreated nuclear morphologies as early as 48 hours after treatment. On day 5, maximum axis length, average mean curvature, and solidity were the top-3 important descriptors. Bounding box volume was once again the most important measure on day 7, followed by minimum axis length, average mean curvature, solidity, and extent. Overall, the feature ranking profiles for days 3 and 7 were more alike comparing to that for day 5.

2.2 VPA induced increased nuclear size

The majority of both volumetric and surface-derived morphometric features were those that measured nuclear size. Previously published studies reported increased 2D cellular and nuclear area and diameter under VPA treatment[16,37]. Our findings confirmed and extended those results, as the most apparent change in morphometric measures over time for the control (NHA) and treatment (VPA) cell conditions was the larger nuclear size for the latter. As Figure 3 shows, univariate statistical analysis on each recorded day of treatment demonstrated a trend for the size increase over the course of treatment as characterized by S_{11} measures. Bounding box volume was not selected as a part of S_{11} feature set, but is reported here for the reference. We reported each relative difference as the percentage change from the control group, along with a *p*-value obtained using two-sided Mann-Whitney *U* test with Holm–Šidák multiple testing correction (α=0.05), along with the difference of means (*DoM*), and the common language effect size statistic that in the case of Mann-Whitney *U* test is equivalent to the Area under the Receiving Operator Characteristic curve (*AUC*)[46].

Before comparing linear size measures of NHA and VPA-treated cell nuclei, we should note that between days 3 and 5 controls exhibited increase in both major and minor nuclear axes by 10.8%

and 11.9% respectively (p=2.06×10^{-10}, *DoM*=1.96 μm, *AUC*=0.70 for major axis; p=5.20×10^{-11}, *DoM*=1.65 μm, *AUC*=0.71 for median axis). Major axis length of VPA-treated cell nuclei also followed the increasing trend over time, but the difference was even greater than in NHA cells by 9.4%, 11.8%, and 12.8% on days 3, 5, and 7, respectively (p=1.53×10^{-7}, *DoM*=1.72 μm, *AUC*=0.68 for day 3; p=9.95×10^{-7}, *DoM*=2.38 μm, *AUC*=0.72 for day 5; p=9.95×10^{-7}, *DoM*=2.67 μm, *AUC*=0.73 for day 7). The difference in median axis length of treated vs. untreated nuclei was not as drastic on days 3 and 5 (+4.6%, p=8.23×10^{-3}, *DoM*=0.64 μm, *AUC*=0.59 for day 3; +2.9%, p=6.56×10^{-1}, *DoM*=0.46 μm, *AUC*=0.52 for day 5), but increased to 14% on day 7 (p=6.55×10^{-6}, *DoM*=2.23 μm, *AUC*=0.71). Nuclei in the VPA-treated group exhibited shorter minor axes for days 3 and 5 (-4.3%, p=1.98×10^{-3}, *DoM*=0.29 μm, *AUC*=0.61 for day 3; -3.3%, p=6.69×10^{-3}, *DoM*=0.21 μm, *AUC*=0.62 for day 5), but showed an increase over the untreated group by 4.5% on day 7 (p=3.79×10^{-3}, *DoM*=0.29 μm, *AUC*=0.63). Simultaneous increase in major and median axis lengths and shortening of the minor axis were characteristic for both NHA and VPA groups between days 3 and 5, which indicated elongation and flattening of nuclei. By day 7, nuclei of valproate-treated cell nuclei were significantly larger their control counterparts in all three axes.

Volume descriptors exhibited similar intra-group changes. Specifically, NHA nuclear volumes increased significantly between days 3 and 5 by 22.0% (p=1.49×10^{-21}, *DoM*=205.57μm^3, *AUC*=0.81), although this change was less prominent after day 5 (+6%, p=5.04×10^{-3}, *DoM*=76.84 μm^3, *AUC*=0.61). This increase in size of normal controls made differences between treated and untreated nuclear volumes smaller on days 3 and 5 (+5.6%, p=3.57×10^{-3}, *DoM*=53.27 μm^3, *AUC*=0.60 for day 3; +8.7%, p=1.89×10^{-1}, *DoM*=99.63 μm^3, AUC=0.56 for day 5). However, nuclear volumes in VPA group were larger than those of NHA by 22.5% on day 7 (p=1.10×10^{-8}, *DoM*=275.75 μm^3, *AUC*=0.76). Bounding cylinder volume showed even more substantial difference, with VPA-treated nuclei reporting 17.7% and 21.4% larger average values on days 3

and 5 relative to the control group (p=3.56×10^{-9}, DoM=351.77 μm^3, AUC=0.70 for day 3; p=1.71×10^{-4}, DoM=495.72 μm^3, AUC=0.67 for day 5). Finally, the largest relative percentage change and effect size were reported on day 7 when comparing nuclear bounding cylinder volumes of treated cells and untreated nuclei with the former being greater by 40.1% (p=1.52×10^{-12}, DoM=951.78 μm^3, AUC=0.82). These measurements of nuclear volumes and corresponding bounding primitives demonstrated congruent results suggesting that nuclei of VPA-treated cells were increasing in size more dramatically over time, compared those of untreated astrocytes. The difference between the increase in bounding primitive volume and in the nuclear volume itself can indicate the change in the object shape, which is explored in detail in the next subsection.

Examples of an NHA and a VPA-treated astrocyte nuclear surfaces is shown in Figure 3b, with valproate-treated nucleus having larger major axis length and volume. These results confirm and extend previously reported findings that suggested nuclear enlargement as one of indications of the chromatin structure relaxation induced by valproate. Our findings also highlight a transition point between two phases of morphological changes. First, the size of VPA treated nuclei demonstrated an increase in size in the first 48 hours compared to the control population (days 0 to 3). However, by day 5, NHA nuclei also increased in size, which led to smaller difference as measured by the most size features. However, NHA nuclear size did not change as much after day 5, while descriptors of VPA-treated nuclear size showed biggest relative increase and effect sizes on the last day of treatment.

2.2 VPA induced nuclear shape irregularity

Among the selected shape features (Figure 4a), we first looked at those that can be directly derived from the size measured discussed above. Sphericity can be expressed either via volume and surface area or computed using principal axes (see Methods). It assesses the compactness

of the 3D object, i.e. it measures how closely the global object shape resembles that of a perfect sphere. Accordingly, we found that nuclear compactness was decreasing over time for both NHA and VPA cells, but sphericity of treated nuclei was lower by 3.8%, 4.1% and 3.2% for days 3, 5, and 7, respectively (p=2.83×10^{-7}, *DoM*=0.033, *AUC*=0.68 for day 3; p=2.37×10^{-4}, *DoM*=0.034, *AUC*=0.67 for day 5; p=1.10×10^{-3}, *DoM*=0.026, *AUC*=0.65 for day 7). These results are showing that VPA-induced large increase in major and median axes manifested in nuclear shapes that are more elongated up to day 5 and overall larger and less spherical on day 7, compared to NHA cells.

The difference between the increase in bounding primitive volumes and in the nuclear volume itself hinted at the shape change that can be captured by ratios of these measures. Extent (ratio of the object volume to the bounding box volume) and solidity (ratio of the object volume to the convex hull volume) of nuclear surfaces in both groups demonstrated significant decrease on day 7 by 8.3% and 2.1%, correspondingly (p=7.95×10^{-7}, *DoM*=0.049, *AUC*=0.73 for extent; p=1.90×10^{-3}, *DoM*=0.020, *AUC*=0.64 for solidity). Before that, extent of VPA-treated nuclei was also 3.6% lower than NHA on day 3 (p=7.06×10^{-3}, *DoM*=0.020, *AUC*=0.60), which corresponds to increase in bounding box volume and major axis length, see Figure 3c. Unlike sphericity, these size-derived shape descriptors do not quantify roundness of an object, instead they are useful for quantification of the amount and size of concavities (or protrusions) in an object boundary. Thus, our results indicated that by day 7 VPA-treated cells on average had more globally irregularly-shaped nuclei, for instance, as shown by representative kidney-shaped examples in Figure 4c-d. Besides, Figure 4d also shows nuclei with small abnormal deformations termed "blebs" that are also argued to be indicative of chromatin decompaction[3,11]. These shape distortions can also be reflected in size and shape measures, for example, by increasing bounding primitive volumes and by resulting in lower extent values for valproate-treated nuclei.

Mean curvature is an extrinsic measure of shape that comes from differential geometry and that locally describes the curvature, which is then averaged across the object surface[47]. The obtained results showed that VPA-treated nuclei exhibited robust 3.0%, 3.1%, and 3.2% increase in average curvature corresponding to more shape irregularity at each corresponding time point (p=1.15×10^{-10}, *DoM*=0.003, *AUC*=0.72 for day 3; p=6.5×10^{-7}, *DoM*=0.003, *AUC*=0.73 for day 5; p=9.03×10^{-9}, *DoM*=0.003, *AUC*=0.76 for day 7). Notably, there were no significant intra-group differences between mean curvature values for both NHA and VPA cells (p>0.05 for days 3 vs. 5 and 5 vs. 7) suggesting that the VPA-induced increase in mean curvature occurred before imaging on day 3 and remained for the rest of the treatment course. These results were also supported by the high relative importance of mean curvature at each timepoint, as reported by the classification model (Figure 2f).

Gaussian curvature and fractal dimension are the other two features that were selected, but only demonstrated significant differences between treated and untreated morphologies on day 7. Gaussian curvature is an intrinsic (size-invariant) measure of curvature, depending only on distances that are measured on the surface. It has been decreasing in both NHA and VPA groups between days 3 and 5, but on day 7 valproate-treated cells had nuclei with more concave areas on the surface that resulted in 2.5% decrease in average Gaussian curvature compared to untreated counterparts (p=7.31×10^{-4}, *DoM*=5.0×10^{-5}, *AUC*=0.66). Fractal dimension is the measure of the object border complexity and our observations suggested a 0.13% decrease in this feature of VPA-treated cells on day 7 (p=4.40e-02, *DoM*= 0.003, *AUC*=0.61), which corresponds to less intensively folded and wrinkled nuclear surface.

Overall, shape measures indicated that initially nuclear elongation was occurring in VPA-treated cells up to day 5. Towards day 7, these nuclei demonstrated more prominent deformations including global shape irregularities characterized by lower sphericity, blebbing and kidney-like

forms, along with higher local curvature with the presence of surface concavities, bumps, and lower border complexity. Together, these descriptors provided a detailed profile of 4D nuclear morphology changes reflective of VPA-induced altered histone modifications and chromatin decompaction[3,11].

3 Discussion

In this study, we assessed VPA-induced changes in 4D spacetime nuclear morphology of primary human astrocytes. The known ability of VPA to promote pluripotency via HDAC inhibition[24] suggests that valproate-induced chromatin modifications play key role in the process of astrocyte-to-neuron reprogramming. Results of our experiment show that the increase in decompact euchromatin in VPA-treated astrocytes is reflected in quantifiable time-dependent changes in 3D nuclear size and shape. We demonstrated the ability to accurately differentiate between nuclear morphological profiles of normal and valproate-treated human astrocytes using surface-based descriptors and provided detailed characterization of the temporal dynamics of these features over 7 days of treatment.

Although most studies to the date performed morphometric quantification of changes in cells and nuclei in 2D, there is evidence that 3D representations are more informative and allow for better characterization of cell morphological phenotypes[39,48,49]. However, there are multiple ways to measure cell nuclear shapes in 3D depending on their extracted representations[50]. Probably, the most common approach to compute shape features is to apply morphometric measures to the raw 3D geometric objects represented as voxel volumes[51]. The main drawback of voxels-based shape representations is that they can be noisy and may lose fine geometric details or even misrepresent the object's topological structure. On the other hand, accurate shape modeling for 3D objects' surface reconstructions requires an extra step that may be more computationally expensive.

Both voxel and surface-based morphometry representations were able to capture the temporal dynamics of 3D nuclear morphological variability over the course of valproate treatment. In terms of discrimination between NHA and VPA conditions, voxel and surface features produced similar results, with their combination achieving the best average performance. However, we then were able to reduce the full feature set to just 11 surface-based measures without sacrificing accuracy. Tree-based ensemble classifiers, such as random forest and extremely randomized trees, demonstrated the most robust classification performance, which may be attributed to their ability to resists overfitting when dealing with smaller sample sizes[43].

As characterized by the set of 11 selected features, VPA-treated astrocytes exhibited nuclear morphologies that were different from NHA group as early as day 3 and then diverged further from controls towards that end of the treatment course. Low-dimensional non-linear projections showed prominent clusters formed in feature space on day 7 (Figure 2c). These results are further confirmed by reported classification performance using random forest, reaching the highest AUC of 85% on the last day. This machine learning model was also used to compute relevant feature importance. Top selected features included few size measures such as major axis length and bounding box volume and a number of shape characteristics, including average mean curvature, demonstrating how a combination of different measures can aid in improving classification.

Using univariate statistical analysis, we then assessed the difference between treatment and control groups as measured by selected interpretable size and shape descriptors. Under valproate treatment, astrocytes tend to have larger nuclei starting from the first timepoint (day 3) as measured by the principal axes, the volume of the nucleus itself and its bounding primitives (box, cylinder, sphere). Although control cell population also increased in nuclear size between days 3 and 5, VPA-treated nuclei were elongated in the major axis and tended had a greater

bounding cylinder volume. By day 7, all size descriptors reported significant increase in size of VPA-treated nuclei on day 7 ($p<5.0\times10^{-3}$). These results expand prior reports on nuclear enlargement under VPA treatment[16,37,38,52], providing detailed characterization of how decondensed or dilated state of chromatin is manifested in corresponding changes in 3D geometric measures.

Another finding that extended results from previous 2D studies was statistically significant difference in shape[3,11] and fractal dimension[53,54]. Our results indicated both higher local shape curvature and overall less spherical form in VPA-treated nuclei at each time point. While lower sphericity can be attributed to the major axis elongation up to day 5, by day 7 we observed the variety of shape changes identified by lower extent (ratio of the object volume to the bounding box volume) and solidity (ratio of the object volume to the convex hull volume). These results were confirmed qualitatively as kidney-shaped and blebbed nuclei were observed more often in VPA-treated cells. At the same time, decrease in Gaussian curvature and fractal dimension on day 7 showed that treated nuclei had surface with more concavities and lower border complexity that has also been suggested to be indicative of chromatin decompaction, reduced tumorigenesis, and neuroprotection[54]. For example, one prior study reported consistently reduced fractal dimension in HT-29 cells upon treatment with VPA at various concentrations[53] that is in agreement with our result on day 7, although we did not observe significant differences for the images taken on days 3 and 5. These results described the dynamics of VPA-treated nuclear morphological profiles characterized by the increase in size and progressively more irregular, complex shapes.

These observations are the first step to studying time-dependent morphological effects of chromatin reorganization in the astrocyte-to-neuron reprogramming process and relating them to underlying molecular mechanisms. In future studies, this approach can be extended to label additional sub-nuclear components or organelles, such as nucleoli, chromosome territories,

Topologically Associated Domains (TADs), and other compartments using, for example, Cell Painting assay for high-content morphological profiling[55]. While tracking the overall cellular phenotype, this extension will allow us to include many other features in the models and assess their variability, association with cellular and nuclear shape morphology, disease state, and treatment conditions. All of this will set the stage to evaluate the effects of VPA and other small molecules measured with different concentrations and temporal sampling. Correlating these phenotypical cell and nuclear profiles with data from other assays, such as Chromatin Conformation Capture (Hi-C)[56], will likely provide useful insights into how altered functional properties of the genome are correlated with TAD structure, nuclear and cellular morphology, and can it be used for training modern machine learning models to more accurately predict treatment response in model systems and in humans[57].

4 Methods

Sample preparation and image acquisition

Primary human astrocyte cells were purchased from ScienCell (Human Astrocytes-hippocampal, Catalog #1830).

Day 0:

- collect 8 Day 0 samples:
    - fix samples in 4\% PFA for 10mins
    - rinse 3 x 5mins in PBS
    - store samples in PBS at 4deg
- replace media with 50% growth media and 50% N2 media (DMEM/F12 + 1X pen/strep, 1X N2 supplements).

Day 1:

- control samples: for 30ml of N2 media add 36ul DMSO
- VPA-treated (1.5mM VPA): for 30ml of N2 media add 450ul VPA

Day 3:

- collect 6 Day 3 control samples and 6 Day 3 VPA samples:
    - fix samples in 4\% PFA for 10mins
    - rinse 3 x 5mins in PBS
    - store samples in PBS at 4deg
- control samples: for 30ml of N2 media add 187.5ul DMSO.
- VPA-treated (1.5mM VPA): for 30ml of N2 media add 450ul VPA

Day 5:

- Collect 6 Day 5 control samples and 6 Day 5 VPA samples:
    - fix samples in 4\% PFA for 10mins
    - rinse 3 x 5mins in PBS
- control samples: For 30ml of N2 media add 37.5ul DMSO

- VPA-treated (1.5mM VPA): For 30ml of N2 media add 450ul VPA

Cells in both collections were labeled with DAPI (4',6-diamidino-2-phenylindole), a common stain for nuclear DNA. 3D imaging employed Zeiss LSM 710 laser scanning confocal microscope with a 63x PLAN/Apochromat 1.4NA DIC objective. Each 3D volume was then re-sliced into a 1,024 ×1,024×Z lattice (Z={30,50}), where regional sub-volumes facilitate the alignment with the native tile size of the microscope. For every sub-volume, accompanying vendor meta-data was extracted from the original data.

Segmentation

We performed the automatic 3D segmentation of nuclei using Nuclear Segmentation algorithm from the Farsight toolkit[39,41]. This tool was created specifically to segment DAPI-stained nuclei in 2D or 3D, it does not require a labeled training set, has a convenient command line interface, and demonstrated stable results on these data. The algorithm implements multiple steps which include a graph-cut algorithm to binarize the sub-volumes, a multi-scale Laplacian of Gaussian filter to convert the nuclei to blob masks, fast clustering to delineate the nuclei, and nuclear contour refinement using graph-cuts with alpha-expansions[41].

After segmentation of the DAPI channel sub-volumes, data were converted to 16-bit 3D TIFF files, each segmented nucleus was represented as a binary mask, and given a unique index value. Post-segmentation processing of nuclear masks followed the protocol proposed by Kalinin *et al.*[39] and included 3D hole filling and a filtering step that removed the objects if they span the edge of a tile, are connected to other objects, or their compactness or voxel count values were outside of the empirically estimated interval. This quality control protocol allowed to remove most of the artifacts, as confirmed by visual inspection.

3D morphometry

3D binary nuclear masks were then used for voxel-based feature extraction using scikit-image[58] Python library, obtaining feature set *V*. For each nuclear mask we measured volume, bounding box and convex hull volumes, extent, solidity, lengths of principal axes, and inertia tensor principal components. In addition to these measures, we measured mask sphericity, which can be computed as a the ratio of the surface area of a sphere (with the same volume as the given object) to the surface area of the object[59]. We first fitted a 3D ellipsoid to the binary voxel mask using linear least squares[60] and then computed the ellipsoidal sphericity Ψ as following:

$$\Psi = \frac{\pi^{\frac{1}{3}}(6V_p)^{2/3}}{SA_p} \approx \frac{(a*b*c)^{2/3}}{\left(\frac{1}{3}*((a*b)^p+(a*c)^p+(c*b)^p)\right)^{1/p}}$$

where V_p is the volume of an ellipsoid and SA_p is the surface area, computed by the approximate formula using $p \approx 1.6075$ that yields a relative error of at most 1.061%[61].

For surface-based morphometry, we first modeled boundaries of nuclear 3D masks extracted from the microscopic images as genus zero two-dimensional manifolds that are embedded as triangulated surfaces in $\mathbb{R}^3$ [40]. 3D surface modeling process employed an iterative Laplace-Beltrami eigen-projection and a topology-preserving boundary deformation algorithm[42]. This algorithm performs robust reconstruction of the objects' surfaces from their segmented masks using an iterative mask filtering process. We applied the surface reconstruction method[42] that uses the Laplace-Beltrami (LB) spectrum for outlier removal without shrinkage. The LB spectrum can be viewed as a generalization of the Fourier basis onto general surfaces. Building upon this intuitive understanding, the mesh reconstruction method[42] iteratively projects the mask boundary onto a subspace of low frequency LB eigenfunctions and removes outliers with large changes during the projection process. Surface reconstruction was done using Mask2Mesh from the MOCA framework[42] and was followed by mesh simplification and subdivision to 40,000 triangles.

This followed by the extraction of size and shape measures as features quantifying geometric characteristics of the 3D surfaces of nuclear masks, including mesh volume, surface area, curvedness, shape index, and fractal dimension implemented as the shape analysis protocol[40] in LONI Pipeline[40,62]. Another set of surface based features, including mean curvature, bounding oriented box, convex hull, cylinder, and sphere volumes, convex hull surface area, inertia tensor eigenvalues, principal axes lengths were computed using trimesh library[63]. Surface-based extent and solidity were computed from trimesh-derived features as the ratio of the object volume to the bounding box volume and the ratio of the object volume to the convex hull volume, correspondingly. All these descriptors were combined into surface-based feature set *S*.

Both volumetric and surface features we merged by inner join on the per-nucleus basis, filtering out those individual cells, for which feature extraction failed or volume measure exceeded the empirically estimated threshold. Nuclei were also considered outliers and removed if any feature value was out of the [0.05;0.95] percentile range.

Feature and model selection, and analysis

In order to handle multicollinear features, we performed hierarchical clustering on the Spearman rank-order correlations between all features and then used averaged random forest model classification performance to select a threshold for defining feature clusters. From each cluster, we selected one interpretable feature, obtaining the final set S_{11} that consisted of 11 surface-based features that included major, median, and minor principal axis lengths, nuclear and bounding cylinder volumes, sphericity, extent, solidity, mean and Gaussian curvatures, and fractal dimension.

Following sets of features were used to compare volumetric and surface-based shape representations: V, S, V_{sub}, S_{sub}, V+S, S_{11}, where V_{sub} and S_{sub} were subsets of equivalent voxel

and surface feature sets correspondingly and included nuclear, convex hull and bounding box volumes, extent, solidity, inertia eigenvalues, major axis length and sphericity. We assessed classification performance of 8 different classifiers with default hyper-parameters from with scikit-learn[64] Python library, including Gaussian naïve Bayes, k-nearest neighbors, logistic regression, linear support vector machine, random forest, extremely randomized trees, AdaBoost and gradient boosting machine. They were evaluated at each time point using the Area under the Receiving Operator Characteristic curve averaged from 10 repetitions of the internal statistical 5-fold stratified cross-validation. We averaged the performance of each model over time and all feature sets. To compare feature sets, we averaged the AUCs over all models and days. Random subsampling of the prevalent class by the total number of per-fold training samples from an underrepresented class at each iteration of the cross-validation procedure was used to combat class imbalance. Permutation feature importance was computed using *feature_importances_* property of the trained random forest model at each time point.

Univariate statistical analysis of individual features was performed using SciPy package[66] with multiple testing correction using statsmodels[67]. We reported each relative difference as the percentage change from the control group, along with a *p*-value obtained using two-sided Mann-Whitney *U* test with Holm–Šidák multiple testing correction (α=0.05), along with the difference of means (*DoM*), and the common language effect size statistic that in the case of Mann-Whitney *U* test is equivalent to the Area under the Receiving Operator Characteristic curve (*AUC*)[46].

For all analysis tasks we used Python 3.7.4 from the Anaconda distribution[68], with numpy[69], scipy[66], pandas[70], and iPython[71] packages. Visualizations and charts were built with matplotlib[72], seaborn[73], and SOCRAT[74] libraries.

Availability of Materials and Data

The documentation supporting the conclusions of this article together with the derived data, pipeline workflows, and underlying source code will be made publicly available online upon publication of the paper on the project webpage: SOCR 3D Cell Morphometry Project, http://socr.umich.edu/projects/3d-cell-morphometry.

Acknowledgements

The authors thank Samuel Handelman, Matthew O'Meara, Gilbert Omenn and many other colleagues at the Michigan Institute for Data Science (MIDAS), the Department of Computational Medicine and Bioinformatics, and the Statistics Online Computational Resource (SOCR), who provided contributions including ideas, pilot testing, improvement suggestions and other assistance. We also would like to thank colleagues at the Laboratory of Neuro Imaging (LONI), at the Keck School of Medicine, University of Southern California, for providing technical support for the LONI Pipeline environment. This work was partially supported by the National Science Foundation grants 1916425, 1734853, 1636840, 1416953, 0716055, 1023115; the National Institutes of Health grants P20 NR015331, P30 DK089503, UL1TR002240, R01CA233487, R01MH121079. A.K.K. thanks Xin Rong of the University of Michigan for the donation of the NVIDIA Titan X GPU used for this research. I.D.D. gratefully acknowledges the support of NVIDIA Corporation with the donation of the TITAN Xp GPU used for this research.

Contributions

G.V.F. and W.M. performed cell sample preparation and image acquisition. A.A.K., A.A. with input from J.Z.S. and I.D.D. designed and built modules for image segmentation, 3D shape modeling, morphometric feature extraction, and analytics. A.A.K and I.D.D. designed, built, and executed high-throughput Pipeline workflows. A.A.K and X.H. implemented and performed statistical analysis and result interpretation. J.Z.S., X.W. and I.D.D. contributed to conceptualization of study. G.A.H. and B.D.A. conceived the study. A.A.K., X.H., I.D.D. and B.D.A. wrote the

manuscript with input from other authors. All authors participated in numerous project discussions and decision making. All authors read and approved the final manuscript.

Corresponding authors

Correspondence to Brian D. Athey: bleu@umich.edu

Competing Interests

The authors declare no competing interests.

References

1. Mukherjee, R. N., Chen, P. & Levy, D. L. Recent advances in understanding nuclear size and shape. *Nucleus* **7**, 167–186 (2016).
2. Uhler, C. & Shivashankar, G. V. Nuclear Mechanopathology and Cancer Diagnosis. *Trends in Cancer* **4**, 320–331 (2018).
3. Stephens, A. D., Banigan, E. J. & Marko, J. F. Chromatin's physical properties shape the nucleus and its functions. *Current Opinion in Cell Biology* **58**, 76–84 (2019).
4. Chen, H. *et al.* Functional organization of the human 4D Nucleome. *Proceedings of the National Academy of Sciences* **112**, 8002–8007 (2015).
5. Cremer, T. *et al.* The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments. *FEBS letters* **589**, 2931–2943 (2015).
6. Higgins, G. A., Allyn-Feuer, A., Handelman, S., Sadee, W. & Athey, B. D. The epigenome, 4D nucleome and next-generation neuropsychiatric pharmacogenomics. *Pharmacogenomics* **16**, 1649–1669 (2015).
7. Zheng, G. *et al.* Hypothesis: Caco-2 cell rotational 3D mechanogenomic Turing patterns has clinical implications to colon crypts. *Journal of Cellular and Molecular Medicine* (2018) doi:10.1111/jcmm.13853.
8. Göttlicher, M. *et al.* Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *The EMBO journal* **20**, 6969–6978 (2001).
9. Yang, X. & Seto, E. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* **26**, 5310 (2007).

10. Ganai, S. A., Malli Kalladi, S. & Mahadevan, V. HDAC inhibition through valproic acid modulates the methylation profiles in human embryonic kidney cells. *Journal of Biomolecular Structure and Dynamics* **33**, 1185–1197 (2015).
11. Stephens, A. D. *et al.* Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *MBoC* **29**, 220–233 (2018).
12. Zink, D., Fischer, A. H. & Nickerson, J. A. Nuclear structure in cancer cells. *Nature reviews cancer* **4**, (2004).
13. Veltri, R. W. & Christudass, C. S. Nuclear Morphometry, Epigenetic Changes, and Clinical Relevance in Prostate Cancer. in *Cancer Biology and the Nuclear Envelope: Recent Advances May Elucidate Past Paradoxes* (eds. Schirmer, E. C. & de las Heras, J. I.) 77–99 (Springer New York, 2014).
14. Grys, B. T. *et al.* Machine learning and computer vision approaches for phenotypic profiling. *J Cell Biol* **216**, 65–71 (2017).
15. Boyd, J., Fennell, M. & Carpenter, A. Harnessing the power of microscopy images to accelerate drug discovery: what are the possibilities? *Expert Opinion on Drug Discovery* **0**, 1–4 (2020).
16. Marchion, D. C., Bicaku, E., Daud, A. I., Sullivan, D. M. & Munster, P. N. Valproic Acid Alters Chromatin Structure by Regulation of Chromatin Modulation Proteins. *Cancer Res* **65**, 3815–3822 (2005).
17. Biggs, R., Liu, P. Z., Stephens, A. D. & Marko, J. F. Effects of altering histone posttranslational modifications on mitotic chromosome structure and mechanics. *MBoC* **30**, 820–827 (2019).

18. Ververis, K., Hiong, A., Karagiannis, T. C. & Licciardi, P. V. Histone deacetylase inhibitors (HDACIs): multitargeted anticancer agents. *Biologics: targets & therapy* **7**, 47 (2013).

19. Eckschlager, T., Plch, J., Stiborova, M. & Hrabeta, J. Histone deacetylase inhibitors as anticancer drugs. *International journal of molecular sciences* **18**, 1414 (2017).

20. Williams, Aaron M. & Hsu, Cindy H. Histone Deacetylase Inhibitors: A Novel Strategy for Neuroprotection and Cardioprotection Following Ischemia/Reperfusion Injury. *Journal of the American Heart Association* **0**, e016349 (2020).

21. Hsieh, J., Nakashima, K., Kuwabara, T., Mejia, E. & Gage, F. H. Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells. *Proceedings of the National Academy of Sciences* **101**, 16659–16664 (2004).

22. Zhang, L. *et al.* Small molecules efficiently reprogram human astroglial cells into functional neurons. *Cell stem cell* **17**, 735–747 (2015).

23. Gao, L. *et al.* Direct generation of human neuronal cells from adult astrocytes by small molecules. *Stem cell reports* **8**, 538–547 (2017).

24. Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnology* **26**, 795–797 (2008).

25. Jang, S. & Jeong, H.-S. Histone deacetylase inhibition-mediated neuronal differentiation via the Wnt signaling pathway in human adipose tissue-derived mesenchymal stem cells. *Neuroscience letters* **668**, 24–30 (2018).

26. Goldman, S. A. Stem and progenitor cell-based therapy of the central nervous system: hopes, hype, and wishful thinking. *Cell Stem Cell* **18**, 174–188 (2016).

27. Li, H. & Chen, G. In vivo reprogramming for CNS repair: regenerating neurons from endogenous glial cells. *Neuron* **91**, 728–738 (2016).

28. Wang, L.-L. & Zhang, C.-L. Engineering new neurons: in vivo reprogramming in mammalian brain and spinal cord. *Cell and tissue research* **371**, 201–212 (2018).

29. Higgins, G. A. *et al.* Network Reconstruction Reveals that Valproic Acid Activates Neurogenic Transcriptional Programs in Adult Brain Following Traumatic Injury. *Pharm Res* **34**, 1658–1672 (2017).

30. Clarke, L. E. & Barres, B. A. Emerging roles of astrocytes in neural circuit development. *Nature Reviews Neuroscience* **14**, 311 (2013).

31. Bayraktar, O. A., Fuentealba, L. C., Alvarez-Buylla, A. & Rowitch, D. H. Astrocyte Development and Heterogeneity. *Cold Spring Harb Perspect Biol* **7**, a020362 (2015).

32. Amamoto, R. & Arlotta, P. Development-inspired reprogramming of the mammalian central nervous system. *Science* **343**, 1239882 (2014).

33. Cheng, L. *et al.* Direct conversion of astrocytes into neuronal cells by drug cocktail. *Cell research* **25**, 1269 (2015).

34. Qin, H., Zhao, A. & Fu, X. Small molecules for reprogramming and transdifferentiation. *Cell. Mol. Life Sci.* **74**, 3553–3575 (2017).

35. Yin, J.-C. *et al.* Chemical Conversion of Human Fetal Astrocytes into Neurons through Modulation of Multiple Signaling Pathways. *Stem Cell Reports* **12**, 488–501 (2019).

36. Kortenhorst, M. S. *et al.* Valproic acid causes dose-and time-dependent changes in nuclear structure in prostate cancer cells in vitro and in vivo. *Molecular cancer therapeutics* **8**, 802–808 (2009).

37. Felisbino, M. B., Gatti, M. S. V. & Mello, M. L. S. Changes in Chromatin Structure in NIH 3T3 Cells Induced by Valproic Acid and Trichostatin A. *Journal of Cellular Biochemistry* **115**, 1937–1947 (2014).

38. Felisbino, M. B., Costa, T. A. da, Gatti, M. S. V. & Mello, M. L. S. Differential Response of Human Hepatocyte Chromatin to HDAC Inhibitors as a Function of Microenvironmental Glucose Level. *Journal of Cellular Physiology* **231**, 2257–2265 (2016).

39. Kalinin, A. A. *et al.* 3D Cell Nuclear Morphology: Microscopy Imaging Dataset and Voxel-Based Morphometry Classification Results. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* 2272–2280 (2018).

40. Kalinin, A. A. *et al.* 3D Shape Modeling for Cell Nuclear Morphological Analysis and Classification. *Scientific Reports* **8**, 13658 (2018).

41. Al-Kofahi, Y., Lassoued, W., Lee, W. & Roysam, B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Transactions on Biomedical Engineering* **57**, 841–852 (2010).

42. Shi, Y. *et al.* Robust surface reconstruction via Laplace-Beltrami eigen-projection and boundary deformation. *IEEE transactions on medical imaging* **29**, 2009–2022 (2010).

43. Breiman, L. Random Forests. *Machine Learning* **45**, 5–32 (2001).

44. van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **9**, 2579–2605 (2008).

45. Wattenberg, M., Viégas, F. & Johnson, I. How to Use t-SNE Effectively. *Distill* **1**, e2 (2016).

46. Mason, S. J. & Graham, N. E. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society* **128**, 2145–2166 (2002).

47. Cohen-Steiner, D. & Morvan, J.-M. Restricted delaunay triangulations and normal cycle. in *Proceedings of the nineteenth annual symposium on Computational geometry* 312–321 (Association for Computing Machinery, 2003). doi:10.1145/777792.777839.

48. Choi, H.-J. & Choi, H.-K. Grading of renal cell carcinoma by 3D morphological analysis of cell nuclei. *Computers in Biology and Medicine* **37**, 1334–1341 (2007).

49. Meyer, M. G. *et al.* Automated cell analysis in 2D and 3D: A comparative study. *Pattern Recognition* **42**, 141–146 (2009).

50. Pincus, Z. & Theriot, J. A. Comparison of quantitative methods for cell-shape analysis. *Journal of Microscopy* **227**, 140–156 (2007).

51. Dufour, A. C. *et al.* Signal processing challenges in quantitative 3-d cell morphology: More than meets the eye. *IEEE Signal Processing Magazine* **32**, 30–40 (2015).

52. Kortenhorst, M. S., Carducci, M. A. & Shabbeer, S. Acetylation and histone deacetylase inhibitors in cancer. *Analytical Cellular Pathology* **28**, 191–222 (2006).

53. Yi, J., Stypula-Cyrus, Y., Blaha, C. S., Roy, H. K. & Backman, V. Fractal Characterization of Chromatin Decompaction in Live Cells. *Biophysical Journal* **109**, 2218–2226 (2015).

54. Metze, K., Adam, R. & Florindo, J. B. The fractal dimension of chromatin - a potential molecular marker for carcinogenesis, tumor progression and prognosis. *Expert Review of Molecular Diagnostics* **19**, 299–312 (2019).

55. Bray, M.-A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nature Protocols* **11**, 1757–1774 (2016).

56. Lieberman-Aiden, E. *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**, 289–293 (2009).

57. Kalinin, A. A. *et al.* Deep learning in pharmacogenomics: from gene regulation to patient stratification. *Pharmacogenomics* **19**, 629–650 (2018).

58. van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).

59. Wadell, H. Volume, shape, and roundness of quartz particles. *The Journal of Geology* **43**, 250–280 (1935).

60. Fitzgibbon, A., Pilu, M. & Fisher, R. B. Direct least square fitting of ellipses. *IEEE Transactions on pattern analysis and machine intelligence* **21**, 476–480 (1999).

61. Xu, D. *et al.* The ellipsoidal area ratio: an alternative anisotropy index for diffusion tensor imaging. *Magnetic resonance imaging* **27**, 311–323 (2009).

62. Dinov, I. *et al.* Efficient, distributed and interactive neuroimaging data analysis using the LONI pipeline. *Frontiers in neuroinformatics* **3**, 22 (2009).

63. Dawson-Haggerty et al. *trimesh*. (2019).

64. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

65. Dinov, I. & Christou, N. *Probability and statistics EBook*. (SOCR Resource, 2010).

66. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17**, 261–272 (2020).

67. Seabold, S. & Perktold, J. statsmodels: Econometric and statistical modeling with python. in *9th Python in Science Conference* (2010).

68. Continuum Analytics & others. Anaconda software distribution. *Computer software. Vers* 2–2 (2016).

69. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science Engineering* **13**, 22–30 (2011).

70. McKinney, W. Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference* 56–61 (2010) doi:10.25080/Majora-92bf1922-00a.

71. Perez, F. & Granger, B. E. IPython: A System for Interactive Scientific Computing. *Computing in Science Engineering* **9**, 21–29 (2007).

72. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* **9**, 90–95 (2007).

73. Michael Waskom *et al. mwaskom/seaborn: v0.10.1 (April 2020)*. (Zenodo, 2020). doi:10.5281/zenodo.3767070.

74. Kalinin, A. A., Palanimalai, S. & Dinov, I. D. SOCRAT Platform Design: A Web Architecture for Interactive Visual Analytics Applications. in *Proceedings of the 2nd Workshop on Human-In-the-Loop Data Analytics* 8:1–8:6 (ACM, 2017). doi:10.1145/3077257.3077262.

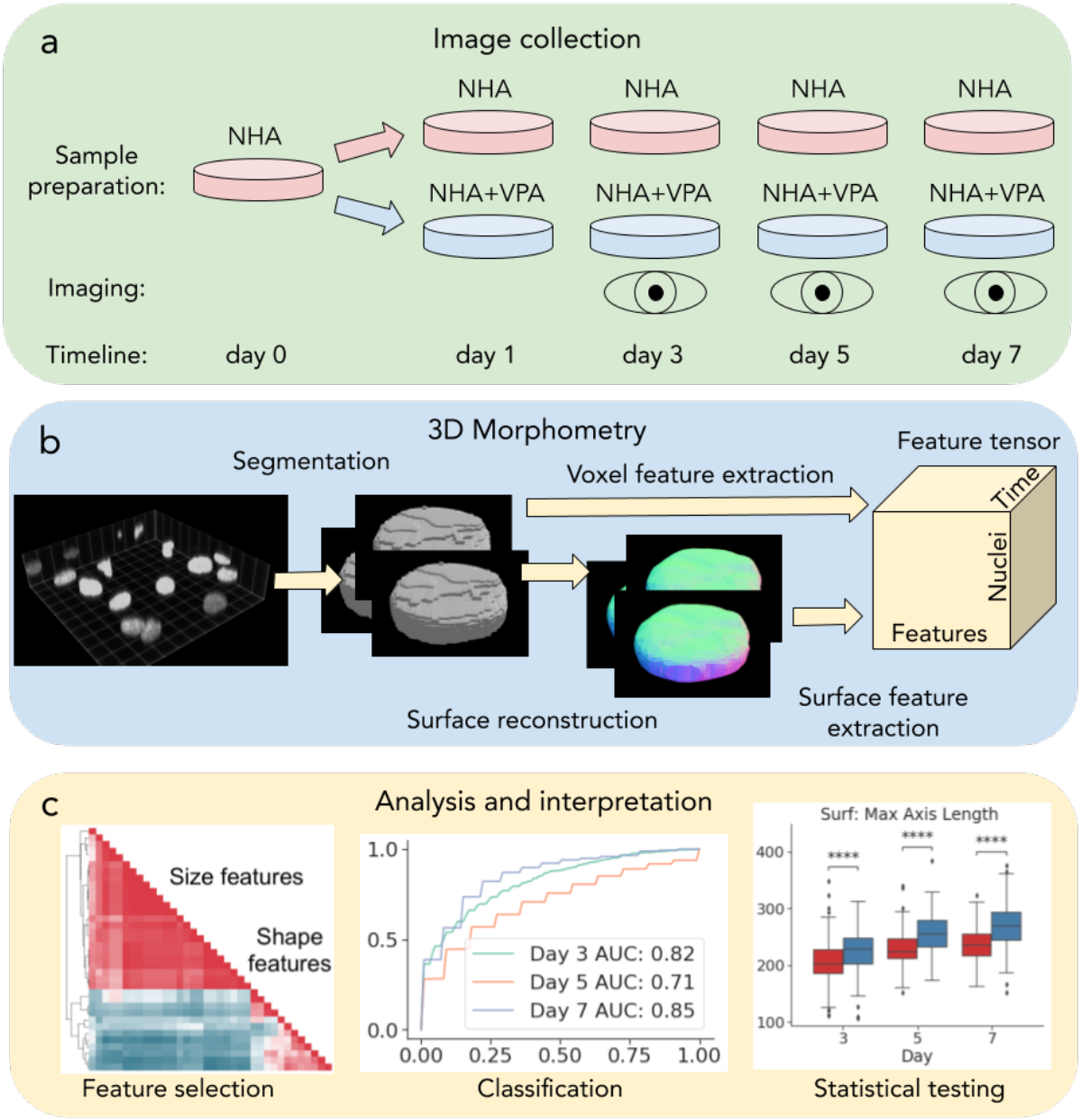

Figure 1. A schematic overview of our approach: (a) sample preparation, treatment, and imaging; (b) 3D nuclear segmentation, shape modeling, and feature extraction; (c) feature selection, univariate statistical and machine learning analysis.

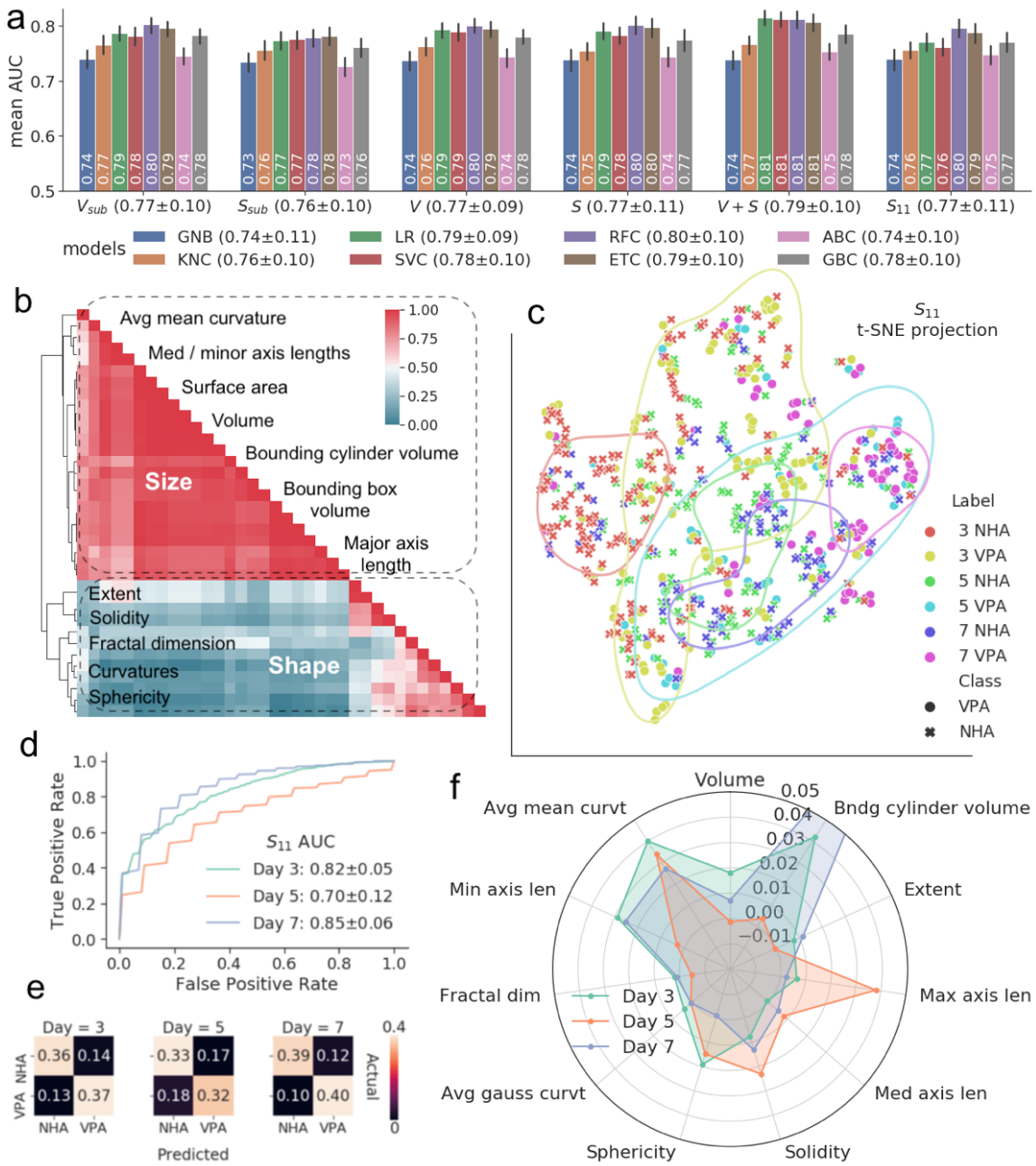

Figure 2. Classifier and feature set selection, and classification performance: (a) mean AUCs (with 95% bootstrapped confidence intervals) over days 3, 5, 7 for different classifiers (GNB–Gaussian naïve Bayes; KNC–k nearest neighbors; LR–logistic regression; SVC–support vector machine; RFC–random forest; ETC–extra-randomized trees; ABC–AdaBoost; GBC–gradient boosting) and feature sets (*V*–volumetric features; *S*–surface-based features; V_{sub} and S_{sub}–subsets of 10 common features extracted from voxels and surfaces, correspondingly; S_{11}–the

subset of 11 surface-based features selected with hierarchical clustering); (b) hierarchical clustering of all voxel and surface features (*V*+*S*), showing representative size and shape descriptors; (c) 2D t-SNE projection of the S_{11} feature space showing NHA and VPA-treated nuclear morphological clustering over time; (d) ROC curves for the random forest model with S_{11} features on days 3, 5, and 7; (e) average normalized confusion matrices for the random forest model on the S_{11} features; (f) permutation importance of S_{11} features for distinguishing morphologies on each day.

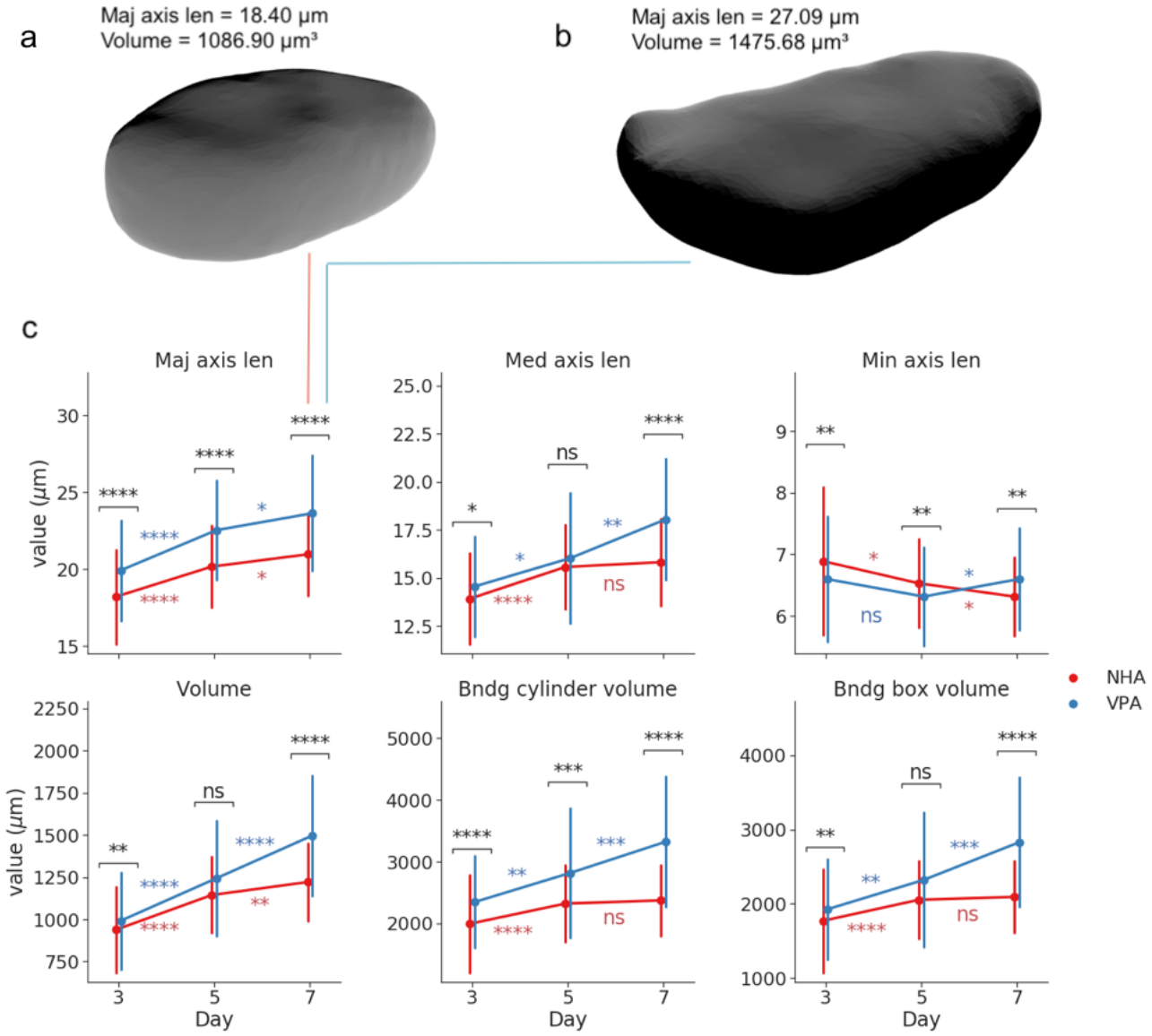

Figure 3. Visualization and univariate statistical analysis of size changes under VPA treatment. (a) – reconstructed surface of a representative NHA nucleus on day 7. (b) – reconstructed surface of a representative VPA nucleus on day 7. (c) – time-dependent changes in morphometric measures of nuclear sizes (error bars show standard deviation; *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, Holm–Šidák-corrected two-sided Mann-Whitney U test).

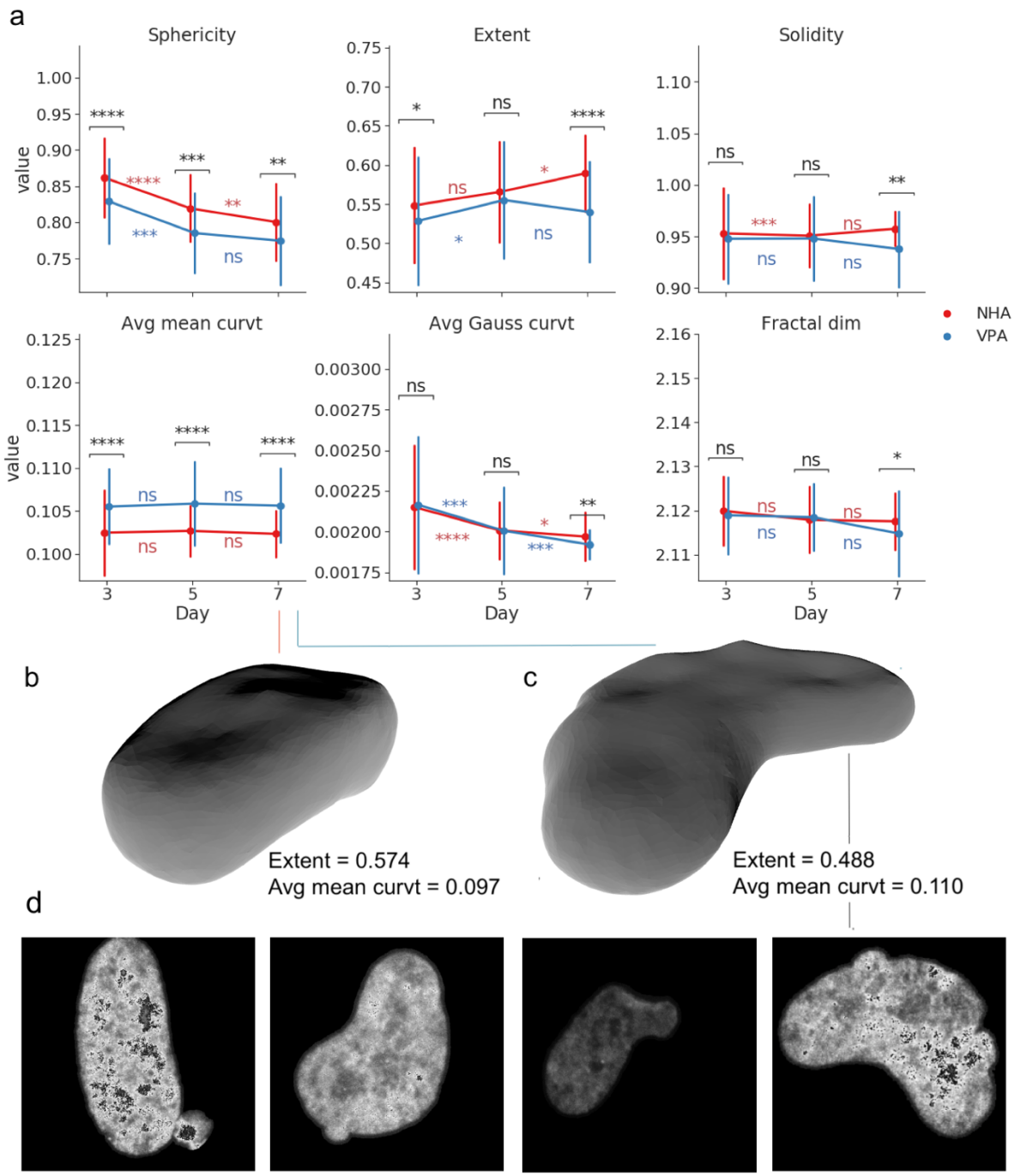

Figure 4. Visualization and univariate statistical analysis of shape changes under VPA treatment. (a) – reconstructed surface of a representative NHA nucleus on day 7. (b) – reconstructed surface of a representative VPA nucleus on day 7. (c) – time-dependent changes in morphometric measures of nuclear shapes (error bars show standard deviation; *p<0.05, **p<0.01, ***p<0.001,

****p<0.0001, Holm–Šidák-corrected two-sided Mann-Whitney U test). (d) – exemplar 2D projections of irregular shaped nuclei as a result of VPA treatment on day 7.