Electronic Supplementary Material (ESI) for Digital Discovery.

Supplementary Material

An Open-Source Environmental Chamber for Materials-Stability Testing

Rodolfo Keesey[§,1], Armi Tiihonen[§,2,3], Alexander E. Siemenn[3], Thomas W. Colburn[4], Shijing Sun[3], Noor Titan Putri Hartono[3], James Serdy[3], Margaret Zeile[1], Keqing He[1], Cole A. Gurtner[4], Austin C. Flick[4], Clio Batali[3], Alex Encinas[3], Richa R. Naik[3], Zhe Liu[3], Felipe Oviedo[3], I. Marius Peters[3], Janak Thapa[3], Siyu Isaac Parker Tian[5], Reinhold H. Dauskardt[4], Alexander J. Norquist [1], Tonio Buonassisi[,3]

[§] Contributed equally
[1] Haverford College, 370 Lancaster Avenue, Haverford, 19041, Pennsylvania, USA
[2] Department of Applied Physics, Aalto University, P.O. Box 11100, Aalto, FI-00076, Finland
[3] Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4] Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305, USA
[5] Low Energy Electronic Systems (LEES), Singapore-MIT Alliance for Research and Technology (SMART), Singapore 138602, Singapore

Section S1: Evidence for correlation of color change and efficiency based on Hashmi et al. [40]

Color and efficiency change of the $MAPbI_3$-based carbon perovskite solar cells presented in Hashmi et al. in [40] is correlated, as shown in the reproduced illustration of the original work (Figure S1). The timepoints of measuring efficiency and color are different in the original work, preventing the direct plotting of color change as a function of efficiency. To investigate correlation directly, the efficiency data shown in Figure S1 and published in [40] was interpolated to the times of color measurements with linear interpolation between the two nearest measured efficiency points.

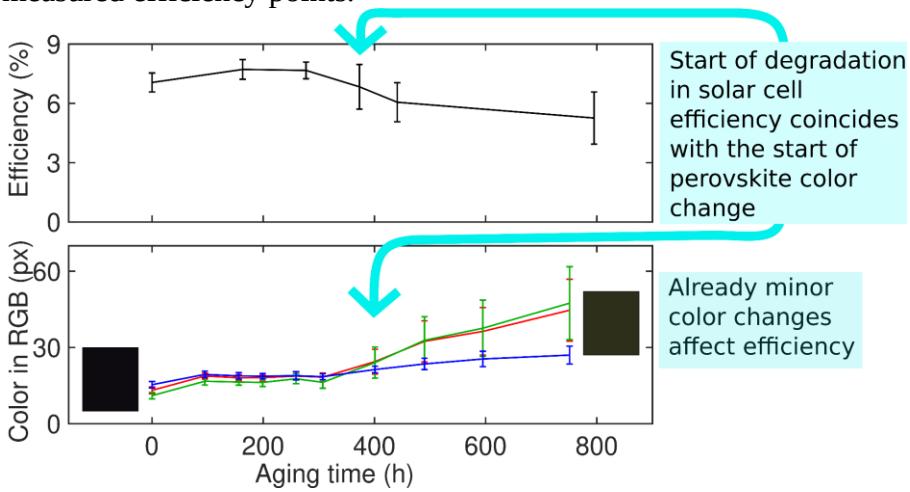

Figure S1. Efficiencies of $MAPbI_3$-based perovskite solar cells and the colors of the perovskite layers measured by Hashmi *et al*. in [40]. This data is used as the basis for Figure 2a. The colored patches represent the sample color before and after the aging test. Figures reprinted and adapted upon request to *Journal of Materials Chemistry A*.

Section S2: Supplementary Figures

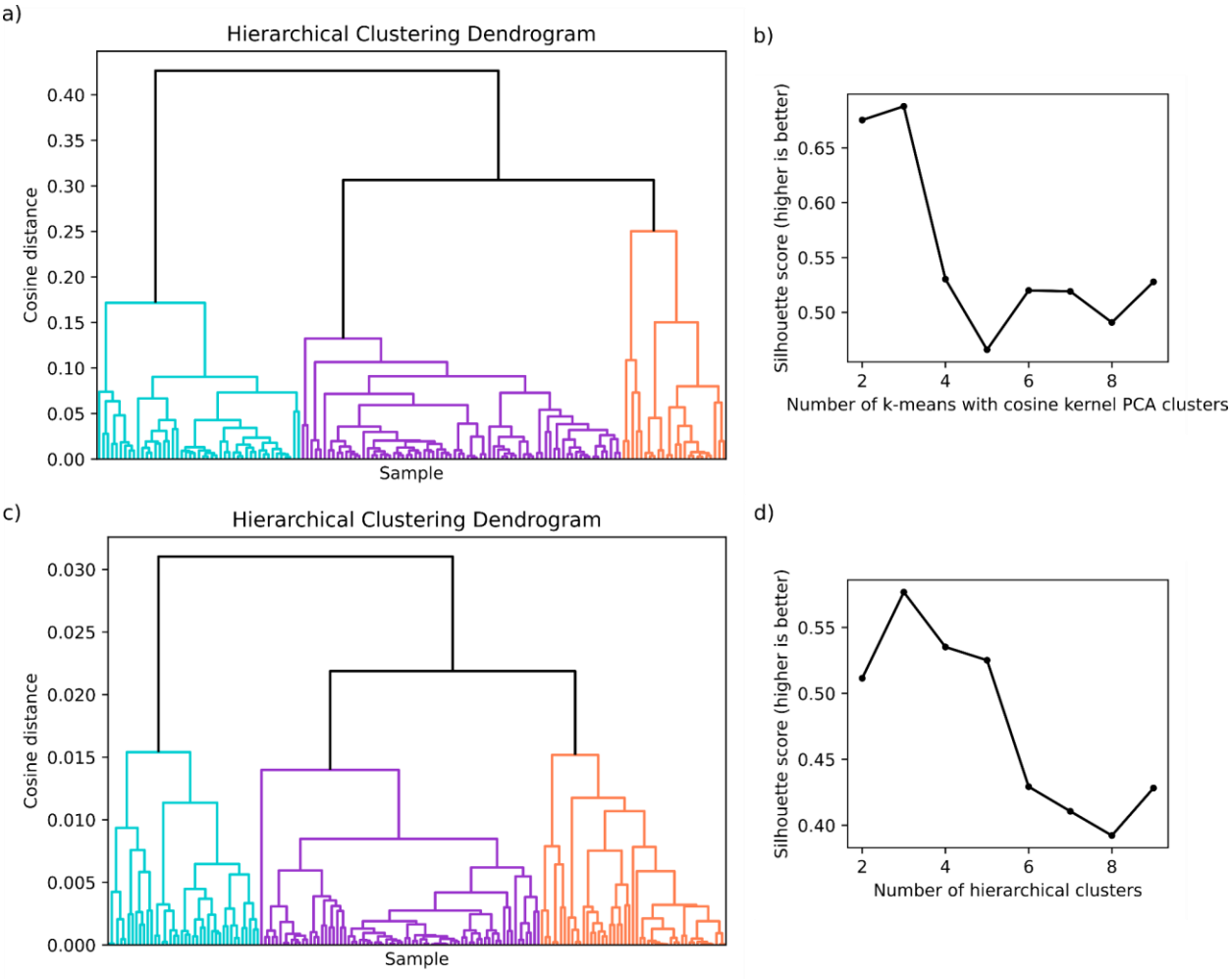

Figure S2. Hierarchical clustering of XRD (a-b) and camera RGB (c-d) data using raw data by Sun *et al*. [43]. a), c) Dendrograms of the raw data and b) , d) silhouette scores of hierarchical clustering repeated with different numbers of clusters suggest there are three distinct XRD spectra types (a-b) and degradation behaviors (c-d) among the samples investigated.

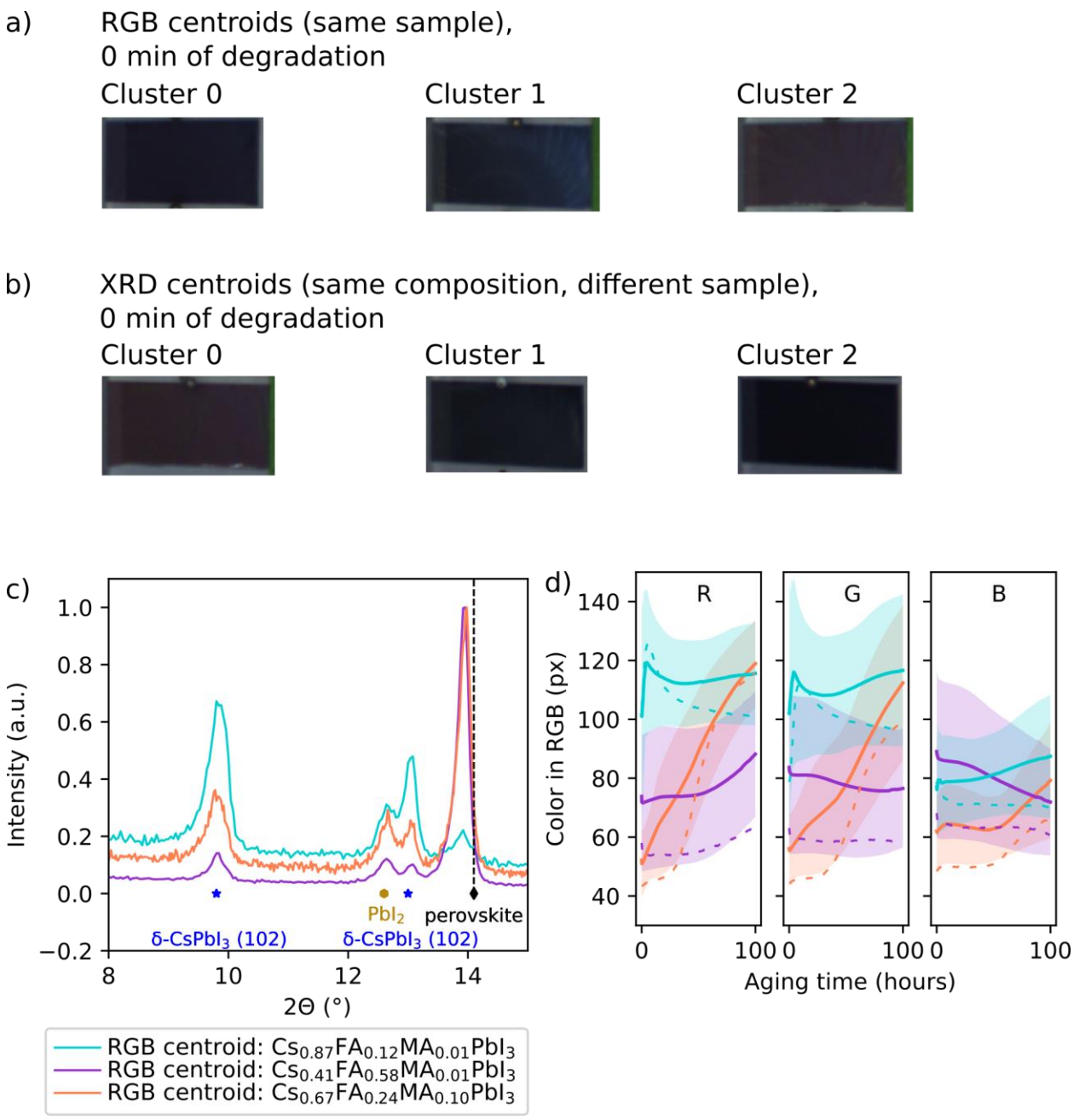

Figure S3. a) RGB and b) XRD cluster centroid samples compared via photographs at the beginning of the aging tests in the environmental chamber. Pictures are raw data (not color calibrated) from the degradation chamber. All the shown samples are dark at 0min, but nevertheless, differing phases are seen in the c) XRD measurements of the RGB centroids, and differing degradation patterns are visible in the d) aging test of the XRD centroids. The cross comparisons show that the XRD spectra of the RGB centroids represent similar differences than the XRD spectra of the XRD centroids, and the RGB curves of the XRD centroids similar differences than the RGB curves of the RGB clusters, *i.e.*, XRD and RGB clusters are similar.

Section S3: Chamber specifications

Chamber specifications are shown in **Table S1**. Temperature and humidity ranges have been measured with the sensors integrated into each chamber. For Stanford Gen. 2 chamber, the intensity measurements were taken on a solar reference calibrated OES system (StellarNet BLACK-Comet Concave Grading Spectrometer, S/N: 13071712, with an F600-UVVIS-SR fiber optic cable and CR2 Cosine Polymer Diffuser and CR2-AP 10% aperture) and then compared with an InfinityPV IsoSun, No.:Q1-19-AH as a standard AM1.5 solar simulator calibrated with a silicon photodiode standard to give relative light intensities from 280 – 849.5 nm with a 0.5 nm step size. Trapezoidal integration was utilized to calculate the intensity for the environmental chamber white LEDs and the InfinityPV IsoSun system to allow for direct comparison given the difference in spectral output from the white LEDs compared to an AM 1.5 source. For MIT Gen. 1 and Haverford Gen. 2 chambers, the intensity was measured via the short circuit current of a photodiode with known spectral response. The current was transformed to the illumination intensity experienced by a typical $MAPbI_3$ perovskite solar cell under AM1.5G by utilizing the spectral sensitivities of the photodiode and a typical perovskite solar cell, AM1.5G spectrum, and the spectrum of the lamp. The intensity ranges for MIT Gen. 2 chamber are as declared by the manufacturer.

Chamber	*Sample Type*	*Trackable Samples*	*Sample Temperature Range*	*Ambient Temperature Range*	*Humidity Range*	*Light Range*
MIT Gen. 1	Thin Films	28	Ambient – 100C	Uncontrolled	Ambient – 85 ± 5%	0.15 ± 0.01 sol
Haverford Gen. 2	Bulk Crystals	9	Ambient – 100C	Uncontrolled	Ambient – 85 ± 5%	0.15 ± 0.01 sol
MIT Gen. 2	Thin Films	28	Ambient – 100C	Ambient – 90C	15 – 80 ± 3% (dewpoint control)	0.1 – 1.1 sol
Stanford Gen. 2	Full Devices	36	Ambient – 300C	Uncontrolled	< 0.0%	2.0E-4 – 0.43 sol *

Table S1. The main specifications of each chamber. The built chamber has to be calibrated before and inspected frequently during use to confirm that the listed values are reached. Misaligned or aged components may cause significant variations to the values.

*0.0002 Sun with the lamp drive current of 75 mA corresponding to the designed use in studying thermal degradation and minimizing the photodegradation, 0.003 Sun with the maximum drive current of 750 mA with the designed chamber geometry, and 0.43 Sun when measured directly under the lamps with the maximum drive current.

Section S4: Calibration Data for Chambers

Typical examples of chamber calibration data are represented in **Figures S4-S7**. It should be noted that this level of uniformity of conditions is reached only in well maintained, frequently calibrated chambers. As conditions across the chamber are not fully uniform, it is good practice to test new sample types in duplicate to ensure the sufficient uniformity of the conditions across the sample holder. Additionally, sample locations on the sample holder should be randomized to avoid location bias.

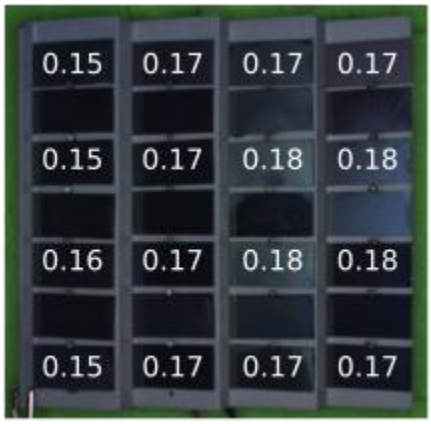

Figure S4. Exemplary light intensity calibration result of MIT Gen. 1 chamber. Intensities shown as Suns experienced by $MAPbI_3$ solar cells.

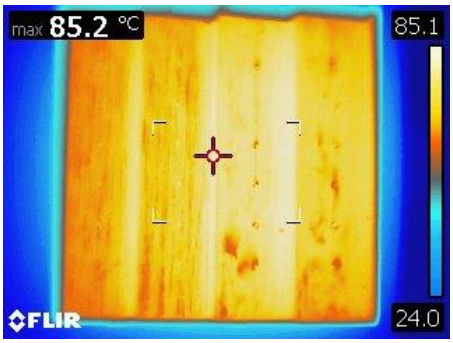

Figure S5. Uniformity of the temperature across the sample holder in MIT Gen. 1 chamber measured by FLIR infrared camera (emissivity 1.0).

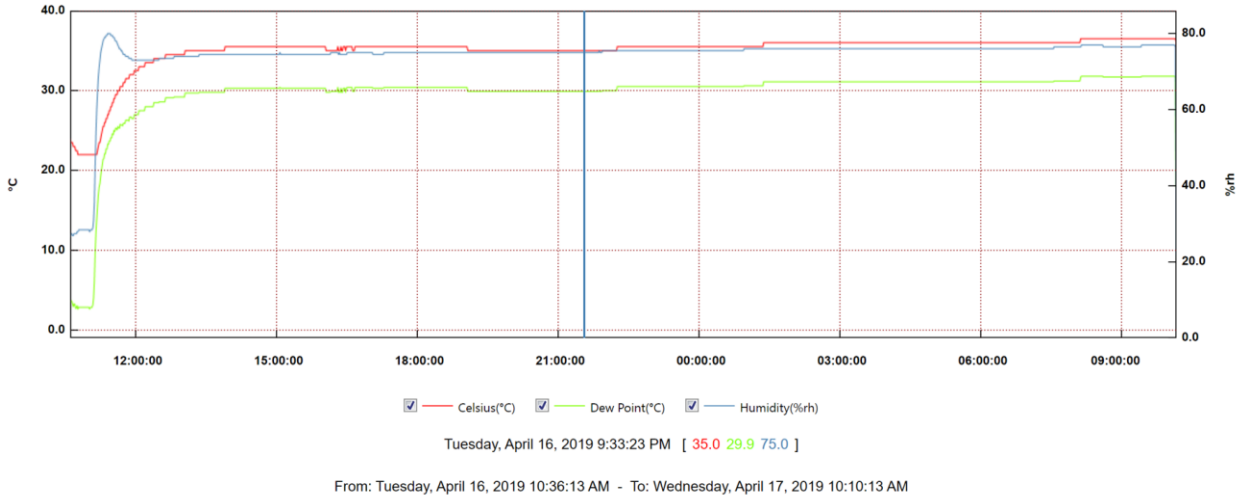

Figure S6. Longitudinal stability of air humidity and temperature in MIT Gen. 1 chamber during a 23.5-hour aging test. Relative air humidity (%rh) shown with blue color, temperature (°C) shown with red color, and dew point (%rh) shown with green color. Calibration graph is a direct output from the control software of Lascar Electronics EL–USB–2 humidity tracker integrated into the chamber.

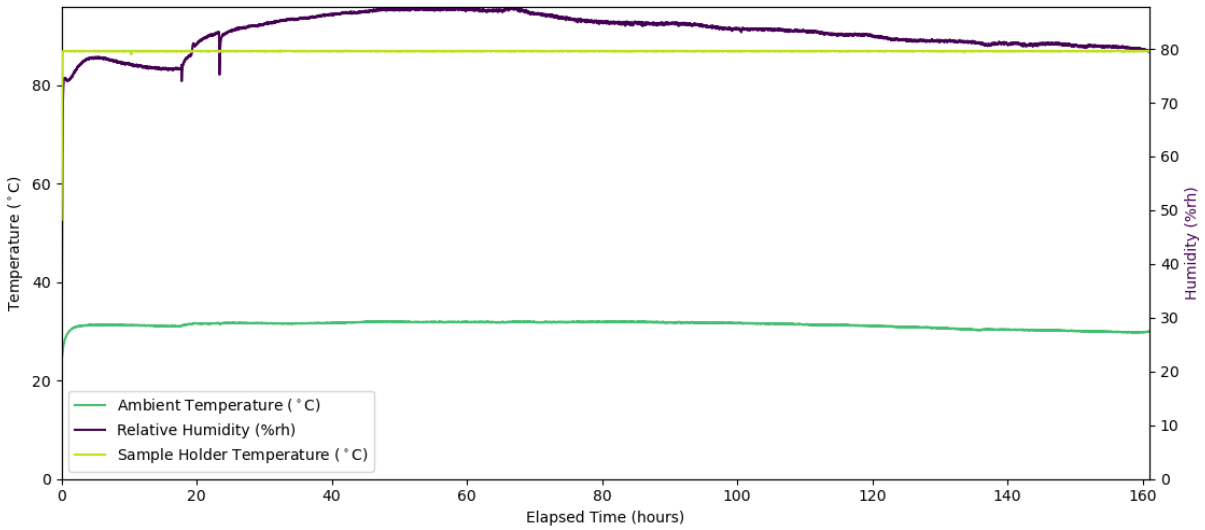

Figure S7. Longitudinal stability of air humidity, chamber ambient temperature, and sample holder temperature in the Haverford Gen. 2 chamber during the aging of the reported bulk perovskite samples. Relative air humidity (%rh) is shown in purple, chamber ambient temperature (°C) is shown in green, and sample holder temperature (°C) is shown in yellow. The chamber is preheated for the first 24 hours. The two humidity drops during the preheat period are due to humidity water-reservoir refilling, and sample loading, respectively.

Section S5: Cluster Number Analysis for Hierarchical Clustering

Silhouette analysis on the composition plane and Davies-Bouldin index on the composition plane for XRD and RGB data shown in **Figures S8-S9**, respectively. Negative silhouette coefficient values suggest the sample has been clustered to a wrong cluster, and overall high silhouette coefficient values (preferably higher than the average silhouette score of the whole dataset marked with a red dashed line) suggest the cluster division leads to well-separated clusters. Davies-Bouldin index suggests two or three clusters would be a suitable number of clusters for the XRD data, and the silhouette analysis suggests 3 clusters, so three is selected. For RGB data, both methods suggest three clusters.

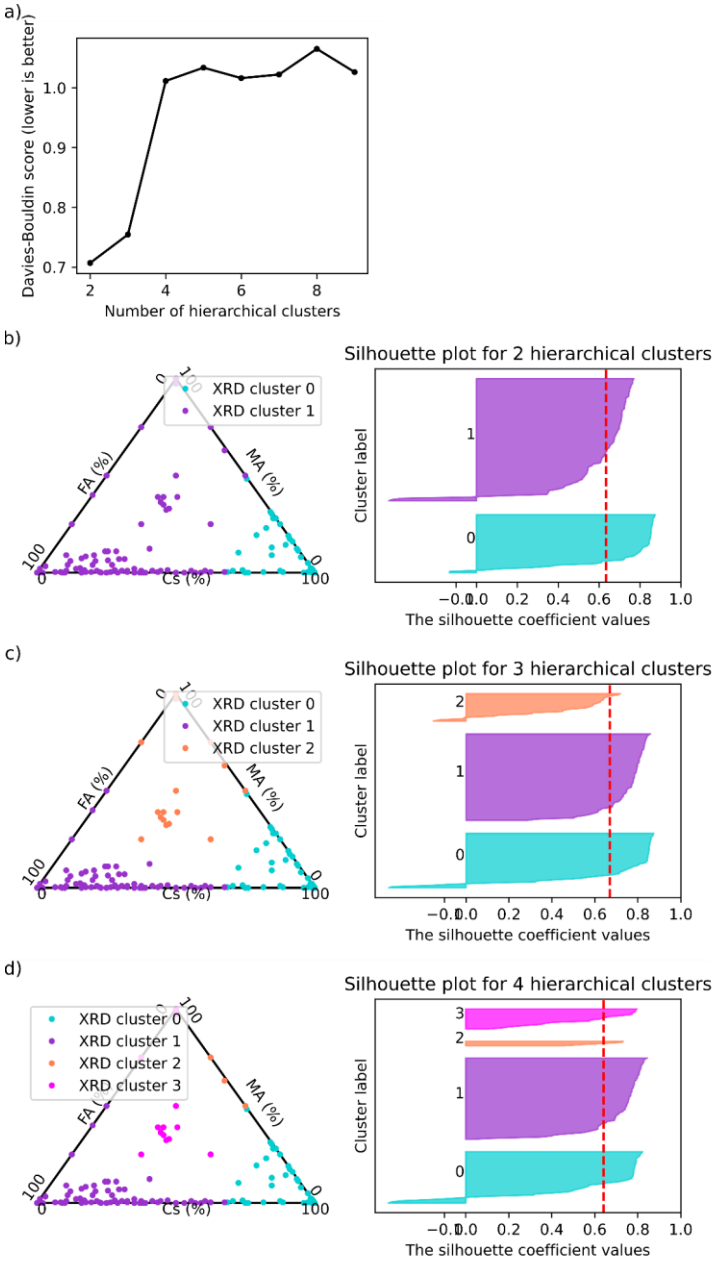

Figure S8. Clustering robustness analysis for the XRD data. a) Davies-Bouldin index of the clusters on the compositional plane. Silhouette analysis is repeated for b) two, c) three, and d) four clusters. Within each cluster, the individual samples are in the silhouette plot stacked to the y axis. The average silhouette score of the whole dataset is marked with a red dashed line.

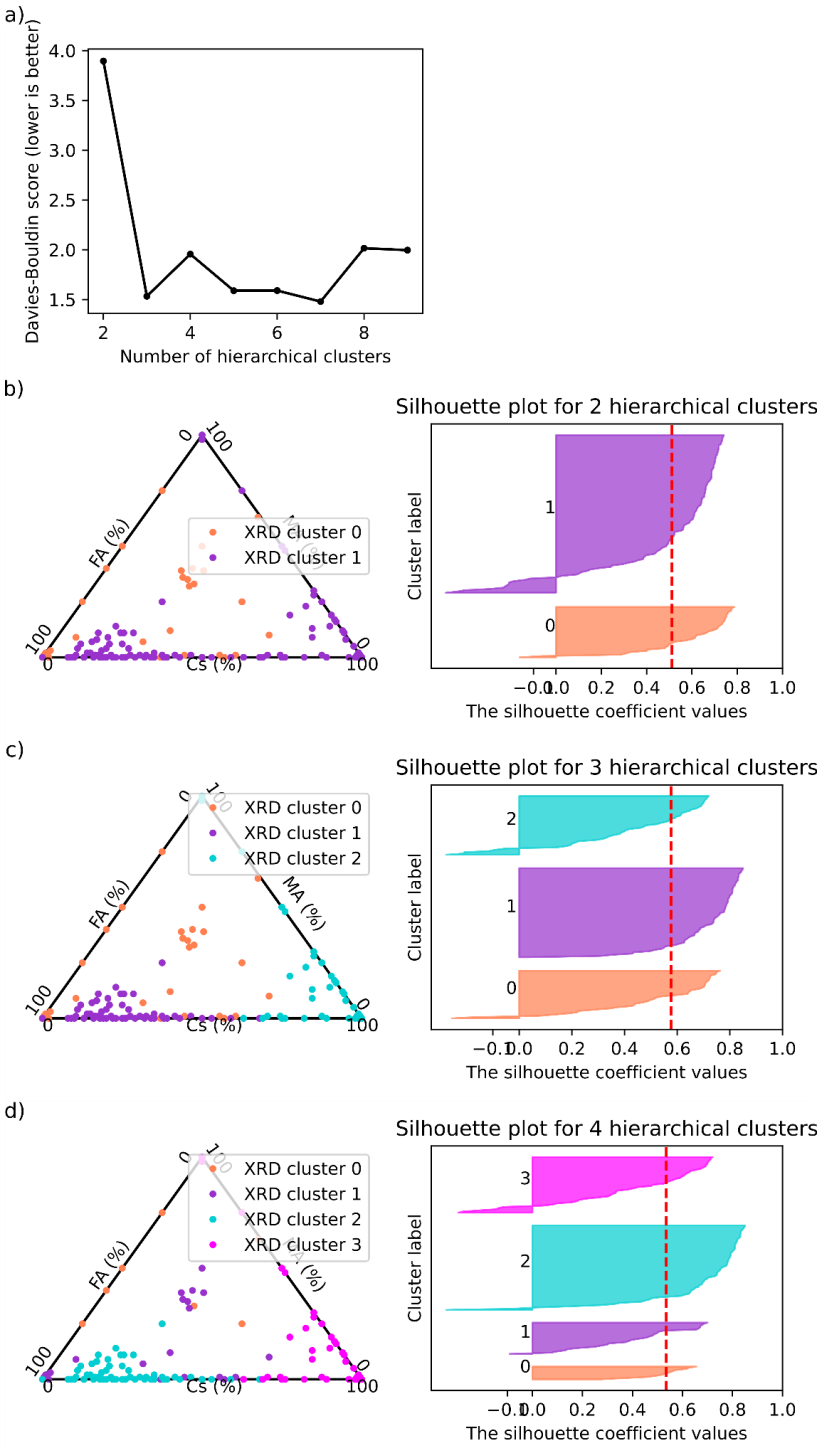

Figure S9. Cluster number analysis for the RGB data. a) Davies-Bouldin index of the clusters on the compositional plane. Silhouette analysis is repeated for b) two, c) three, and d) four clusters. Within each cluster, the individual samples are in the silhouette plot stacked to the y axis. The average silhouette score of the whole dataset is marked with a red dashed line.

Section S6: Clustering Robustness with K-means Clustering

The selection of the number of the clusters and the clustering are repeated with k-means algorithm in order to confirm the robustness of the clustering. The results shown in **Figures S10-S11** are similar to those gained with hierarchical clustering.

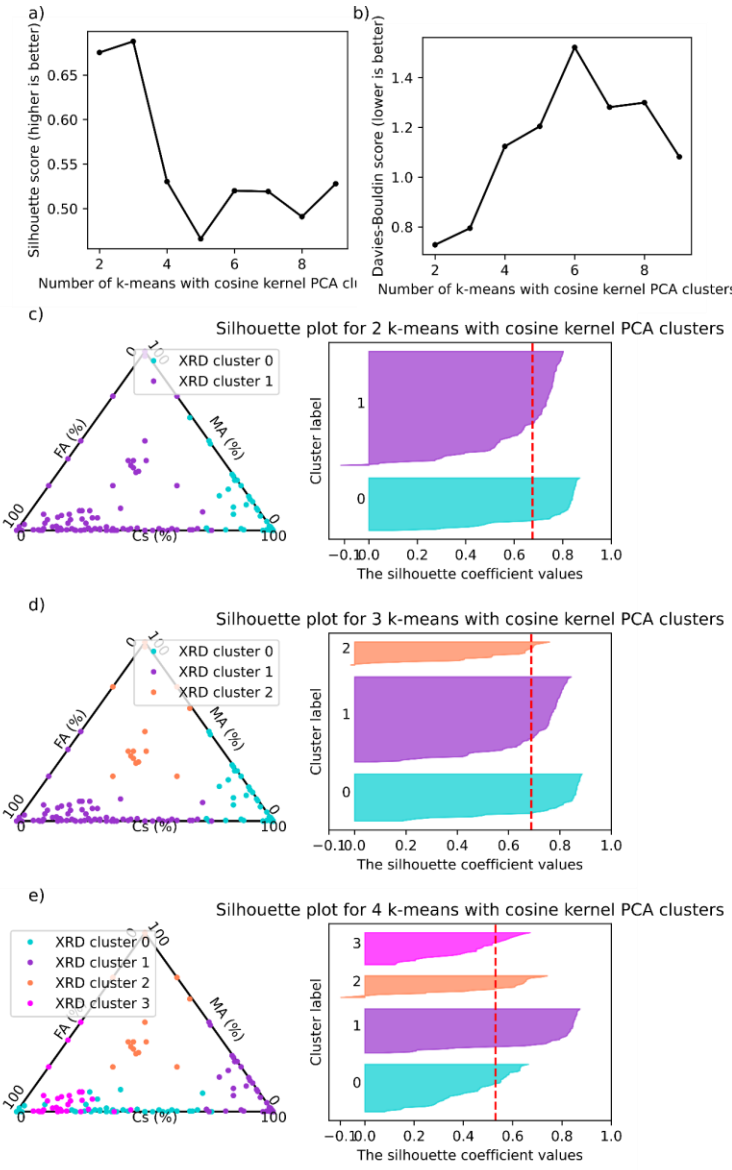

Figure S10. K-means clustering of the XRD data as a reference method for investigating the robustness of clustering. a) The average silhouette score with cosine metric on the XRD plane, b) Davies-Bouldin index on the composition plane, and sample-wise silhouette analysis shown for c) two, d) three, and e) four clusters. Within each cluster, the individual samples are in the silhouette plot stacked to the y axis. The average silhouette score of the whole dataset is marked with a red dashed line.

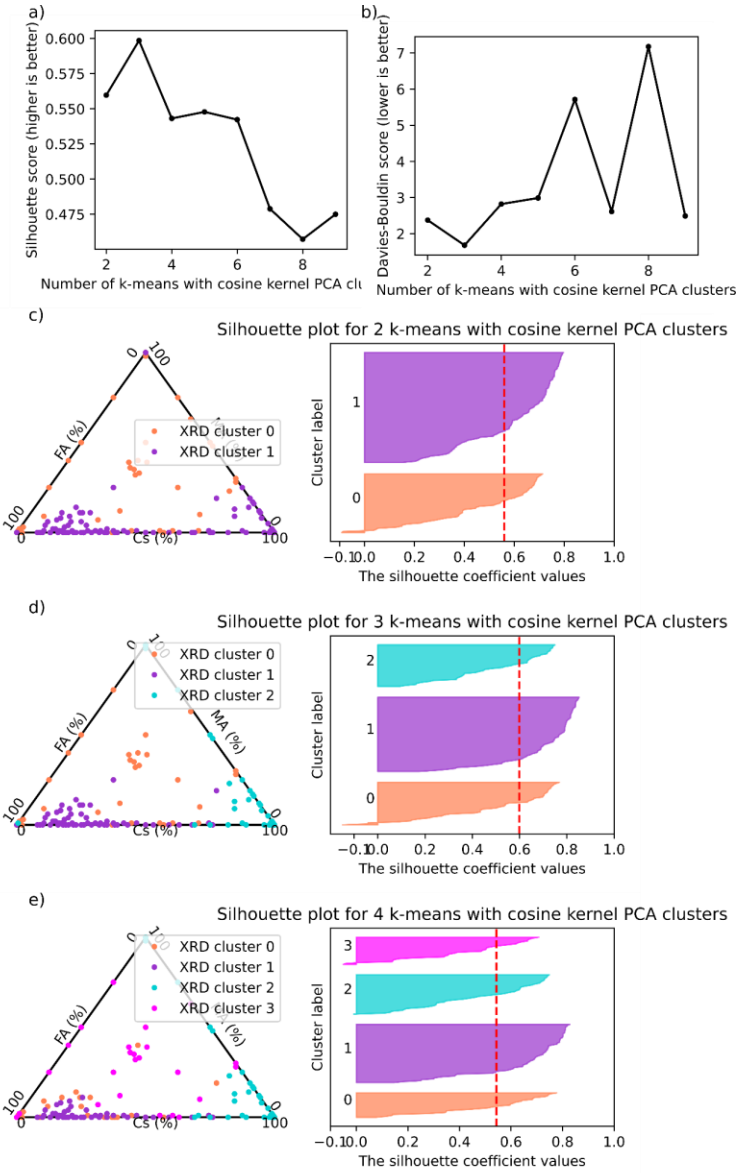

Figure S11. K-means clustering of the RGB data as a reference method for investigating the robustness of the clustering. Silhouette analysis with a) average silhouette score, and sample-wise analysis shown for b) two, c) three, and d) four clusters. Within each cluster, the individual samples are in the silhouette plot stacked to the y axis. The average silhouette score of the whole dataset is marked with a red dashed line.