eLIFE
elifesciences.org

Figures and figure supplements

Diagnostically relevant facial gestalt information from ordinary photos

Quentin Ferry, et al.

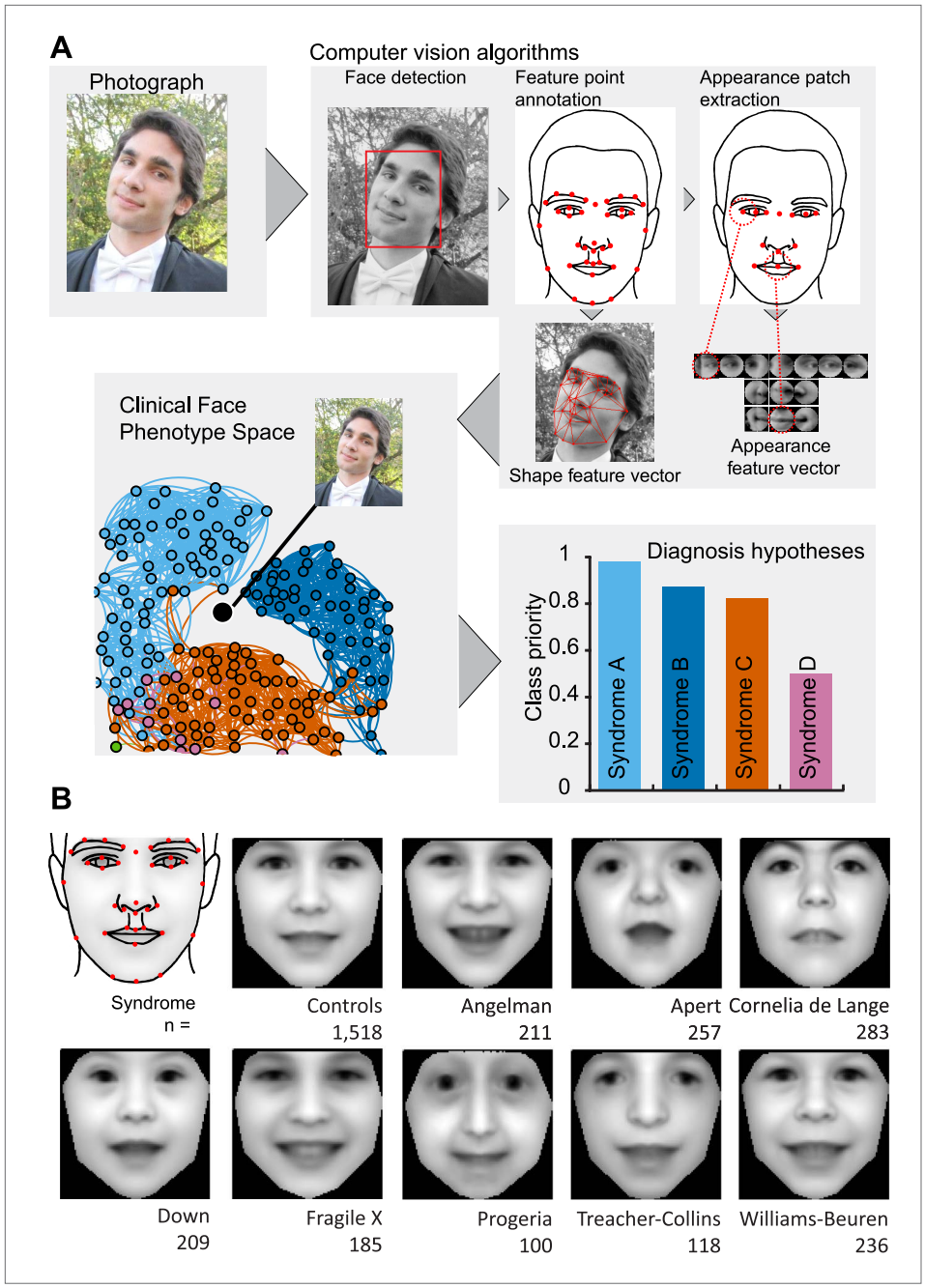

Figure 1. Overview of the computational approach and average faces of syndromes. (**A**) A photo is automatically analyzed to detect faces and feature points are placed using computer vision algorithms. Facial feature annotation points delineate the supra-orbital ridge (8 points), the eyes (mid points of the eyelids and eye canthi, 8 points), nose (nasion, tip, ala, subnasale and outer nares, 7 points), mouth (vermilion border lateral and vertical midpoints, 6 points) and the jaw (zygoma mandibular border, gonion, mental protrubance and chin midpoint, 7 points). Shape and Appearance feature vectors are then extracted based on feature points and these determine the photo's location in Clinical Face Phenotype Space (further details on feature points in ***Figure 1—figure supplement 1***). This location is then analyzed in the context of existing points in Clinical Face Phenotype Space to extract phenotype similarities and diagnosis hypotheses (further details on Clinical Face Phenotype Space with simulation examples in ***Figure 1—figure supplement 2***). (**B**) Average faces of syndromes in the database constructed using AAM models ('Materials and methods') and number of individuals which each average face represents. See online version of this manuscript for animated morphing images that show facial features differing between controls and syndromes (***Figure 2***).

DOI: 10.7554/eLife.02020.003

Figure 1. Continued on next page

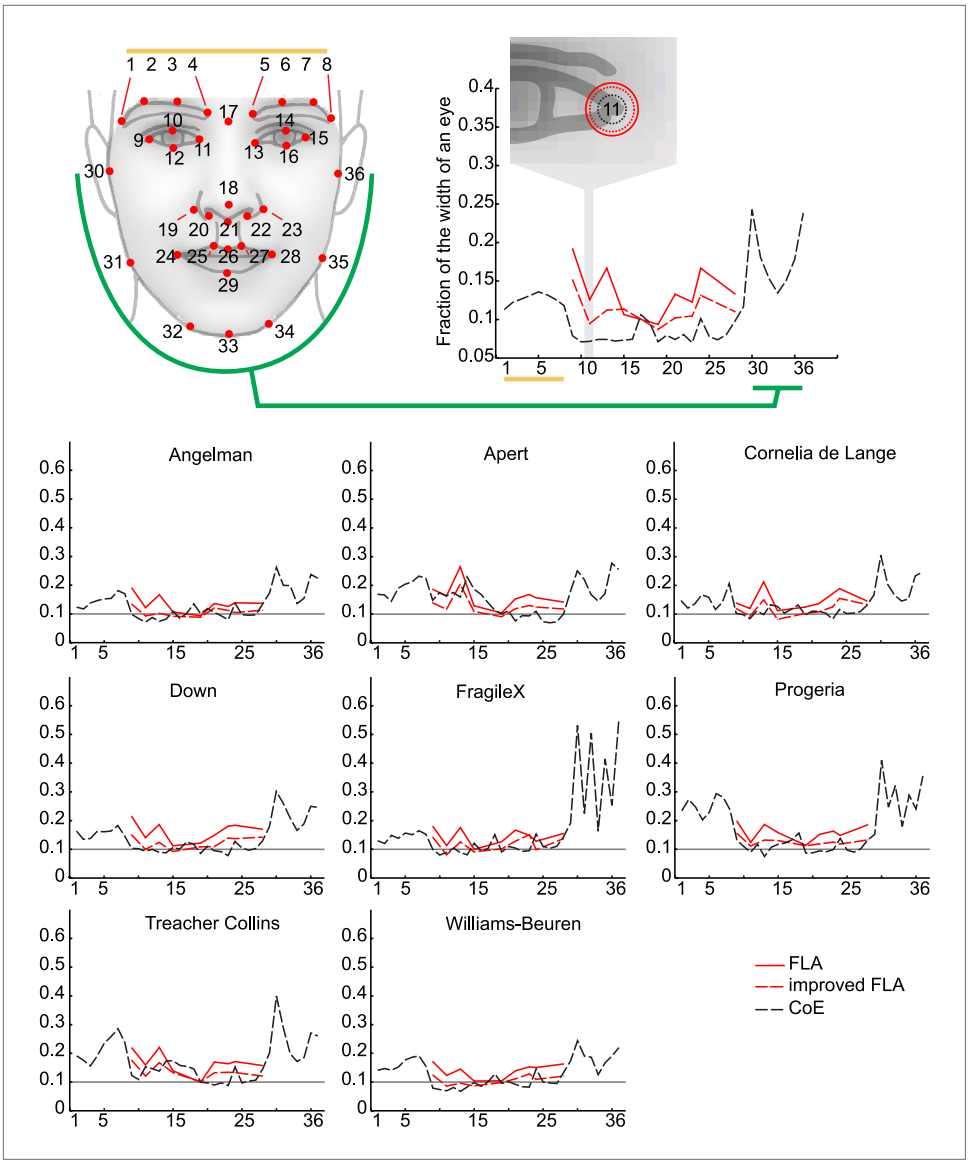

Figure 1—figure supplement 1.

DOI: 10.7554/eLife.02020.004

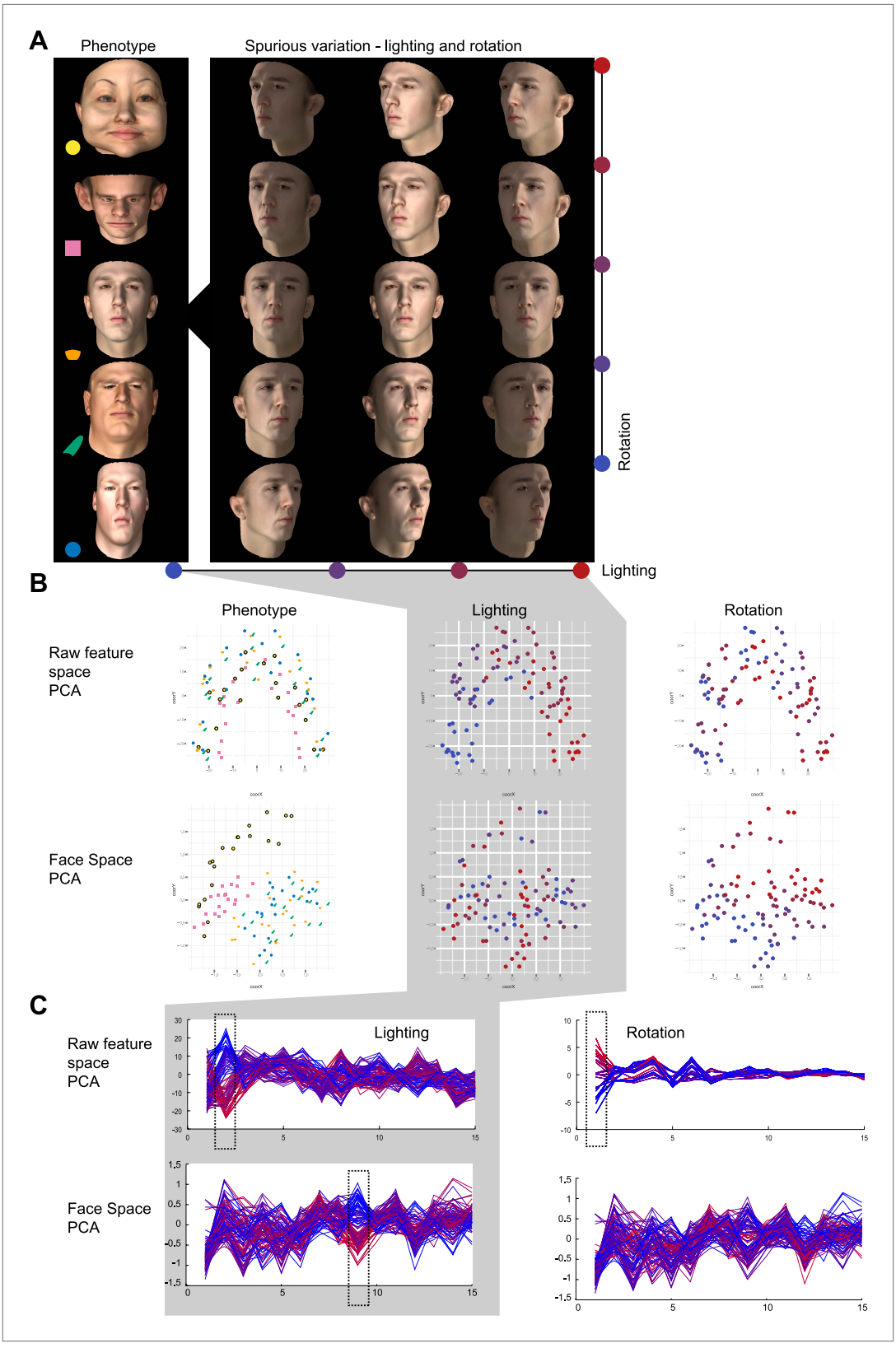

Figure 1—figure supplement 2. Phenotypic vs spurious feature variation in Clinical Face Phenotype Space using simulated faces.

DOI: 10.7554/eLife.02020.005

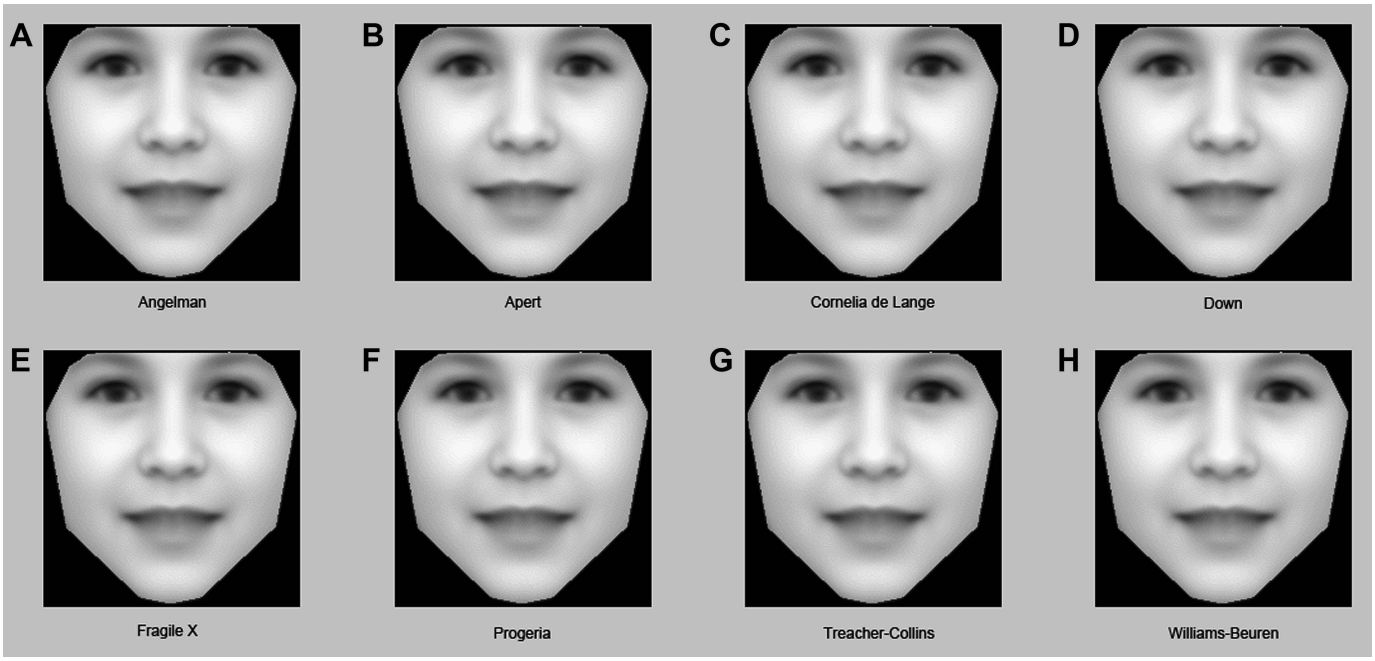

Figure 2. Animated morphs of average faces from controls to syndromes. (**A**) Angelman, (**B**) Apert, (**C**) Cornelia de Lange, (**D**) Down, (**E**) Fragile X, (**F**) Progeria, (**G**) Treacher-Collins, (**H**) Williams-Beuren. Delineation of syndrome gestalt relative to controls with distortion graphs in ***Figure 2—figure supplement 1***.

DOI: 10.7554/eLife.02020.008

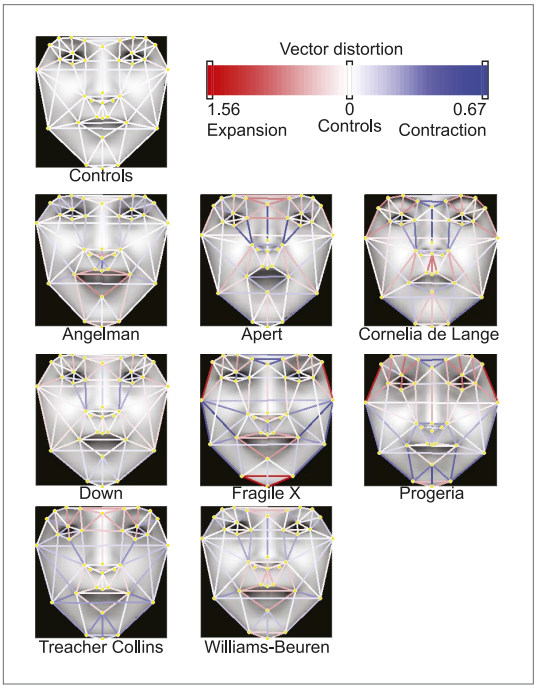

Figure 2—figure supplement 1. Distortion graphs representing the characteristic deformation of syndrome faces relative to the average control face.

DOI: 10.7554/eLife.02020.009

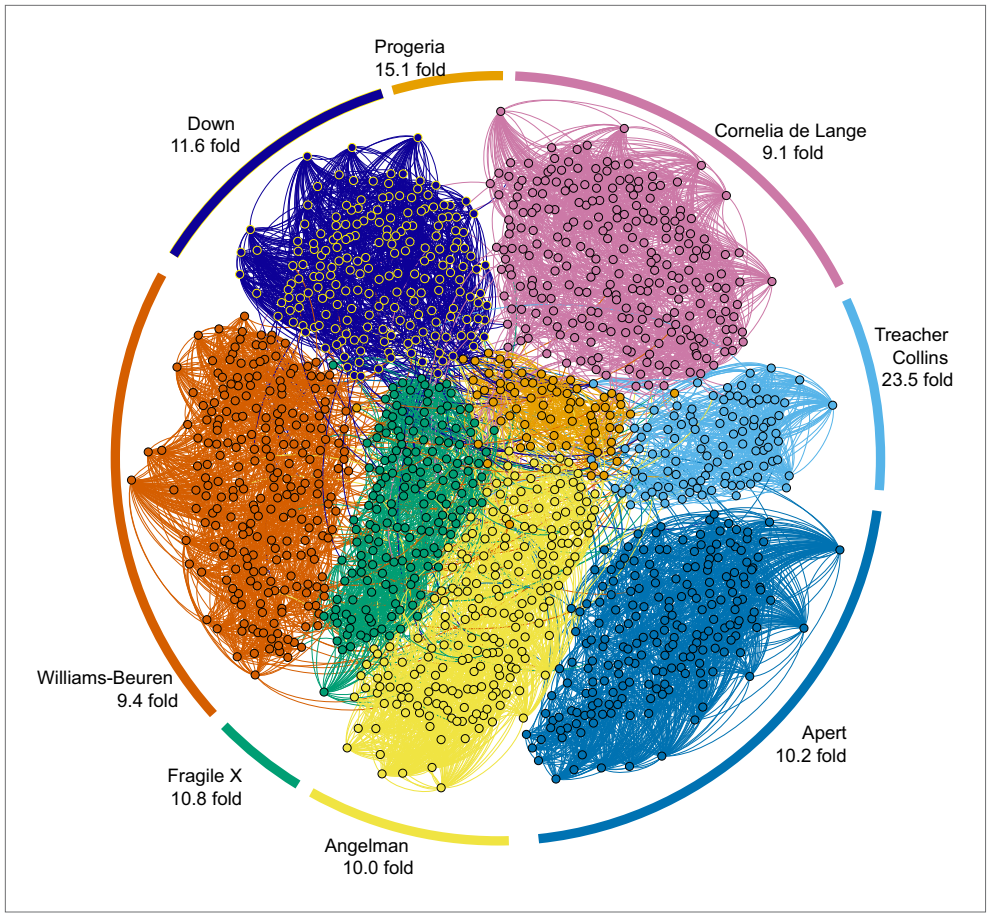

Figure 3. Clinical Face Phenotype Space enhances the separation of different dysmorphic syndromes. The graph shows a two dimensional representation of the full Clinical Face Phenotype Space, with links to the 10 nearest neighbors of each photo (circle) and photos placed with force-directed graphing. The Clustering Improvement Factor (CIF, fold better clustering than random expectation) estimate for each of the syndromes is shown along the periphery.

DOI: 10.7554/eLife.02020.010

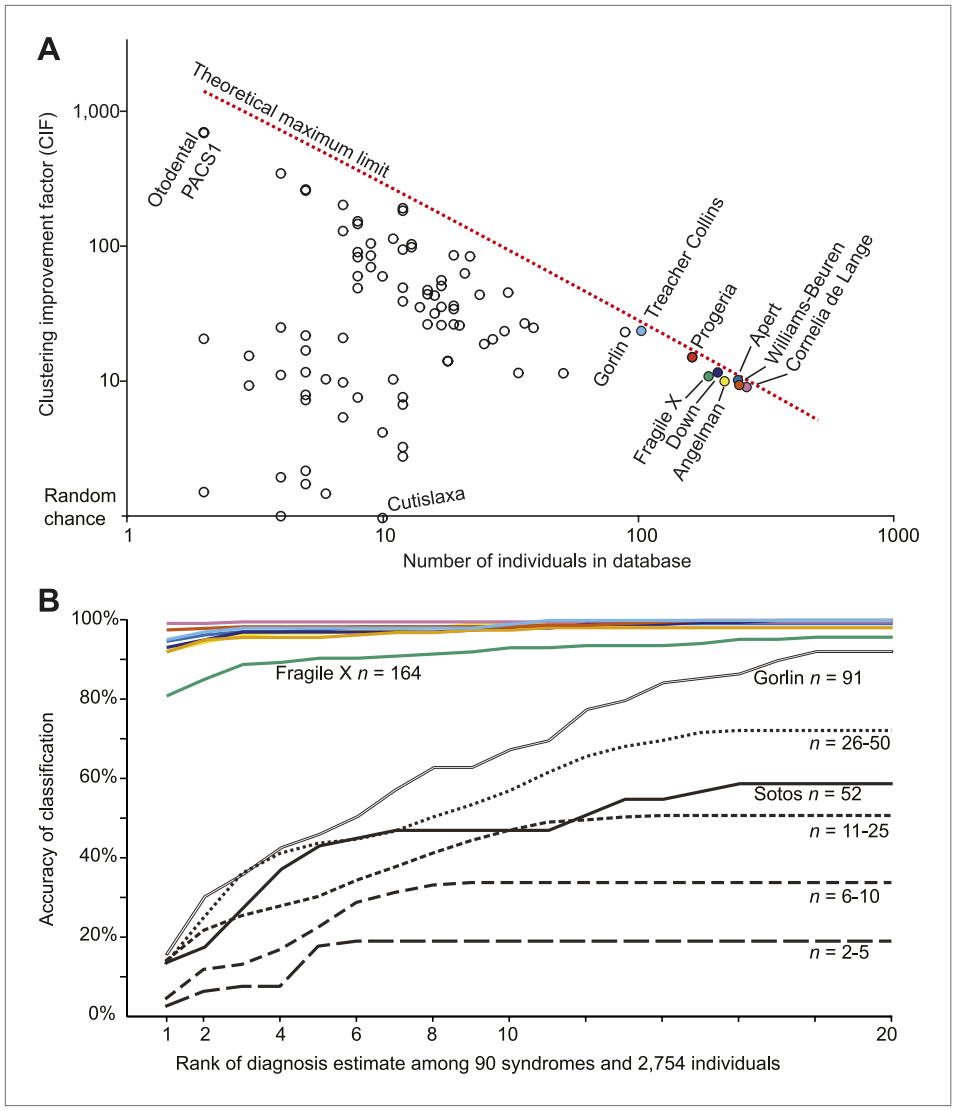

Figure 4. Clinical Face Phenotype Space is generalizable to dysmorphic syndromes that are absent from a training set. (**A**) Clustering Improvement Factor (CIF) estimates are plotted vs the number of individuals per syndrome grouping in the Gorlin collection or patients with similar genetic variant diagnoses. As expected, the stochastic variance in CIF is inversely proportional to the number of individuals available for sampling. The median CIF across all groups is 27.6-fold over what is expected by clustering syndromes randomly. That is to say, the CIF of a randomly placed set is 1. The maximum CIF is fixed by the total number of images in the database and by the cardinality of a syndrome set: the theoretical maximal CIF upper bound is plotted as a red dotted line. The CIF for the minimum and maximum, Cutislaxa syndrome and Otodental syndrome, were 1.0 and 700.0 respectively. (**B**) Average probabilistic classification accuracies of each individual face placed in Clinical Face Phenotype Space (class prioritization by 20 nearest neighbors weighted by prevalence in the database). The 8 initial syndromes used to train Clinical Face Phenotype Space are shown in color. For syndromes with fewer than 50 examples, accuracies were averaged across all syndromes binned by data set size (i.e., the average accuracy is shown for syndromes with 2–5, 6–10, 11–25, and 26–50 images in the database, ***Supplementary file 1***). Classification accuracies increase proportional to the number of individuals with the syndrome present in the database. Accuracies using support vector machines with binary and forced choice classifications are shown in ***Figure 4—figure supplement 1*** and ***Figure 4—figure supplement 2***. A simulation example of probabilistic querying of Clinical Face Phenotype Space is shown in ***Figure 4—figure supplement 3***.
DOI: 10.7554/eLife.02020.011

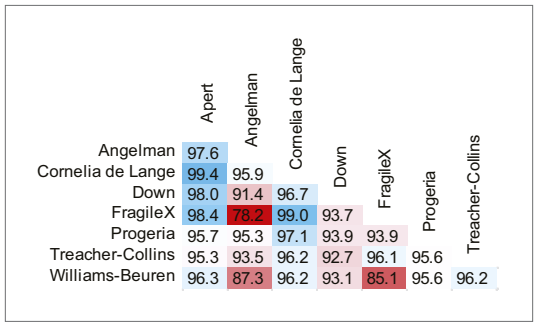

Figure 4—figure supplement 1. SVM binary classification accuracies among the 8 syndromes in ***Table 1***.

DOI: 10.7554/eLife.02020.012

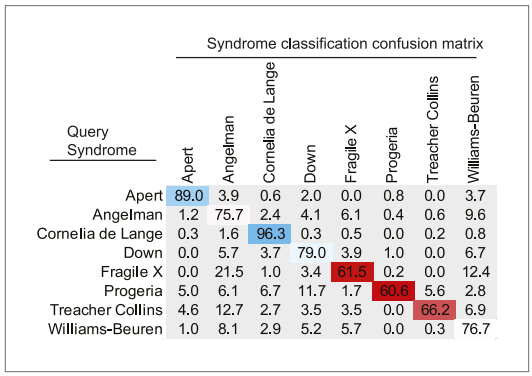

Figure 4—figure supplement 2. SVM forced choice classification accuracies among the 8 syndromes in ***Table 1***.

DOI: 10.7554/eLife.02020.013

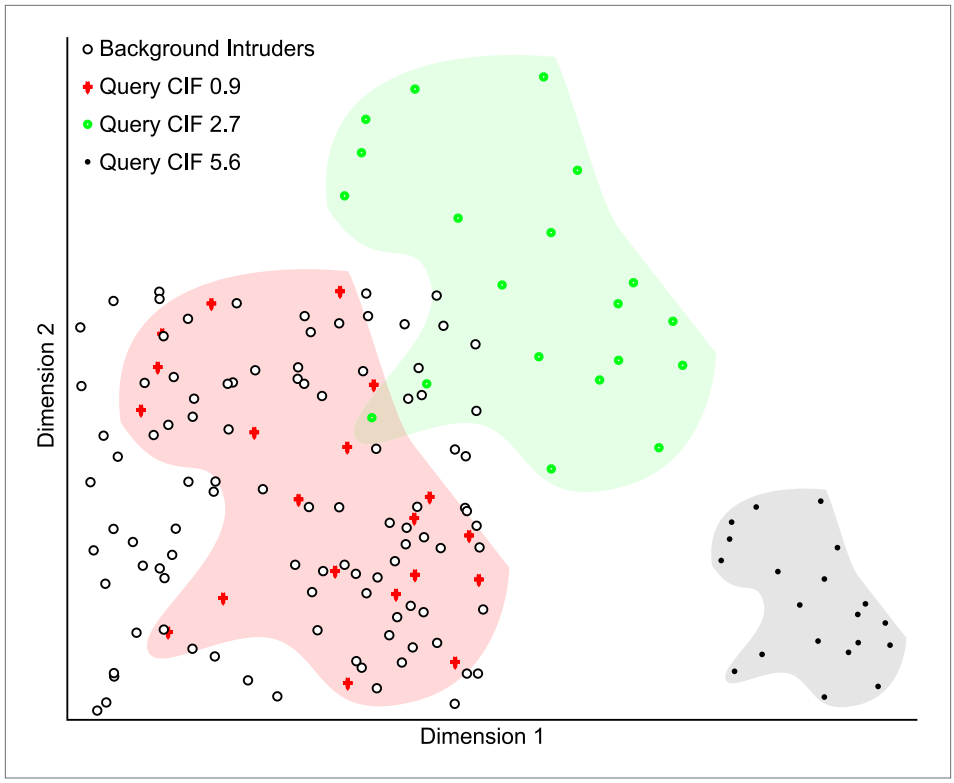

Figure 4—figure supplement 3. Simulated example illustrating the Clustering Improvement Factor.

DOI: 10.7554/eLife.02020.014

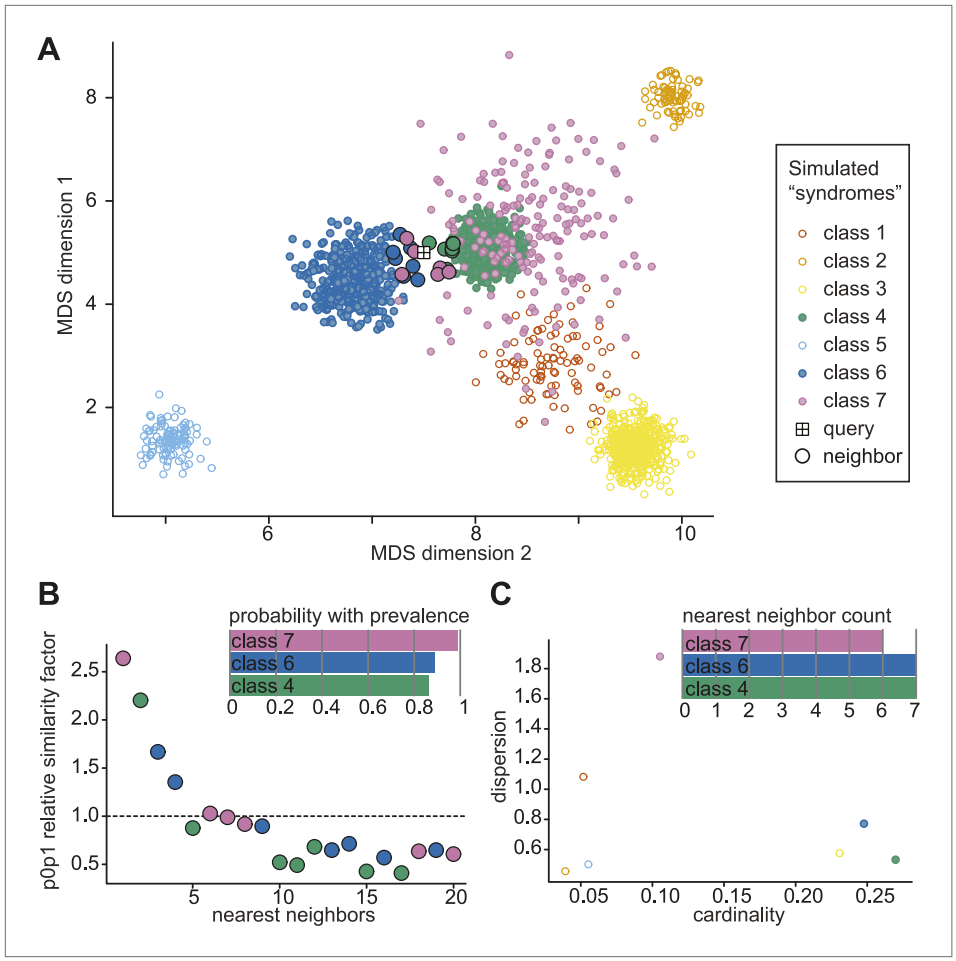

Figure 4—figure supplement 4. Simulated example of probabilistic querying of Clinical Face Phenotype Space.

DOI: 10.7554/eLife.02020.015

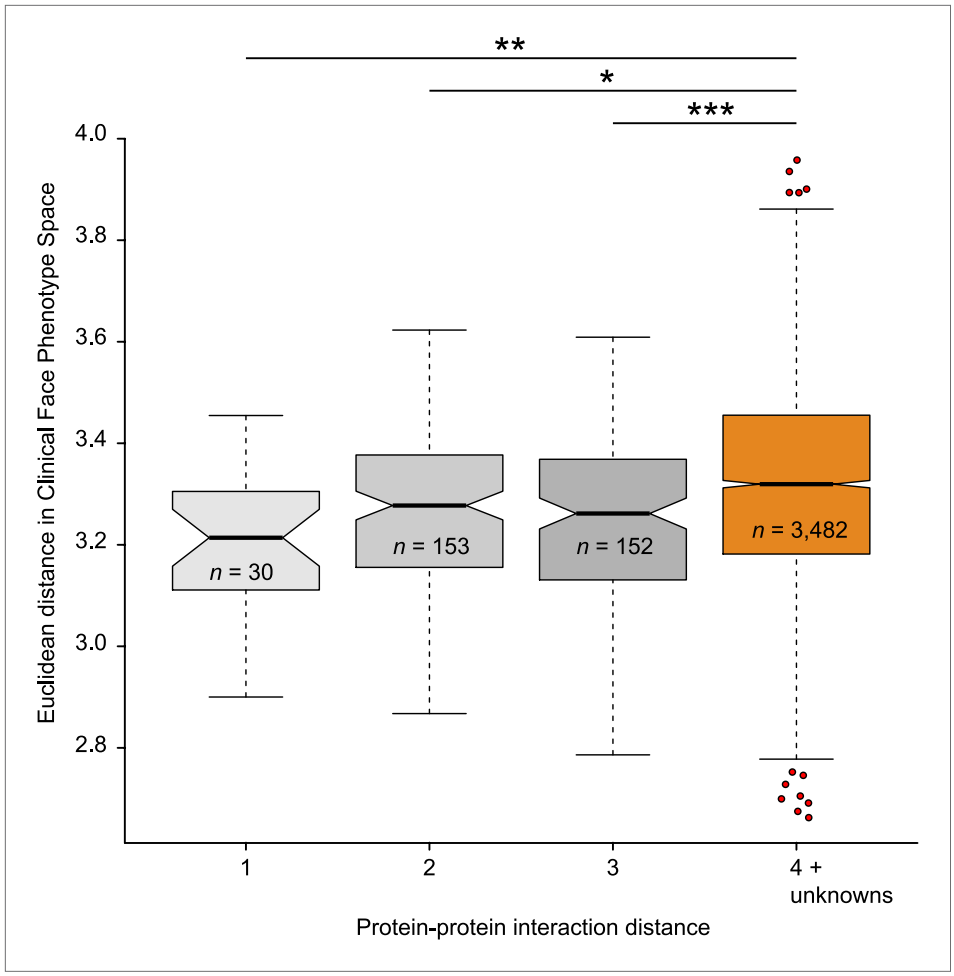

Figure 5. Clinical Face Phenotype Space recapitulates features of functional gene links between syndromes. Protein–protein interaction distances of 1–3 for genetically characterized syndromes are associated with significantly shorter Euclidean distance (arbitrary units) between syndromes in Clinical Face Phenotype Space as compared to syndromes with distance 4 or no known interaction distance (shown in orange) (Kruskal–Wallis tests with Bonferroni corrected p-values indicated as *p<0.05, **p<0.01, ***p<0.001). The Spearman correlation across all distances was r = 0.09, p<0.001. The numbers of pairwise syndrome comparisons underlying each of the interaction distances are listed within the respective boxes.

DOI: 10.7554/eLife.02020.016

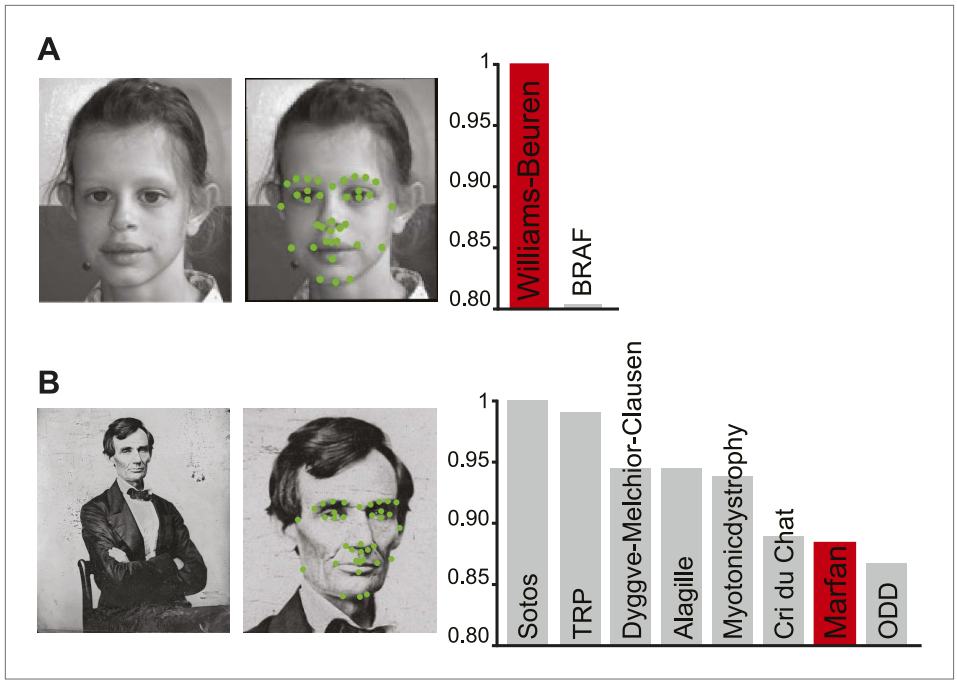

Figure 6. Class priority of diagnostic classifications for images. The full computer vision algorithm and Clinical Face Phenotype Space analysis procedure with diagnostic hypothesis generation exemplified by: (**A**) a patient (***Ferrero et al., 2007***) with Williams-Beuren (this patient figure was published in 'Giovanni Battista Ferrero, Elisa Biamino, Lorena Sorasio, Elena Banaudi, Licia Peruzzi, Serena Forzano, Ludovica Verdun di Cantogno, Margherita Cirillo Silengo. Presenting phenotype and clinical evaluation in a cohort of 22 Williams–Beuren syndrome patients. European Journal of Medical Genetics; 2007;50(5):327–337 Copyright 2007 Elsevier Masson SAS. All rights reserved'.). Note that panel **A** does not fall under a creative commons license and would need permissions from the copyright holders for future reproductions. (**B**) Abraham Lincoln. The former US President is thought to have had a marfanoid disorder, if not Marfan syndrome (***Gordon, 1962***; ***Sotos, 2012***). Bar graphs show class prioritization of diagnostic hypotheses determined by 20 nearest neighbors weighted by prevalence in the database. As expected, the classification of Marfan is not successfully assigned in the first instance as there were only 18 faces of individuals with Marfan in the database (making this an example of a difficult case with the current database). However, the seventh suggestion is Marfan, despite this being among 90 different syndromes and 2754 faces.
DOI: 10.7554/eLife.02020.017