Journal of Translational Genetics and Genomics

Original Article

DeepGestalt analysis of the *SETD5*-associated intellectual disability syndrome

Giulia Pascolini

Medical Genetics, Department of Molecular Medicine, Sapienza University, San Camillo-Forlanini Hospital, Rome 00152, Italy.

Correspondence to: Dr. Giulia Pascolini, Medical Genetics, Department of Molecular Medicine, Sapienza University, San Camillo-Forlanini Hospital, Circonvallazione Gianicolense 87, Rome 00152, Italy. E-mail: giupascolini@gmail.com

Abstract

Aim: This is the first computer-assisted study focused on the craniofacial features of the intellectual disability (ID)/developmental delay (DD) syndrome related to haploinsufficiency of the *SETD5* gene (SET domain-containing protein 5, MIM#615743), which is a chromatin regulator. The purpose of this novel research is to better delineate the facial phenotype of this condition and identify the associated dysmorphic features to consider for clinical diagnosis.

Methods: A total of 18 2D frontal images of previously published pediatric individuals (aged 1-14 years, Caucasian ethnicity) with *SETD5* mutations (SETD5, cohort 1) were uploaded to the RESEARCH application of the Face2Gene online platform (V.19.1.3) (FDNA Inc., Boston, MA, USA). Images from this group of patients were compared with 36 photos of individuals with two other known chromatin disorders, specifically KBG (KBGS, cohort 2, 18 images) and Koolen-de Vries syndromes (KdVS, cohort 3, 18 images), which share with the *SETD5*-related ID syndrome a very similar facial gestalt and peculiar dysmorphisms. An additional cohort of 18 unaffected controls that were matched for age and ethnicity (Ctrl., controls, cohort 4) was also included in the comparison experiment.

Results: Results obtained from the binary comparison analysis were expressed in terms of Area Under the Curve and its Receiver Operating Characteristic curve for aggregated splits. A high facial overlap between the *SETD5*-related phenotype and KBGS was demonstrated. Other conditions considered for the study were well recognized by the system and differentiated using the unaffected controls.

Conclusion: This study confirms the presence of distinctive dysmorphic features that characterize the *SETD5*-related facial phenotype, providing observations about its possible role in facial morphogenesis.

Keywords: *SETD5*, dysmorphisms, intellectual disability, chromatin, Face2Gene

INTRODUCTION

Heterozygous mutations in the *SETD5* gene (SET domain-containing protein 5, MIM#615743; KIAA1757) have been associated with a syndromic intellectual disability (ID)/developmental delay (DD), also termed ID autosomal dominant 23 (MIM#615761)[1].

The SET domain-containing protein 5 gene, mapped to the 3p25.3 genomic locus, encodes a histone-lysine *N*-methyltransferase, which is involved in chromatin remodeling and is strongly expressed in both the adult brain, as well as the spinal cord. Molecular bases of the disorder have also been recently demonstrated to be associated with nonsense-mediated decay, resulting in gene haploinsufficiency[2].

The *SETD5*-related phenotype is mainly characterized by ID/DD, language delay and dysmorphic features. To date, few patients (mostly pediatric) have been reported, being a very rare condition and initial phenotype delineation has been provided by previous authors. The main craniofacial features are represented by brachycephaly, long and smooth philtrums, micrognathia, synophrys, abnormal eyebrow, upslanted palpebral fissures, bulbous nose with depressed nasal bridge and anteverted nares, thin upper lip vermilion, downturned corners of the mouth, and dental crowding.

In this study, the *SETD5*-facial phenotype was analyzed with the DeepGestalt technology (V.19.1.3) (FDNA Inc., Boston, MA, USA; https://www.face2gene.com) for the first time, to further define the craniofacial characteristics, which should be considered as a key feature for clinical diagnosis.

METHODS

"SETD5" and "3p25.3 deletion/haploinsufficiency" terms have been digitized in the PubMed database (https://www.ncbi.nlm.nih.gov/pmc/) to select all scientific works related to the *SETD5*-ID syndrome. Deletions comprising not only *SETD5* were included, based on evidence that it is the main gene determining the typical core phenotype, as previously postulated[2]. Only papers containing patients' photographs were considered for the experiment (for references details see Supplementary Materials). A total of 18 facial two-dimensional (2D) images of individuals with *SETD5* mutations were uploaded to the CLINIC and then RESEARCH applications of the Face2Gene suite (SETD5, Cohort 1) [Figure 1A]. All patient images were anonymized and uploaded to the personal account of the user (the author), which is protected by a password and made inaccessible to others. All photos were processed by the system to generate unrecognizable, composite matrices, which were then used for the comparison study (see below). The DeepGestalt technology is based on deep learning algorithms built for syndrome-specific computational-based classifiers (syndrome gestalts), converting patient photos into de-identified mathematical facial descriptors. The patient's facial descriptors are then compared to syndrome gestalts to quantify similarity (gestalt scores) resulting in a prioritized list of syndromes with similar morphology. To date, available individuals with *SETD5* mutations are mostly pediatric (1-14 years) and Caucasian. Only one patient of Asian origin was included in the present analysis[3]. Cohort 1 was compared with three other cohorts: (1) one containing 18 published images of patients with KBG syndrome (KBGS, MIM#148050) (KBGS, Cohort 2); (2) 18 images of patients with Koolen-de Vries syndrome (KdVS, MIM#610443) (KdVS, Cohort 3); and (3) the last consisting of 18 images of unaffected individuals (Ctrl., controls, Cohort 4). KBGS and KdVS were included in the study because both are chromatin disorders, which are caused respectively by mutations in the chromatin modulator *ANKRD11* (Ankyrin repeat domain-containing protein 11, MIM#611192) and *KANSL1* (KAT8 regulatory NSL complex subunit 1, MIM#612452). Both conditions share strongly overlapping facial features, including nasal and mouth dysmorphisms with the *SETD5*-related ID syndrome. The DeepGestalt technology based on the Face2Gene platform was performed, as described previously[4].

RESULTS

Multiclass comparison analysis generated a confusion matrix, in which errors (false positives and false negatives) were represented [Figure 1B]. The highlighted diagonal line indicates true positives values, which
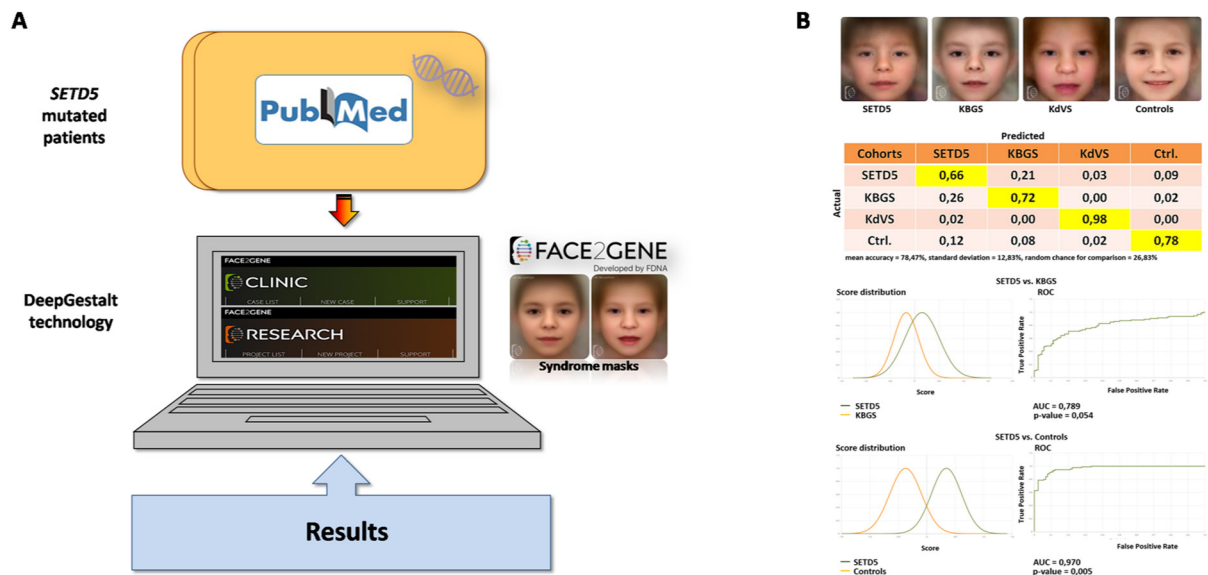
Figure 1. Schematic representation of the main phases of the present study and DeepGestalt technology experiment results. A: a summary of the main phases of the study is provided; B: composite photos for the comparison analysis between the 4 study cohorts (SETD5, Cohort 1; KBGS, Cohort 2; KdVS, Cohort 3, CTRL Ctrl.; unaffected controls, Cohort 4) (at the top). For each group, 18 2D frontal images were analyzed. Confusion matrix (in the middle). The technology automatically detects facial points and recognizes dysmorphic features through images analysis, evaluating similarities to known facial patterns (gestalt). Results of the SETD5 *vs.* KBGS and SETD5 *vs.* Ctrl. binary comparison experiment. Note the overlapping curves for the first comparison analysis, indicating a facial overlap of the 2 studied conditions (at the bottom). AUC: area under the curve; ROC: receiver operating characteristic

range between 0.66 and 0.98 [Figure 1B]. The mean accuracy was 78.47% with a standard deviation of 12.83%, and the random chance for comparison was 26.03% [Figure 1B]. Results of the binary comparison experiment are expressed in terms of Area Under the Curve (AUC) and its Receiver Operating Characteristic (ROC) curve for aggregated splits. Values result to be comprised between 0.735 (P = 0.050) and 1.000 (P = 0.000) (all experimental results are available in Supplementary Materials).

DISCUSSION

This is the first reported work to study the facial phenotype of the *SETD5*-associated ID, a very rare malformation condition that is mainly characterized by neurodevelopmental delay, variable congenital defects and dysmorphic craniofacial features. It can be included in the genetically heterogeneous group of chromatin disorders, which represents a specific set of ID syndromes. These are mainly caused by mutations in the various components of the chromatin remodeling BAF-complex (which are typically accompanied by hypertrichosis as often happens in ID[5]) or in the histone's modifiers, including *SETD5*. The latter encodes a ubiquitously expressed methyltransferase containing a conserved domain of 130 amino acids and 2 signature motifs (ELxF/YDY and NHS/CxxPN), acting as essential gene for normal embryo development and survival, as has been demonstrated by previous *in vivo* studies[6-9].

Dysmorphisms are frequently encountered in patients with *SETD5* mutations, as previously reported[1]. To better define the related phenotype, the facial appearance of affected individuals was compared, for the first-time, with two other conditions with strong *SETD5*-ID clinical overlap and caused by mutations in molecules involved in histones modification. Specifically, KBG (KBGS) and Koolen-de Vries (KDVS) syndromes, which are respectively associated with mutations in *ANKRD11*[10] and *KANSL1*[11], were considered. Both are clinically defined by ID/DD, malformations and distinctive dysmorphisms, strongly resembling the *SETD5*-ID. This was also confirmed by the CLINIC application of the Face2Gene platform, which identified KBGS (18/18, 100%) as the most probable clinical diagnosis in patients with *SETD5* variant, based exclusively on facial features. KdVS was also proposed by the system for 4/18 (22%) individuals. KBGS is characterized by microcephaly, a round-

triangular face, long philtrum, large and protruding ears, ocular anomalies (hypertelorism, telecanthus, long palpebral fissures), broad and thick eyebrows, an abnormal nose with anteverted nares and underdeveloped alae nasi, and malformed teeth (macro-oligodontia, wide upper central or fused incisors, ridged teeth). On the other hand, a long face with a broad forehead, broad chin, large and protruding ears with overfolded helices, upslanted palpebral fissures, high and narrow palate with cleft lip/palate, everted lower lip vermilion, small widely spaced teeth and typical nose morphology (tubular-pear shaped nose with broad nasal tip) define KdVS. Interestingly, macrodontia is a frequently observed clinical sign in KBGS (> 80% of patients)[12], as well as an abnormal nose can also be considered a distinctive KdVS facial dysmorphism[13]. Dental and nasal malformations, such as a bulbous nose with broad nasal bridge, are also frequently identified in the *SETD5*-syndrome, demonstrating a strong overlapping facial gestalt between these three conditions. These observations could be explained by the common biological cause of the diseases, or mutations in genes encoding key molecules for histones changes.

Results of the DeepGestalt experiment are in line with the above hypothesis. Specifically, the SETD5-KBGS comparison analysis demonstrated AUC and ROC values compatible with a high degree of facial overlap, which was shown by the two overlapping curves in the binary comparison and by the ROC graphic [Figure 1B]. A minor facial similarity was registered in the SETD5-KdVS comparison. All three conditions were distinctly recognized by the DeepGestalt technology and the obtained results seem to indicate that this group of disorders are all characterized by distinctive dysmorphisms, confirming what has recently been illustrated in a clinical phenotyping study on ID syndromes related to mutations in histones modifiers[14].

Furthermore, the DeepGestalt technology has been demonstrated to not be influenced by ethnicity, according to previously published works, which analyzed the facial features of individuals with different ancestries and affected by diverse genetic conditions[15-17]. These studies showed that the technology identifies facial gestalt independently from an individual's ancestry.

Craniofacial malformations are possibly related to mutations in *SETD5*, which could be involved in facial morphogenesis, analogously to other chromatin remodeling genes. For example, *ANKRD11* has been demonstrated to influence bone development[18] and facial morphogenesis[10]. Interestingly, the nose and mouth regions are often malformed in the *SETD5*-related phenotype. The possible connection between this gene and the embryological pathways regulating nose and mouth formation could be an interesting area of research. In this scenario, deregulation of embryonic structures controlling physiological facial development such as the neural crests, could be further investigated in affected individuals. However, all these have to be verified in future studies.

In conclusion, a novel study regarding the ID syndrome related to *SETD5* mutations by using the DeepGestalt technology has been here illustrated. The present results highlight the strong facial resemblance between affected subjects and those with KBGS, suggesting to perform *SETD5* analysis in negatives cases for *ANKRD11* mutations. Furthermore, nose and mouth abnormalities should be evaluated with careful attention in ID individuals because distinctive dysmorphic facies is an important clinical handle. Finally, a possible role of *SETD5* in facial deformities could be hypothesized, based on the demonstrated action of other histones modifiers on craniofacial development.

DECLARATIONS

Acknowledgments

Authors' contributions
The author contributed solely to the article.

Availability of data and materials

All data regarding this study are available by the corresponding author upon reasonable request.

Financial support and sponsorship

None.

Conflicts of interest

The author declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

All patient images were anonymized and uploaded to the personal account of the user (the author), which is protected by a password and made inaccessible to others. All photos were processed by the system to generate unrecognizable, composite matrices, which were then used for the comparison study.

Copyright

© The Author(s) 2020.

REFERENCES

1. Grozeva D, Carss K, Spasic-Boskovic O, Parker MJ, Archer H, et al. De novo loss-of-function mutations in SETD5, encoding a methyltransferase in a 3p25 microdeletion syndrome critical region, cause intellectual disability. Am J Hum Genet 2014;94:618-24.
2. Kuechler A, Zink AM, Wieland T, Lüdecke HJ, Cremer K, et al. Loss-of-function variants of SETD5 cause intellectual disability and the core phenotype of microdeletion 3p25.3 syndrome. Eur J Hum Genet 2015;23:753-60.
3. Kellogg G, Sum J, Wallerstein R. Deletion of 3p25.3 in a patient with intellectual disability and dysmorphic features with further definition of a critical region. Am J Med Genet A 2013;161A:1405-8.
4. Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, et al. Identifying facial phenotypes of genetic disorders using deep learning. Nat Med 2019;25:60-4.
5. Pezzani L, Milani D, Tadini G. Intellectual disability: when the hypertrichosis is a clue. J Pediatr Genet 2015;4:154-8.
6. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ. Altered Hox expression and segmental identity in Mll-mutant mice. Nature 1995;378:505-8.
7. O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, et al. The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol 2001;21:4330-6.
8. Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 2002;16:1779-91.
9. Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, et al. NSD1 is essential for early post-implantation development and has a catalytically active SET domain. EMBO J 2003;22:3153-63.
10. Sirmaci A, Spiliopoulos M, Brancati F, Powell E, Duman D, et al. Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. Am J Hum Genet 2011;89:289-94.
11. Koolen DA, Pfundt R, Linda K, Beunders G, Veenstra-Knol HE, et al. The Koolen-de Vries syndrome: a phenotypic comparison of patients with a 17q21.31 microdeletion versus a KANSL1 sequence variant. Eur J Hum Genet 2016;24:652-9.
12. Low K, Ashraf T, Canham N, Clayton-Smith J, Deshpande C, et al. Clinical and genetic aspects of KBG syndrome. Am J Med Genet Part A 2016;170:2835-46.
13. Koolen DA, Kramer JM, Neveling K, Nillesen WM, Moore-Barton HL, et al. Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. Nat Genet 2012;44:639-41.
14. Pascolini G, Fleischer N, Ferraris A, Majore S, Grammatico P. The facial dysmorphology analysis technology in intellectual disability syndromes related to defects in the histones modifiers. J Hum Genet 2019;64:721-8.
15. Vorravanpreecha N, Lertboonnum T, Rodjanadit R, Sriplienchan P, Rojnueangnit K. Studying down syndrome recognition probabilities in Thai children with de-identified computer-aided facial analysis. Am J Med Genet Part A 2018;176:1935-40.
16. Fung JLF, Rethanavelu K, Luk HM, Ho MSP, Lo IFM, et al. Coffin-Lowry syndrome in Chinese. Am J Med Genet Part A 2019;179:2043-8.
17. Narayanan DL, Ranganath P, Aggarwal S, Dalal A, Phadke SR, et al. Computer-aided facial analysis in diagnosing dysmorphic syndromes in Indian children. Indian Pediatr 2019;56:1017-9.
18. Zhang A, Yeung PL, Li CW, Tsai SC, Dinh G, et al. Identification of a novel family of ankyrin repeats containing cofactors for p160 nuclear receptor coactivators. J Biol Chem 2004;279:33799-805.