CHARGE syndrome and related disorders

Ufartes, Roser; Grün, Regina; Salinas, Gabriela; Sitte, Maren; Kahl, Fritz; Wong, Monica T Y; van Ravenswaaij-Arts, Conny M A; Pauli, Silke

Published in:
Human Molecular Genetics

DOI:
10.1093/hmg/ddab183

Publication date:
2021

Human Molecular Genetics, 2021, Vol. 30, No. 23 2215–2224

https://doi.org/10.1093/hmg/ddab183
Advance Access Publication Date: 6 July 2021
General Article

GENERAL ARTICLE

CHARGE syndrome and related disorders: a mechanistic link

Roser Ufartes[1], Regina Grün[1], Gabriela Salinas[2], Maren Sitte[2], Fritz Kahl[3], Monica T.Y. Wong[4], Conny M.A. van Ravenswaaij-Arts[4] and Silke Pauli[1,*]

[1]Institute of Human Genetics, University Medical Center Göttingen, Göttingen 37073, Germany, [2]NGS Integrative Genomics Core Unit, University Medical Center Göttingen, Göttingen 37073, Germany, [3]Department of General-, Visceral- and Pediatric Surgery, University Medical Center Göttingen, UMG, Göttingen, Germany and [4]University of Groningen, University Medical Center Groningen, Department of Genetics, RB, Groningen 9700, The Netherlands

*To whom correspondence should be addressed. Tel: +49 5513967591; Fax: +49 5513967567; Email: silke.pauli@med.uni-goettingen.de

Abstract

CHARGE syndrome is an autosomal dominant malformation disorder caused by pathogenic variants in the chromatin remodeler CHD7. Affected are craniofacial structures, cranial nerves and multiple organ systems. Depending on the combination of malformations present, its distinction from other congenital disorders can be challenging. To gain a better insight into the regulatory disturbances in CHARGE syndrome, we performed RNA-Seq analysis on blood samples of 19 children with CHARGE syndrome and a confirmed disease-causing *CHD7* variant in comparison with healthy control children. Our analysis revealed a distinct CHARGE syndrome pattern with downregulation of genes that are linked to disorders described to mimic the CHARGE phenotype, i.e. *KMT2D* and *KDM6A* (Kabuki syndrome), *EP300* and *CREBBP* (Rubinstein-Taybi syndrome) and *ARID1A* and *ARID1B* (Coffin-Siris syndrome). Furthermore, by performing protein–protein interaction studies using co-immunoprecipitation, direct yeast-two hybrid and *in situ* proximity ligation assays, we could demonstrate an interplay between CHD7, KMT2D, KDM6A and EP300. In summary, our data demonstrate a mechanistic and regulatory link between the developmental disorders CHARGE-, Kabuki- and Rubinstein Taybi-syndrome providing an explanation for the overlapping phenotypes.

Introduction

CHARGE syndrome (MIM 214800) is a developmental disorder affecting craniofacial structures, cranial nerves and multiple organ systems. The acronym CHARGE is derived from the main clinical features, i.e. coloboma, heart defects, atresia of the choanae, retarded growth and development, genital hypoplasia, ear anomalies and deafness (1). Additional pathognomonic features are hypoplasia of the semicircular canals, facial nerve palsy, cleft lip/palate and tracheoesophageal fistula (2–4). The syndrome is caused by heterozygous loss-of-function mutations in *CHD7*, a chromodomain helicase DNA-binding protein acting as an ATP-dependent chromatin remodeler (5).

In about 90% of patients who fulfill the clinical diagnostic criteria of Blake *et al.* (6) and/or Verloes (7) a disease-causing variant can be found (3,8), while in only 30–60% of patients who are suspected of having CHARGE syndrome can a pathogenic variant be detected (8–10).

Recently, Moccia *et al.* performed whole exome sequencing (WES) in a cohort of 28 patients who fulfilled the clinical diagnostic criteria for CHARGE syndrome and were either tested negative for pathogenic *CHD7* mutations or did not have a

Received: May 5, 2021. **Revised:** June 23, 2021. **Accepted:** June 23, 2021

previous genetic test. In this preselected cohort, they found in about 53% of individuals a disease-causing *CHD7* mutation and in about 14% a pathogenic variant in genes linked to other developmental disorders like Rubinstein-Taybi syndrome (*EP300*), Kabuki syndrome (*KMT2D*), Verheij syndrome (*PUF60*, MIM: 615583) and RERE-associated developmental disorder (*RERE*, MIM: 616975) (11). Their results show that a number of syndromes can mimic the clinical presentation of CHARGE syndrome.

Kabuki syndrome is one of the disorders with a well-described phenotype overlapping with CHARGE. We and others have previously described that the two developmental disorders can be difficult to distinguish in young patients with a multitude of organ malformations fitting the spectrum of CHARGE syndrome, also due to the fact that the characteristic facial gestalt of Kabuki syndrome may be lacking in early infancy (12–14,11).

Kabuki syndrome is caused by loss-of-function mutations in either the histone methyltransferase KMT2D (MIM 602113) or the lysine-specific demethylase KDM6A (MIM 300128) (15–17). KMT2D and KDM6A have diverse but complementary functions. While KMT2D adds up to three methyl groups to histone 3 lysine 4 (H3K4), KDM6A removes the trimethylation from histone 3 lysine 27 (H3K27me3). Both actions lead to open chromatin structures and allow gene expression.

Because CHD7, KMT2D and KDM6A are all components of the histone machinery and as remodeler, writer or eraser are involved in epigenetic regulation, a common pathomechanism explaining the overlapping phenotypes of CHARGE and Kabuki syndromes has been assumed. Although DNA methylation signatures differ between CHARGE and Kabuki syndrome patients, direct or indirect interactions via a common chromatin modification and remodeling complex or regulation of same target genes have been suggested (12,14,13).

To gain a better insight into the regulatory disturbances in individuals affected by CHARGE syndrome, we performed RNA-Seq analysis on blood samples of 19 children with CHARGE syndrome and a confirmed disease-causing *CHD7* variant. Compared to healthy control children, we observed a distinct CHARGE syndrome pattern. Interestingly, the downregulated genes include the Kabuki-related genes *KMT2D* and *KDM6A* as well as genes linked to disorders described to mimic the CHARGE phenotype like *EP300* and *CREBBP* (Rubinstein-Taybi syndrome) or *ARID1A* and *ARID1B* (Coffin-Siris syndrome). Furthermore, our protein–protein interaction studies confirmed an interplay between CHD7, KMT2D, KDM6A and EP300. In summary, our data demonstrate a mechanistic and regulatory link between the developmental disorders of CHARGE, Kabuki and Rubinstein-Taybi syndromes.

Results

RNA-Seq analysis performed on blood samples of CHARGE patients revealed a downregulation of genes mutated in syndromes which can mimic the CHARGE syndrome phenotype

To study the downstream effects of CHD7 and to identify commonly misregulated genes/pathways in CHARGE syndrome patients, we performed RNA-Seq analysis on human RNA. RNA was isolated from blood samples of 19 CHARGE patients (8 female and 11 male) and 10 control children (5 female and 5 male). Among differentially expressed genes with a fold change greater than 2 [log2 fold-change (FC) ≤ -1 or ≥ 1] and a false discovery rate (FDR) of $< 5\%$ are genes that are involved in epigenetic regulation processes like chromatin-remodeling and modification (Fig. 1). Interestingly, this includes genes affected in syndromes which are described as being differential diagnoses of CHARGE syndrome, for example, *KMT2D* and *KDM6A* (Kabuki syndrome, MIM 147920 and 300 867), *CREBBP, EP300* (Rubinstein-Taybi syndrome, MIM 180849 and 613 684) and *ARID1A, ARID1B* (Coffin-Siris syndrome, MIM 614607 and 135 900). A gender specific effect could not be observed.

In our analysis, ribosome biogenesis is one of the most extensively altered pathways (Fig. 2). Many genes encoding proteins of both the small and the large subunit of ribosomes (e.g. *RPS5, RPS6, RPS8, RPS9, RPS10, RPS11, RPS14, RPS19, RPS20, RPS21, RPS27, RPS27A, RPL10L, RPL11, RPL14, RPL15, RPL18, RPL18A, RPL28, RPL31, RPL32, RPL36, RPL41*) and genes coding for mitochondrial ribosomal proteins (e.g. *MRPL11, MRPL12, MRPL14, MRPL20, MRPL23, MRPL24, MRPL27, MRPL33, MRPL51, MRPL58*) are upregulated. Furthermore, *POLR2F*, a gene encoding the subunit F of the DNA-dependent RNA polymerase II, which is responsible for synthesizing messenger RNA is also upregulated.

To validate the RNA-Seq results, we performed quantitative reverse transcription PCR (qRT-PCR) for selected genes. For this, we chose genes from each category, i.e. downregulated (*CHD7, KMT2D, KDM6A, EP300*), not regulated (*WDR5*) and upregulated (RPS9, RPS11). QRT-PCR was performed using the same RNA as for the RNA-Seq analysis from 4 female and 4 male CHARGE patients with missense, splice site, nonsense, and frameshift variants as well as from 6 healthy children (3 girls and 3 boys) and from 5 healthy adults (1 male, 4 female). Except for *KMT2D* (comparison CHARGE patients with healthy children controls), the results of the RNA-Seq analysis were validated by qRT-PCR for all other selected genes in comparison CHARGE patients with healthy children and adults (Fig. 3A and B).

CHD7 co-immunoprecipitates with KMT2D and KDM6A

To investigate a potential interaction between CHD7, KMT2D and KDM6A, we performed co-immunoprecipitation (Co-IP) studies using protein extracts from transfected HEK293 cells. After transfection and culturing the cells for 24 h, proteins were isolated from total cell lysates.

For analyzing a potential CHD7-KDM6A interaction, the cells were double transfected with CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis and KDM6A full-length-pCMV-HA plasmids. KDM6A was immunoprecipitated by using an anti-KDM6A antibody, while an anti-FLAG antibody was used for CHD7 detection. Western blot analysis showed a protein band of ~336 kDa corresponding to the estimated size of the CHD7 full-length protein (Fig. 4A). The reciprocal experiment, immunoprecipitation with the anti-CHD7 antibody and use of the anti-HA antibody for detection, showed a specific protein band of ~154 kDa, the estimated size of full-length KDM6A (Fig. 4B). To exclude unspecific binding of the antibody in the immunoprecipitation and anti-CHD7 antibody Co-IPs were performed on protein extract isolated from cells transfected with the empty pCMV-HA plasmid.

For the KMT2D-CHD7 analysis, two different transfections were performed due to the large size of KMT2D of 593 kDa. First, cells were transfected with KMT2D full-length-pCMV-HA and CHD7-CR1–3-pCMV-c-Myc, which contains a part of CHD7. Endogenous KMT2D was immunoprecipitated using an anti-KMT2D antibody and the CHD7 part was detected with an anti-c-Myc antibody. A protein band of the expected size of ~67 kDa was detected (Fig. 4C). Second, cells were

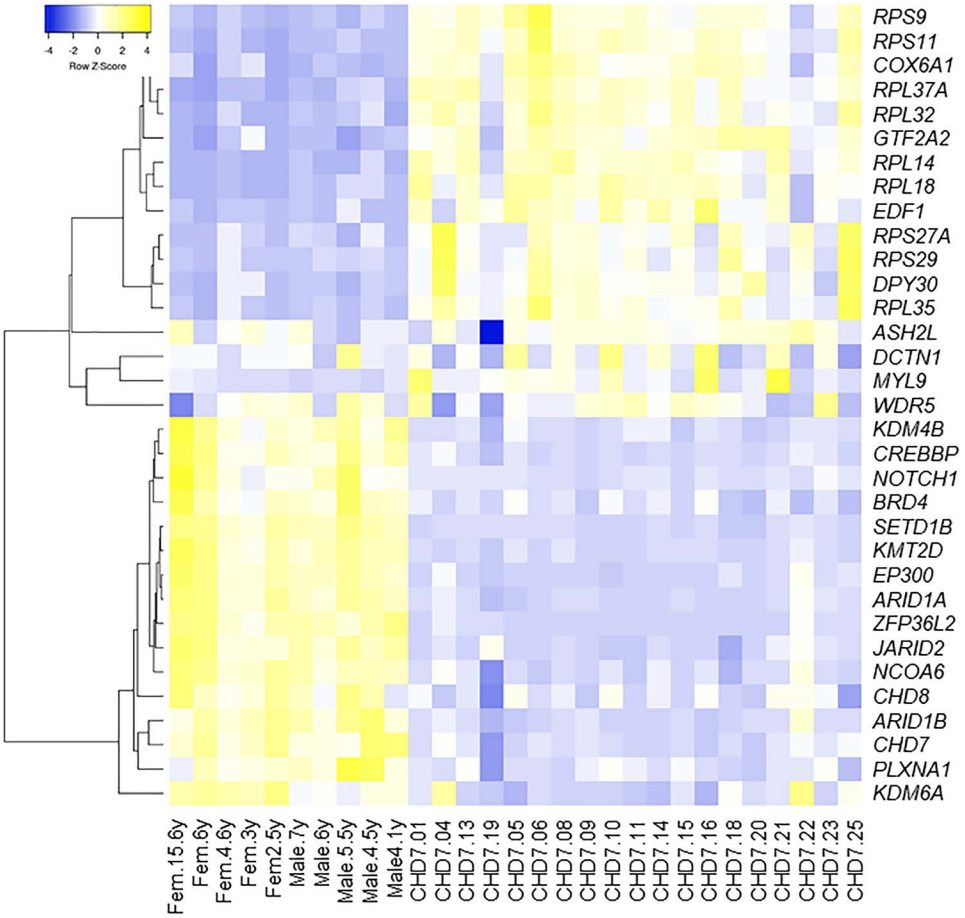

Figure 1. Heatmap: Heatmap showing a selection of significantly (FDR < 5%) upregulated (log2 FC ≥ 1, yellow) and downregulated (log2 FC ≤ −1, blue) genes as well as the confirmed not regulated gene *WDR5*. The expression profiles of 19 CHARGE patients were compared to 10 healthy children controls.

transfected with KMT2D full-length-pCMV-HA and CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis. Endogenous CHD7 was immunoprecipitated using an anti-CHD7 and KMT2D was detected with an anti-HA antibody. A weak protein band of 593 kDa was detected (Fig. 4D).

In situ proximity ligation assays (PLA) verify the interaction between CHD7 and KMT2D and between CHD7 and KDM6A.

To verify the Co-IP results, Duolink® PLA was performed in transfected HEK293 cells. The cells were either double transfected with the CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis and KMT2D full-length pCMV-HA plasmids or with the plasmids CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis and KDM6A full-length-pCMV-HA. For verifying a CHD7 and KMT2D interaction, an anti-CHD7 antibody and an anti-KMT2D antibody was used. The cell nuclei were stained with DAPI. The CHD7 and KMT2D interaction was visualized in the cell nuclei of transfected cells (Fig. 4E).

In case of the CHD7 and KDM6A double transfection, the cells were incubated with an anti-CHD7 and anti-HA antibody. In 10–30% of the cells, demonstrating the transfection efficiency, many red PLA signals could be observed, indicating a protein–protein interaction. As a positive control, HEK293 cells were transfected with the KDM6A full-length-pCMV-HA plasmid. The cells were incubated with an anti-HA and anti-KMT2D antibody to visualize the known interaction between KDM6A and KMT2D. As a negative control, cells were transfected with the KDM6A full-length-pCMV-HA plasmid and were incubated with an anti-KDM6A and an anti-SC35 antibody (marker for nuclear speckles, recognizing the non-snRNP (small nuclear ribonucleoprotein particles) factor SC-35). Unspecific PLA signals could be excluded.

The CHD7 interaction with KMT2D and KDM6A is likely indirect

To determine whether CHD7 interacts with either KMT2D or KDM6A direct or indirectly via additional linker proteins, we performed direct yeast two-hybrid (Y2H) studies. As bait, the recently described CHD7 plasmids (CHD7–2-pGBKT7, CHD7–3-pGBKT7, CHD7–4-pGBKT7, CHD7-CR1–3-pGBKT7) were used. Due to the large size of KMT2D, ten plasmids were generated while KDM6A full length was used as prey. Autoactivation tests using the empty pGBKT7 bait vector together with the prey plasmids showed no autoactivation of the yeast strain reporter-genes. The yeast two-hybrid experiments revealed no direct interaction of the tested CHD7 parts with the KMT2D and KDM6A parts, leading to the suggestion that the assembly of all three proteins is indirect via additional complex members (Supplementary Material, Fig. S2). As a positive control, KMT2D and KDM6A plasmids were used.

EP300 co-immunoprecipitates with CHD7 as well as with KMT2D and KDM6A

To examine potential protein interactions between EP300 and CHD7, KDM6A or KMT2D, Co-IP studies were performed using

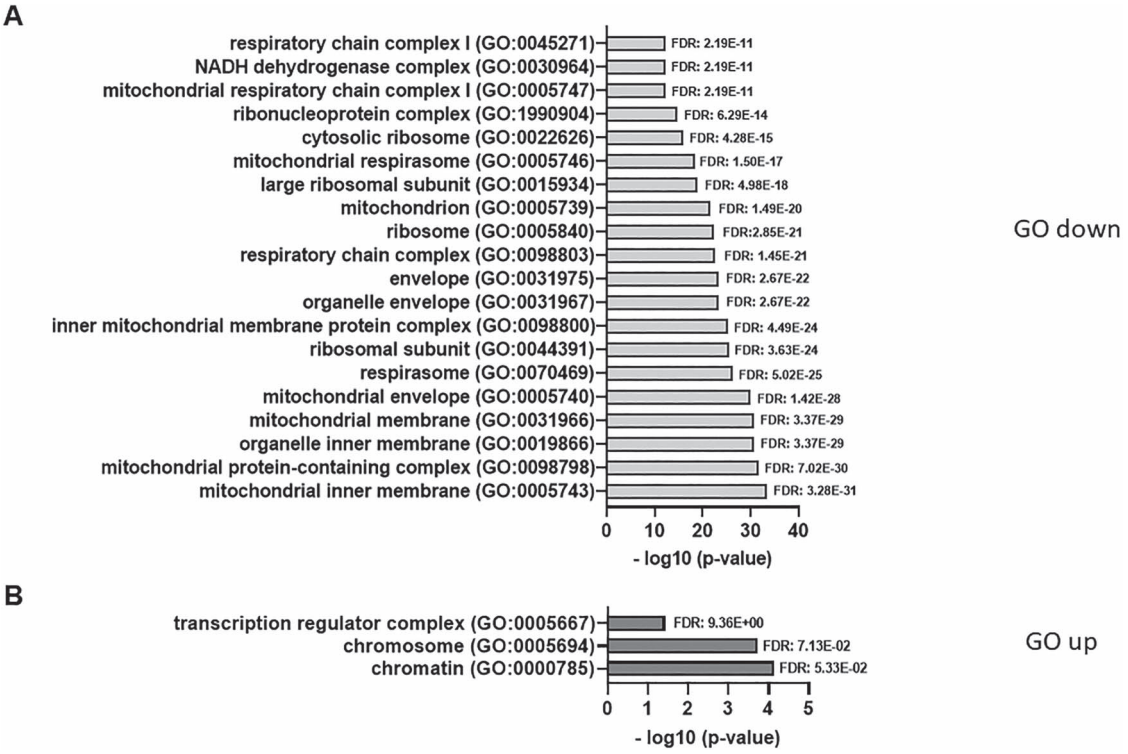

Figure 2. RNA-Seq GO: Bar graphs show significantly enriched GO cellular components of commonly down- (**A**) and upregulated (**B**) genes.

protein extracts from transfected HEK293 cells. HEK293 cells were double transfected with EP300 full-length-pCMV-c-Myc and one of the following plasmids: CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis or KDM6A full-length-pCMV-HA or KMT2D full-length-pCMV-HA. Co-IP experiments were performed in both directions using an anti-c-Myc antibody for immunoprecipitation or detection of EP300 and an anti-CHD7 or anti-KDM6A or anti-KMT2D antibody. Bands corresponding to the expected size of full-length EP300 (~300 kDa) (Fig. 5A) or for the reciprocal experiments for the estimated sizes of CHD7 (~336 kDa) (Fig. 5B), KDM6A (154 kDa) (Fig. 5C) or KMT2D (~593 kDa) (Fig. 5D) were detected.

The interaction between EP300 and CHD7, KMT2D or KDM6A was confirmed by PLA and takes place in the nucleus

Duolink® PLA experiments were carried out in HEK293 cells to confirm the Co-IP results. Cells were double transfected with the EP300 full-length-pCMV-c-Myc plasmid and CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis or KDM6A full-length-pCMV-HA or KMT2D full-length-pCMV-HA following incubation with an anti-c-Myc antibody for EP300 and a specific antibody against CHD7, KDM6A or KMT2D. In all three cases, many PLA signals (Fig. 5E, red dots) were detected in approximately 10–30% of the cell nuclei, representing the transfection efficiency. As a positive control, the known interaction of KDM6A with KMT2D was used, while anti-SC35 served as a negative control. The protein interaction between EP300 and KMT2D, KDM6A and CHD7 was confirmed and takes place in the nucleus as expected.

EP300 is most probably not bridging the protein interaction between CHD7 and KMT2D/KDM6A

The yeast two-hybrid system was used to examine a direct interaction between CHD7 and EP300. The above described CHD7 plasmids (CHD7–2-pGBKT7, CHD7–3-pGBKT7, CHD7–4-pGBKT7, CHD7-CR1–3-pGBKT7) were used as bait and one of the EP300 parts (EP300–1-pGADT7, EP300–2-pGADT7 and EP300–3-pGADT7) as prey. The autoactivation test using the empty pGBKT7 bait vector together with the prey plasmids demonstrated that no activation of the yeast strain reporter-genes takes place. As we did not observe any direct interaction of the tested CHD7 parts with the EP300 parts, we suggest that a linker function of EP300 between KMT2D, KDM6A and CHD7 is less likely (Supplementary Material, Fig. S2).

Discussion

Several reports have described that the complex CHARGE phenotype can be mimicked by other developmental syndromes. The clinical overlap between CHARGE and Kabuki syndrome is well characterized (12) and Moccia *et al.* additionally reported on the clinical overlap between CHARGE syndrome and Rubinstein-Taybi (*EP300*), Verheij (*PUF60*) and RERE-associated syndromes (11). A phenotypic overlap between Rubinstein-Taybi and Kabuki syndrome was also reported (18). A shared molecular pathology of these syndromes was suggested, as all involved proteins have roles in chromatin remodeling and modification (11).

KMT2D belongs to the family of type 2 lysine methyltransferases (KMT2), which catalyze mono-, di-, or trimethylation of the fourth lysine on histone 3 (H3K4me1, H3K4me2 and H3K4me3) (19). KMT2D has been identified as a major histone methyltransferase for H3K4 monomethylation at enhancer

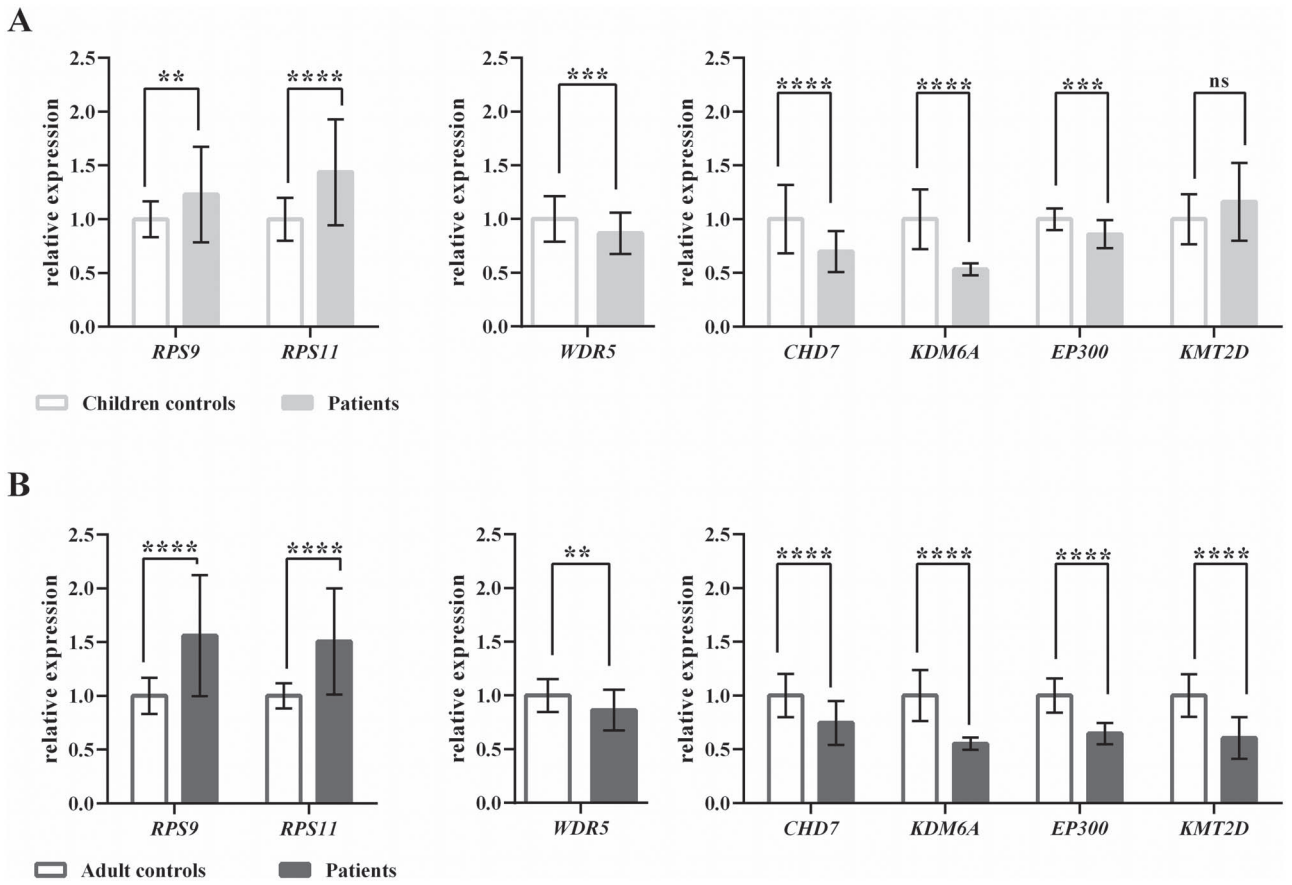

Figure 3. QRT-PCR: QRT-PCR was used to validate downregulation of the clinical relevant genes *CHD7* (CHARGE syndrome), *KMT2D* and *KDM6A* (Kabuki syndrome) and *EP300* (Rubinstein-Taybi) as well as upregulation of two representatives for ribosomal proteins (*RPS11* and *RPS9*) in CHARGE patients in comparison with children **(A)** and adult controls **(B).** To confirm upregulated, downregulated and non-regulated genes *WDR5* was included as a non-regulated gene. Three biological replicates were used. Gene expression was normalized against the housekeeping gene *TBP* (TATA box binding protein). The relative gene expression is visualized as the mean +/− standard deviation. Significance was calculated via Mann–Whitney test (ns: pval > 0.05; *: pval ≤ 0.05; **: pval ≤ 0.01, ***: pval ≤ 0.001; ****: pval ≤ 0.0001). Except for *KMT2D*, the results of the RNA-Seq analysis could be validated by qRT-PCR for all other selected genes.

elements and is therefore implicated in enhancer activation (20). Its enzymatic activity is driven by a conserved SET domain and the catalytic function can be increased by binding to the WRAD complex, consisting of WDR5, RbBP5, ASH2L and DPY-30 (21–23). In addition to the core KMT2D-WRAD complex, other binding partners can be included in the multi-protein complex COMPASS (complex of proteins associating with Set1). One of these additional non-WRAD members is KDM6A. KDM6A encodes a type 2 demethylase protein that activates gene transcription through the removal of di- and trimethyl groups from lysine 27 of histone 3 residues, an inactive mark (24). Furthermore, ChIP-Seq studies demonstrated co-localization of KMT2D and EP300 with H3K4me1 and histone 3 lysine 27 acetylation (H3K27ac) marks (20,25). EP300 and its family member CREBBP function as transcriptional coactivators as well as histone acetyltransferases, setting H3K27ac marks at both enhancers and promoters, correlating with active gene transcription (26). An epigenetic crosstalk between KMT2D, KDM6A and EP300 was demonstrated in which KDM6A recruits KMT2D and combines KMT2D and EP300 function in a demethylase activity-independent manner. Protein–protein interactions between these three proteins were shown and it was assumed that the KDM6A-KMT2D complex enhances EP300-dependent H3K27 acetylation, while KMT2D needs the support of EP300 function for H3K4 monomethylation, which further enhances EP300-dependent transcription (25). CHD7 is an ATP-dependent chromatin remodeler that functions tissue specific in enhancer-mediated transcription (27). Furthermore, it was shown by ChIP-Seq that CHD7 co-localizes with EP300 (28). In an earlier study, we demonstrated that CHD7 co-immunoprecipitates with the WAR complex members WDR5, ASH2L and RbBP5 (12). Therefore, we wondered whether there is a direct protein–protein interaction between CHD7 and KMT2D, KDM6A and EP300.

We performed co-immunoprecipitation studies and PLA as well as yeast two-hybrid studies. Indeed, we could demonstrate by co-immunoprecipitation experiments that CHD7 is associated with KDM6A, KMT2D and EP300. Due to the large size of KMT2D, the CHD7-KMT2D interaction was demonstrated for a small part of CHD7 (aa 1535–1852) and only a weak band was observed for the full-length protein. Nevertheless, using the PLA method, we confirmed the interaction in the cell nuclei of HEK293 cells transfected pairwise with plasmids containing the full-length sequence of CHD7, KMT2D, KDM6A and EP300. Further, to test whether CHD7 interacts directly with one of the above described proteins, we performed yeast two-hybrid studies. In the yeast system, we did not observe any direct interaction between CHD7 and KMT2D or KDM6A with the used plasmids. Neither did we observe a direct interaction between

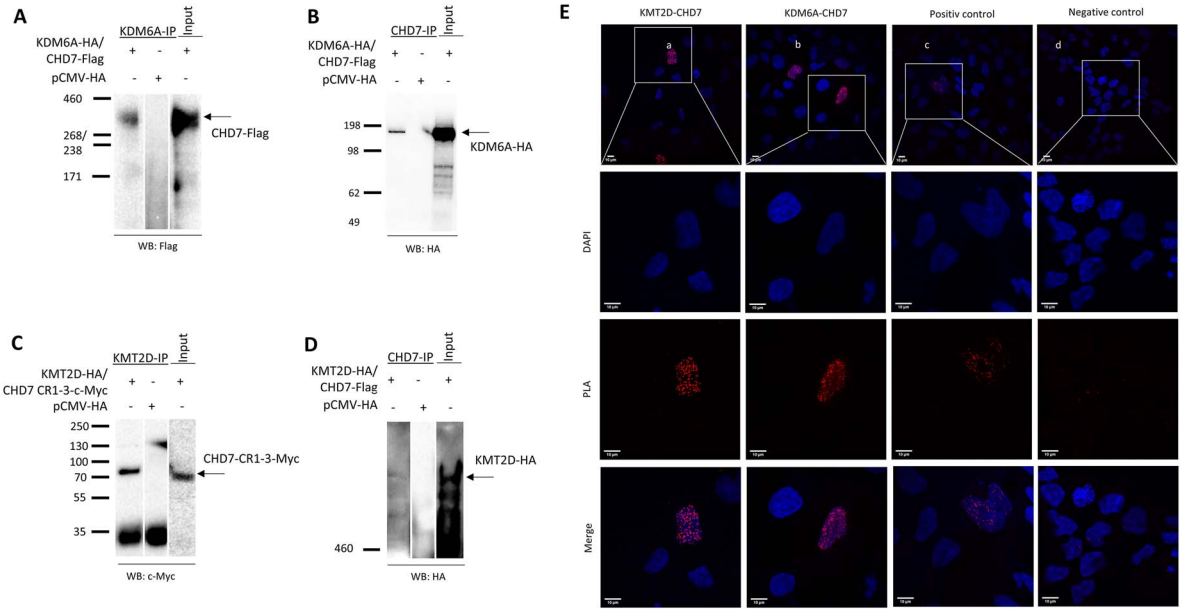

Figure 4. Co-IP and PLA CHD7- KMT2D, CHD7-KDM6A: **A:** HEK293 cells were double transfected with CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis and KDM6A full-length-pCMV-HA plasmids. KDM6A was immunoprecipitated by using an anti-KDM6A antibody, while an anti-FLAG antibody was used for CHD7 detection. Western blot analysis showed a protein band of ~336 kDa corresponding to the estimated size of the CHD7 full-length protein. **B:** Reciprocal experiment, a specific protein band of ~154 kDa, the estimated size of full-length KDM6A was detected. Unspecific antibody binding was excluded by performing immunoprecipitation on protein extracts isolated from cells transfected with the empty pCMV-HA plasmid. For the KMT2D-CHD7 analysis, two different transfections were performed due to the large size of KMT2D of 593 kDa. **C:** Cells were transfected with KMT2D full-length-pCMV-HA and CHD7-CR1–3-pCMV-c-Myc. A protein band of the expected size of ~67 kDa was detected **D:** Cells were transfected with KMT2D full-length-pCMV-HA and CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis. Endogenous CHD7 was immunoprecipitated using an anti-CHD7 and KMT2D was detected with an anti-HA antibody. A weak protein band of 593 kDa was detected **E:** Duolink PLA method on HEK293 cells. The cells were either double transfected with the CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis and KMT2D full-length pCMV-HA plasmids or with the plasmids CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis and KDM6A full-length-pCMV-HA. Positive PLA signals, indicative for a protein–protein interaction could be detected in the nucleus (red dots) using an anti-CHD7 antibody and an anti-KMT2D antibody or an anti-CHD7 and anti-HA antibody. As a positive control, the known interaction of KDM6A with KMT2D was used, while anti-SC35 served as a negative control. The cell nuclei were stained with DAPI (blue).

CHD7 and EP300, which leads us to the suggestion that EP300 does not serve as a bridging protein in a CHD7 and KMT2D, KDM6A containing complex.

By incubating a recombinant CHD7 to a protein array, Yan *et al.* could demonstrate that CHD7 binds directly to WDR5. They mapped the interaction area between both proteins to the CHD7 region containing amino acids 1535–1852 (29). This region overlaps with the CHD7 part (amino acids 1591–2181), which we used recently for the co-immunoprecipitation studies between CHD7 and WAR-complex members (12), as well as between CHD7 and KMT2D. If this region is indeed a direct interaction area between CHD7 and WDR5, it is possible that WDR5 mediates the association between CHD7 and KMT2D, KDM6A and EP300. This could explain our observation of an indirect interaction between CHD7 and KMT2D, KDM6A and EP300 using this CHD7 region as a template for Co-IPs and Yeast two hybrid studies.

In summary, our interaction studies demonstrate a closer protein association of CHD7 with the Kabuki-related genes *KMT2D* and *KDM6A* as well as with the Rubinstein-Taybi-related gene *EP300*, which might explain the phenotypic overlap of the three syndromes.

To identify regulatory disturbances in individuals affected by CHARGE syndrome, we performed RNA-Seq analysis on blood samples of 19 children with CHARGE syndrome and a confirmed disease-causing *CHD7* variant in comparison with samples of 10 healthy children.

In line with the well-known *CHD7* haploinsufficiency in CHARGE syndrome we identified a downregulation of *CHD7*. Interestingly, we also observed downregulation of other genes involved in epigenetic regulation, among them are the above described genes *KMT2D, KDM6A* and *CREBBP, EP300*.

A reverse effect was described in KDM6A mutant murine neural crest cells, where Sphargel and colleagues observed a downregulation of Chd7 (30). Further studies are needed to analyze whether these observations are due to a direct effect (e.g. the involved genes can regulate each other) or are secondary (e.g. due to epigenetic changes).

In addition to the above described misregulation of genes involved in chromatin modification, pathway analysis of our RNA-Seq data revealed that ribosome biogenesis is one of the most extensively dysregulated pathways. A role for CHD7 in ribosomal biogenesis has already been described. Zentner and colleagues demonstrated that CHD7 is associated with hypomethylated, active rDNA and that siRNA-mediated depletion of CHD7 results in hypermethylation of the rDNA promoter and a reduction of 20–30% in pre-rRNA levels (31). It has been shown that the polycomb group protein Bmi1 binds to the promoter region of multiple ribosomal protein genes and positively regulates their expression. Furthermore, it has been shown that the loss of Bmi1 in erythroid progenitor cells leads to a downregulation of multiple ribosomal protein genes and, as a consequence, to impaired ribosomal biogenesis (32). As CHD7 belongs to the trithorax group, counteracting polycomb group activities, a reversal effect on the expression level of ribosomal proteins is conceivable. Further studies are needed to explore the question of to what extent the misregulation of several ribosomal proteins is responsible for some of the symptoms seen in CHARGE patients.

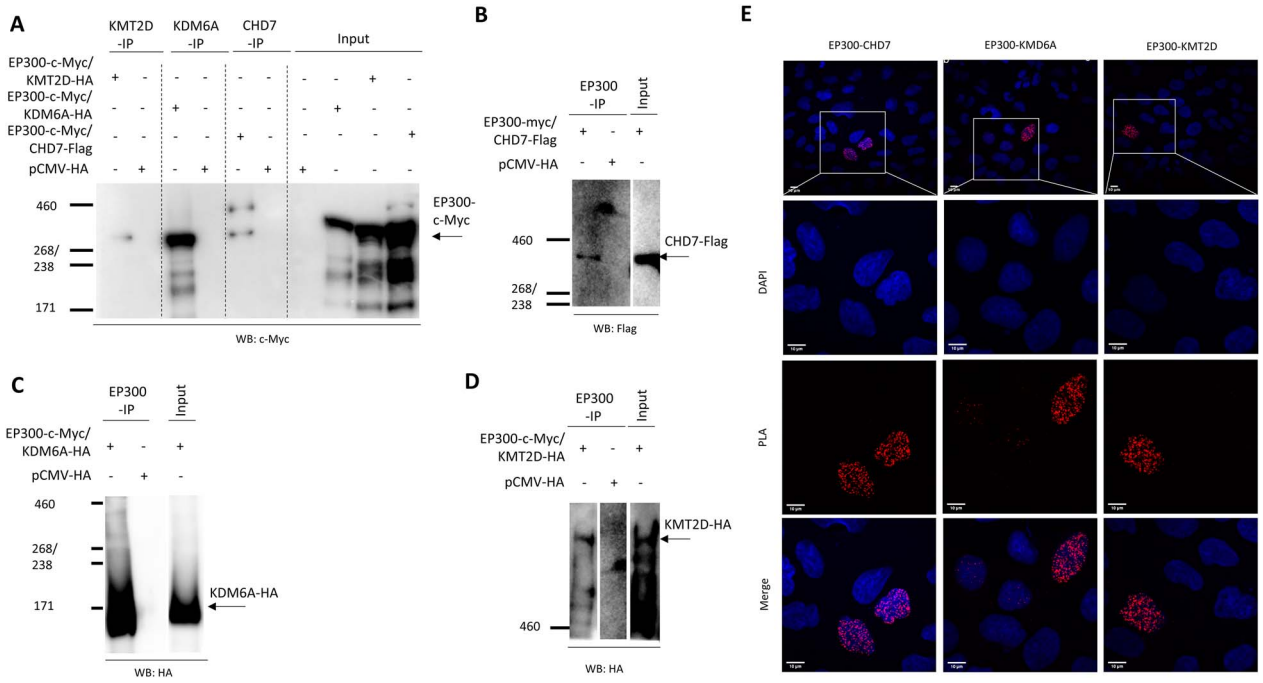

Figure 5. Co-IP and PLA CHD7, KMT2D, KDM6A and EP300: HEK293 cells were double transfected with EP300 full-length-pCMV-c-Myc and one of the following plasmids: CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis or KDM6A full-length-pCMV-HA or KMT2D full-length-pCMV-HA. Co-IP experiments were performed in both directions using an anti-c-Myc antibody for immunoprecipitation or detection of EP300 and an anti-CHD7 or anti-KDM6A or anti-KMT2D antibody. Bands corresponding to the expected size of **A:** full-length EP300 (~300 kDa),or the reciprocal experiments for the estimated sizes of **B:** CHD7 (~336 kDa), **C:** KDM6A (154 kDa) or **D:** KMT2D (~593 kDa) were detected. **E:** Duolink PLA method on HEK293 cells. Cells were double transfected with the EP300 full-length-pCMV-c-Myc plasmid and CHD7 full-length-pcDNA3.1(+)-FLAG-6xHis or KDM6A full-length-pCMV-HA or KMT2D full-length-pCMV-HA following incubation with an anti-c-Myc antibody for EP300 and a specific antibody against CHD7, KDM6A or KMT2D. In all three cases, positive PLA signals, indicative for a protein–protein interaction could be detected in the nucleus (red dots).

In summary, we were able to demonstrate that CHD7 is associated with KMT2D, KDM6A and EP300. Furthermore, we identified a misregulation of these genes in CHARGE syndrome patients, which provides an explanation for the overlapping phenotypes of CHARGE, Kabuki and Rubinstein-Taybi syndromes.

Material and Methods

Patients

Nineteen children with CHARGE syndrome and a confirmed *CHD7* variant (missense, nonsense or splice site) were recruited from the University Medical Center Groningen (UMCG). An overview of the patients mutations are given in Supplementary Material, Table S1. Ten control children (5 girls and 5 boys) at the age of 2 to 15 years were included in the study, as well as 5 adults (1 male and 4 female). Informed consent was obtained from the parents of all patients at the UMCG or at the University Medical Center Göttingen (UMG). The study has been exempted from ethical review by the Medical Ethical Review Committee of the UMCG and UMG.

RNA-seq library preparation and sequencing

Quality and integrity of RNA was assessed with the Fragment Analyzer from Advanced Analytical by using the standard sensitivity RNA Analysis Kit (DNF-471). All samples selected for sequencing exhibited an RNA integrity number over 8. RNA-seq libraries were performed using 500 ng total RNA of a non-stranded RNA Seq, massively-parallel mRNA sequencing approach from Illumina (TruSeq RNA Library Preparation Kit v2, Set A; 48 samples, 12 indexes, Cat. NRS-122-2001). Specifically we first optimized the ligation step diluting the adapter concentration to increase ligation efficiency (>94%), and finally we reduced the number of PCR cycles (10 cycles) to avoid PCR duplication artifacts as well as primer dimers in the final library product. Libraries were prepared on the automation Beckman Coulter's Biomek FXP workstation.

For accurate quantitation of cDNA libraries a fluorometric based system, the QuantiFluor™dsDNA System from Promega was used. The size of final cDNA libraries was determined by using the dsDNA 905 Reagent Kit (Fragment Analyzer from Advanced Bioanalytical) exhibiting a sizing of 300 bp in average. Libraries were pooled and sequenced on the Illumina HiSeq 4000 (SE; 1 × 50 bp; 30–35 Mio reads/sample).

Data analysis

Raw read and quality check. Sequence images were transformed with Illumina software BaseCaller to BCL files, which were demultiplexed to fastq files with bcl2fastq v2.20. The sequencing quality was asserted using FastQC (Andrews S. (2014). FastQC A Quality Control tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Mapping and normalization. Sequences were aligned to the reference genome *Homo sapiens* (GRCh38, https://www.ensembl.org/Homo_sapiens/Info/Index) using the RNA-Seq alignment tool (version 2.5.2a) (33) allowing for 2 mismatches within 50 bases. Subsequently, read counting was performed using featureCounts (34) with gene version 99. Read counts were analyzed in the R/Bioconductor environment (version 3.6.3, www.bioconductor.org) using the DESeq2 (35) package version

1.29.4. Candidate genes were filtered using an absolute log2 fold-change >1 and FDR-corrected P-value < 0.05. Gene annotation was performed using *Homo sapiens* entries via biomaRt R package version 2.44.1 (36).

A heatmap of candidate genes was generated using Heatmapper (37). Genes were hierarchical clustered based on Pearson distance of z-scores data and average linkage. Gene ontology analysis was performed using ToppFun of ToppGene Suite (38).

The data set can be shared upon request.

QRT-PCR

To confirm the RNA-Seq results, 450 ng of RNA of CHARGE patients and controls was reverse transcribed using SuperScript™ IV VILO™ Master-Mix and Random Primer (Invitrogen). QRT-PCR was performed using 5 µL SYBR ® Green Quantitect (Qiagen), 2.5 µL primer mix (forward and reverse, each 1 mM) and 2.5 µL cDNA. The reactions were incubated in a 384-well plate (4titude) at 50°C for 2 min, at 95°C for 15 min, followed by 40 cycles of 95°C for 15 s, 58°C for 30 s, and 72°C for 30 s, finished by one cycle of 95°C for 30 s, 60°C for 15 s, and 95°C for 15 s using the ABI Prism 7900HT Sequence Detection System. Control DNA was used to perform a standard curve for each gene with a serial dilution of 1:5. A triple determination was performed for each sample and normalized against *TBP* (TATA-box binding protein). The generated data were analyzed using SDS Version 2.1 software (PE Applied Biosystems) and the relative expression was determined by the standard curve method. The relative gene expression is visualized as the mean +/− standard deviation. For statistical analysis GraphPad Prism 6 software was used. Significance was calculated via Mann–Whitney test (ns: pval > 0.05; *: pval ≤ 0.05; **: pval ≤ 0.01, ***: pval ≤ 0.001; ****: pval ≤ 0.0001). Each experiment was repeated three times. Primer sequences will be sent on request.

Cell culture, transient transfection and plasmids

Cell culture with HEK293 cells and transfection with jetPRIME® or jetOPTIMUS® (Polyplus) were performed as described previously (39). For transient transfection the following plasmids were used: CHD7-flag-his-pcDNA 3.1(+) (41), CHD7-CR1–3-pCMV-myc, KDM6A-pCMV-HA (was a gift from Kristian Helin, Addgene plasmid #24168, (40), KMT2D-pCMV-HA (kindly provided by Laura Pasqualucci, (42), p300-pCMV-myc (originally named pCMVβ-p300-myc, was a gift from Tso-Pang Xao, Addgene plasmid #30489) and pCMV-HA as empty vector (Clontech). The CHD7-CR1–3-pCMV-myc was generated by subcloning the corresponding CHD7-CR1–3 region (amino acids 1591–2181, NP_065971.2) from the pGBKT7 vector (plasmid published in (12)) into the pCMV-myc vector (Clontech) using EcoR1/SalI.

Western blot analysis and immunoprecipitation

Whole-cell lysates from transfected HEK293 were prepared using modified RIPA buffer with proteinase inhibitor cocktail Complete (SIGMA) as described previously (39). Two different gels and markers were used depending on the protein size. For large size proteins (> 150 kDa), 30 µg protein was loaded on a 3–8% Tris-acetate gel (Life Technologies) and gel electrophoresis was performed at 140 V for about 4 h. High Marker (Invitrogen) was used to determine the size of the protein. Afterwards, proteins were transferred to an activated PVDF membrane (Biorad) using a Tank system (Biorad) at 25 V for 19 h at 4°C. For proteins with a size of up to 150 kDa, 30 µg protein was loaded on a 4–12% Tris-Acetate gel (Life Technologies) and gel electrophoresis was performed at 140 V for about 2 h. Seeblue Plus 2 marker (Invitrogen) was used to determine the size of the protein. Afterwards, proteins were transferred to a nitrocellulose membrane (Amersham) using a Tank system (Biorad) at 250 A for 1 h at 4°C. After blocking with 5% milk/TBST for 60 min, the following primary antibodies were used in 2% milk/TBST: Rat anti-HA (11 867 431 001, Sigma-Aldrich) at a dilution of 1:2000, mouse anti-Flag M2 (Invitrogen) at a dilution of 1:1000 and mouse anti-cmyc (Cell Signaling) at a dilution of 1:2000. For co-immunoprecipitation assay 800 µg protein was diluted in IP-Buffer (0.1% Triton, 50 mM Tris–HCl pH 7.4, 300 mM NaCl, 5 mM EDTA pH 8.0) in the presence of proteinase inhibitor cocktail Complete and magnetic beads (NEB). After 1 h at 4°C, supernatant was incubated with 2 µg antibody overnight at 4°C. Afterwards, 25 µL protein G magnetic beads were added and incubated for 1 h at 4°C. Beads were washed four times with IP-Buffer and once with TBS. Protein was eluted with 24 µL SDS-MIX ((NuPAGE® LDS Sample Buffer (4x) (Life Technologies)) and 10% of reducing agent (1 M DTT) at 70°C for 10 min. Subsequently, western blotting was performed. Co-immunoprecipitation assay was done in three independent experiments.

Immunocytochemistry and PLA

After the permeabilization step with 0.5% Triton X-100 in PBS for 10 min on fixed HEK293 cells, cells were washed twice with 0.1% Tween in PBS (PBT) and blocked with 3% BSA in PBT for 45 min at RT. Afterwards, the following primary antibodies in 1% BSA in PBT were used overnight at 4°C: mouse anti-HA (#901501,Covance) at a dilution of 1:100, rabbit anti-HA (#3724, Cell Signaling) at a dilution of 1:800, mouse anti-Flag M2 (Invitrogen) was used at a dilution of 1:500 and mouse anti-cmyc (Cell Signaling) was used at a dilution of 1:100, rabbit anti-CHD7 (Cell Signaling) at a dilution of 1:1000, rabbit anti-KDM6A (Cell Signaling) at a dilution of 1:100, and rabbit anti-KMT2D (Sigma-Aldrich) at a dilution of 1:100. After primary antibody incubation, cells were washed with PBT and secondary antibody incubation (1:800 goat anti-rabbit IgG (H + L)-Alexa Fluor 488, 1:800 goat anti-mouse IgG (H + L)-Alexa Fluor 546 or 1:800 goat anti-rat IgG (H + L); Invitrogen) for 1 h at RT was performed. Finally, cells were washed twice with PBT and mounted using Fluoroshield™ with DAPI (Merck Millipore).

PLA experiments were performed using PLA Duolink® kit (SIGMA). Fixed HEK293 cells were treated as for immunocytochemistry until primary antibody incubation and afterwards cells were treated according to the manufacturer's protocol. After the last step, cells were mounted using Fluoroshield™ with DAPI (Merck Millipore). Immunofluorescent images were acquired using the Olympus BX60 microscope with ×600 magnification. All images were processed with FIJI (an image-processing package based on ImageJ). All experiments were performed three times.

Yeast two-hybrid

In order to perform direct yeast two-hybrid 1 µg of prey and bait were used. The following plasmids were used: CHD7–2-pGBKT7 (amino acids 732–1567, NP_060250.2), CHD7–3-pGBKT7 (amino acids 1533–2380, NP_060250.2), and CHD7–4-pGBKT7 (amino acids 2325–2997, NP_060250.2) and CHD7-CR1–3-pGBKT7 (amino acids 1591–2181, NP_060250.2) which

were described previously (12). Additionally, the following plasmids were generated: KDM6A-pGBTK7 (NP_001278345.1), KMT2D-1-pGADT7 (amino acids 1–393, NP_003473.3), KMT2D-2-pGADT7 (amino acids 346–883, NP_003473.3), KMT2D-3-pGADT7 (amino acids 877–1583, NP_003473.3), KMT2D-4-pGADT7 (amino acids 1571–2248, NP_003473.3), KMT2D-5.1-pGADT7 (amino acids 2240–2610, NP_003473.3), KMT2D-5.2-pGADT7 (amino acids 2600–2990, NP_003473.3), KMT2D-6-pGADT7 (amino acids 2959–3709, NP_003473.3), KMT2D-7-pGADT7 (amino acids 3580–4409, NP_003473.3), KMT2D-8-pGADT7 (amino acids 4294–4879, NP_003473.3) and KMT2D-9-pGADT7 (amino acids 4850–5538, NP_003473.3), dividing the human KMT2D full-length protein into ten overlapping fragments. Furthermore, EP300–1-pGADT7 (amino acids 1–846, NP_001420.2), EP300–2-pGADT7 (amino acids 793–1567, NP_001420.2) and EP300–3-pGADT7 (amino acids 1528–2415, NP_001420.2) were generated, dividing the human EP300 full-length protein into three overlapping fragments. A toxicity and autoactivation test of the Y2H Gold strain (Clontech) was performed for each newly generated plasmid. The direct yeast two-hybrid experiments were done according to the protocol previously described (12). All experiments were performed three times.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank all patients and their parents for participating in this research project. We thank Bernd Wollnik and Karin Boss for critically reading the manuscript.

Conflict of Interest statement. The authors declare no conflict of interest.

Funding

The Deutsche Forschungsgemeinschaft (DFG) [PA 2030/2–1 to S.P.].

References

1. Pagon, R.A., Graham, J.M., Zonana, J. and Yong, S.L. (1981) Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. *J. Pediatr.*, **99**, 223–227.
2. Bergman, J.E.H., Janssen, N., Hoefsloot, L.H., Jongmans, M.C.J., Hofstra, R.M.W. and van Ravenswaaij-Arts, C.M.A. (2011) CHD7 mutations and CHARGE syndrome: the clinical implications of an expanding phenotype. *J. Med. Genet.*, **48**, 334–342.
3. Jongmans, M.C.J., Admiraal, R.J., van der Donk, K.P., Vissers, L.E.L.M., Baas, A.F., Kapusta, L., van Hagen, J.M., Donnai, D., de Ravel, T.J., Veltman, J.A. *et al.* (2006) CHARGE syndrome: the phenotypic spectrum of mutations in the CHD7 gene. *J. Med. Genet.*, **43**, 306–314.
4. Lalani, S.R., Safiullah, A.M., Fernbach, S.D., Harutyunyan, K.G., Thaller, C., Peterson, L.E., McPherson, J.D., Gibbs, R.A., White, L.D., Hefner, M. *et al.* (2006) Spectrum of CHD7 mutations in 110 individuals with CHARGE syndrome and genotype-phenotype correlation. *Am. J. Hum. Genet.*, **78**, 303–314.
5. Vissers, L.E.L.M., van Ravenswaaij, C.M.A., Admiraal, R., Hurst, J.A., de Vries, B.B.A., Janssen, I.M., van der Vliet, W.A., Huys, E.H.L.P.G., de Jong, P.J., Hamel, B.C.J. *et al.* (2004) Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat. Genet.*, **36**, 955–957.
6. Blake, K.D., Davenport, S.L., Hall, B.D., Hefner, M.A., Pagon, R.A., Williams, M.S., Lin, A.E. and Graham, J.M. (1998) CHARGE association: an update and review for the primary pediatrician. *Clin. Pediatr.*, **37**, 159–173.
7. Verloes, A. (2005) Updated diagnostic criteria for CHARGE syndrome: a proposal. *Am. J. Med. Genet. A*, **133A**, 306–308.
8. van Ravenswaaij-Arts, C.M.A., Blake, K., Hoefsloot, L. and Verloes, A. (2015) Clinical utility gene card for: CHARGE syndrome - update 2015. *EJHG*, **23**, 3–4.
9. Janssen, N., Bergman, J.E.H., Swertz, M.A., Tranebjaerg, L., Lodahl, M., Schoots, J., Hofstra, R.M.W., van Ravenswaaij-Arts, C.M.A. and Hoefsloot, L.H. (2012) Mutation update on the CHD7 gene involved in CHARGE syndrome. *Human mutation*, **33**, 1149–1160.
10. Bartels, C.F., Scacheri, C., White, L., Scacheri, P.C. and Bale, S. (2010) Mutations in the CHD7 gene: The experience of a commercial laboratory. *Genetic testing and molecular biomarkers*, **14**, 881–891.
11. Moccia, A., Srivastava, A., Skidmore, J.M., Bernat, J.A., Wheeler, M., Chong, J.X., Nickerson, D., Bamshad, M., Hefner, M.A., Martin, D.M. *et al.* (2018) Genetic analysis of CHARGE syndrome identifies overlapping molecular biology. *Genetics in Medicine: Off J American College of Medical Genetics*, **20**, 1022–1029.
12. Schulz, Y., Freese, L., Mänz, J., Zoll, B., Völter, C., Brockmann, K., Bögershausen, N., Becker, J., Wollnik, B. and Pauli, S. (2014) CHARGE and Kabuki syndromes: a phenotypic and molecular link. *Hum. Mol. Genet.*, **23**, 4396–4405.
13. Patel, N. and Alkuraya, F.S. (2015) Overlap between CHARGE and Kabuki syndromes: more than an interesting clinical observation? *Am. J. Med. Genet. A*, **167A**, 259–260.
14. Butcher, D.T., Cytrynbaum, C., Turinsky, A.L., Siu, M.T., Inbar-Feigenberg, M., Mendoza-Londono, R., Chitayat, D., Walker, S., Machado, J., Caluseriu, O. *et al.* (2017) CHARGE and kabuki syndromes: gene-specific DNA methylation signatures identify epigenetic mechanisms linking these clinically overlapping conditions. *Am. J. Hum. Genet.*, **100**, 773–788.
15. Ng, S.B., Bigham, A.W., Buckingham, K.J., Hannibal, M.C., McMillin, M.J., Gildersleeve, H.I., Beck, A.E., Tabor, H.K., Cooper, G.M., Mefford, H.C. *et al.* (2010) Exome sequencing identifies MLL2 mutations as a cause of kabuki syndrome. *Nat. Genet.*, **42**, 790–793.
16. Bögershausen, N. and Wollnik, B. (2013) Unmasking Kabuki syndrome. *Clin. Genet.*, **83**, 201–211.
17. Miyake, N., Mizuno, S., Okamoto, N., Ohashi, H., Shiina, M., Ogata, K., Tsurusaki, Y., Nakashima, M., Saitsu, H., Niikawa, N. *et al.* (2013) KDM6A point mutations cause Kabuki syndrome. *Hum. Mutat.*, **34**, 108–110.
18. Negri, G., Magini, P., Milani, D., Crippa, M., Biamino, E., Piccione, M., Sotgiu, S., Perrìa, C., Vitiello, G., Frontali, M. *et al.* (2019) Exploring by whole exome sequencing patients with initial diagnosis of Rubinstein-Taybi syndrome: the interconnections of epigenetic machinery disorders. *Hum. Genet.*, **138**, 257–269.
19. Ruthenburg, A.J., Allis, C.D. and Wysocka, J. (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol. Cell*, **25**, 15–30.

20. Hu, D., Gao, X., Morgan, M.A., Herz, H.-M., Smith, E.R. and Shilatifard, A. (2013) The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol. Cell. Biol.*, **33**, 4745–4754.
21. Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D. and Roeder, R.G. (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.*, **13**, 713–719.
22. Lavery, W.J., Barski, A., Wiley, S., Schorry, E.K. and Lindsley, A.W. (2020) KMT2C/D COMPASS complex-associated diseases KCDCOM-ADs: an emerging class of congenital regulopathies. *Clin. Epigenetics*, **12**, 10.
23. Froimchuk, E., Jang, Y. and Ge, K. (2017) Histone H3 lysine 4 methyltransferase KMT2D. *Gene*, **627**, 337–342.
24. Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.-P., Reinberg, D., Di Croce, L. and Shiekhattar, R. (2007) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science (New York, N.Y.)*, **318**, 447–450.
25. Wang, S.-P., Tang, Z., Chen, C.-W., Shimada, M., Koche, R.P., Wang, L.-H., Nakadai, T., Chramiec, A., Krivtsov, A.V., Armstrong, S.A. *et al.* (2017) A UTX-MLL4-p300 transcriptional regulatory network coordinately shapes active enhancer landscapes for eliciting transcription. *Mol. Cell*, **67**, 308–321.e6.
26. Pasini, D., Malatesta, M., Jung, H.R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O.N. *et al.* (2010) Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of polycomb group target genes. *Nucleic Acids Res.*, **38**, 4958–4969.
27. Schnetz, M.P., Bartels, C.F., Shastri, K., Balasubramanian, D., Zentner, G.E., Balaji, R., Zhang, X., Song, L., Wang, Z., Laframboise, T. *et al.* (2009) Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. *Genome Res.*, **19**, 590–601.
28. Schnetz, M.P., Handoko, L., Akhtar-Zaidi, B., Bartels, C.F., Pereira, C.F., Fisher, A.G., Adams, D.J., Flicek, P., Crawford, G.E., Laframboise, T. *et al.* (2010) CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. *PLoS Genet.*, **6**, e1001023.
29. Yan, S., Thienthanasit, R., Chen, D., Engelen, E., Brühl, J., Crossman, D.K., Kesterson, R., Wang, Q., Bouazoune, K. and Jiao, K. (2020) CHD7 regulates cardiovascular development through ATP-dependent and -independent activities. *Proc. Natl. Acad. Sci. U. S. A.* First published on October 30, 2020. 28847–28858.
30. Shpargel, K.B., Starmer, J., Wang, C., Ge, K. and Magnuson, T. (2017) UTX-guided neural crest function underlies craniofacial features of kabuki syndrome. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, E9046–E9055.
31. Zentner, G.E., Hurd, E.A., Schnetz, M.P., Handoko, L., Wang, C., Wang, Z., Wei, C., Tesar, P.J., Hatzoglou, M., Martin, D.M. *et al.* (2010) CHD7 functions in the nucleolus as a positive regulator of ribosomal RNA biogenesis. *Hum. Mol. Genet.*, **19**, 3491–3501.
32. Gao, R., Chen, S., Kobayashi, M., Yu, H., Zhang, Y., Wan, Y., Young, S.K., Soltis, A., Yu, M., Vemula, S. *et al.* (2015) Bmi1 promotes erythroid development through regulating ribosome biogenesis. *Stem cells (Dayton, Ohio)*, **33**, 925–938.
33. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013, 2012) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* First published on October 25. **29**(1), 15–21. Epub 2012 Oct 25. https://doi.org/10.1093/bioinformatics/bts635.
34. Liao, Y., Smyth, G.K. and Shi, W. (2014) Feature counts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**(7), 923–30. Epub 2013 Nov 13. https://doi.org/10.1093/bioinformatics/btt656.
35. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
36. Durinck, S., Spellman, P.T., Birney, E. and Huber, W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**(8), 1184–91. Epub 2009 Jul 23. https://doi.org/10.1038/nprot.2009.97.
37. Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A. and Wishart, D.S. (2016) Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* **44**(W1), W147–53. Epub 2016 May 17. https://doi.org/10.1093/nar/gkw419.
38. Chen, J., Bardes, E.E., Aronow, B.J. and Jegga, A.G. (2009) Topp gene suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**(Web Server issue), W305–11. Epub 2009 May 22. https://doi.org/10.1093/nar/gkp427.
39. Ufartes, R., Berger, H., Till, K., Salinas, G., Sturm, M., Altmüller, J., Nürnberg, P., Thiele, H., Funke, R., Apeshiotis, N. *et al.* (2020) De novo mutations in FBRSL1 cause a novel recognizable malformation and intellectual disability syndrome. *Hum. Genet.* **139**(11), 1363–1379.
40. Agger, K., Cloos, P.A.C., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E. and Helin, K. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature.* **449**(7163), 731–4. Epub 2007 Aug 22. https://doi.org/10.1038/nature06145.
41. Bajpai, R., Chen, D.A., Rada-Iglesias, A., Zhang, J., Xiong, Y., Helms, J., Chang, C.-P., Zhao, Y., Swigut, T. and Wysocka, J. (2010) CHD7 cooperates with PBAF to control multipotent neural crest formation. *Nature*, **463**, 958–962.
42. Zhang, J., Dominguez-Sola, D., Hussein, S., Lee, J.-E., Holmes, A.B., Bansal, M., Vlasevska, S., Mo, T., Tang, H., Basso, K. *et al.* (2015) Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat. Med.* **21**(10), 1190–8. Epub 2015 Sep 14. https://doi.org/10.1038/nm.3940.