Histone Modifier Genes Alter Conotruncal Heart Phenotypes in 22q11.2 Deletion Syndrome

Tingwei Guo,[1] Jonathan H. Chung,[1] Tao Wang,[2] Donna M. McDonald-McGinn,[3] Wendy R. Kates,[4] Wanda Hawuła,[5] Karlene Coleman,[6] Elaine Zackai,[3] Beverly S. Emanuel,[3] and Bernice E. Morrow[1,*]

We performed whole exome sequence (WES) to identify genetic modifiers on 184 individuals with 22q11.2 deletion syndrome (22q11DS), of whom 89 case subjects had severe congenital heart disease (CHD) and 95 control subjects had normal hearts. Three genes including *JMJD1C* (jumonji domain containing 1C), *RREB1* (Ras responsive element binding protein 1), and *SEC24C* (SEC24 family member C) had rare (MAF < 0.001) predicted deleterious single-nucleotide variations (rdSNVs) in seven case subjects and no control subjects ($p = 0.005$; Fisher exact and permutation tests). Because *JMJD1C* and *RREB1* are involved in chromatin modification, we investigated other histone modification genes. Eighteen case subjects (20%) had rdSNVs in four genes (*JMJD1C*, *RREB1*, *MINA*, *KDM7A*) all involved in demethylation of histones (H3K9, H3K27). Overall, rdSNVs were enriched in histone modifier genes that activate transcription (Fisher exact $p = 0.0004$, permutations, $p = 0.0003$, OR $= 5.16$); however, rdSNVs in control subjects were not enriched. This implicates histone modification genes as influencing risk for CHD in presence of the deletion.

The 22q11.2 deletion syndrome (22q11DS; DiGeorge syndrome/velo-cardio-facial syndrome [MIM: 192430, 188400]) is a congenital malformation disorder that occurs in 1/4,000 live births. The affected individuals have a de novo, hemizygous 3 million base pair (Mb) deletion on chromosome 22q11.2. A total of 60%–70% have congenital heart disease (CHD), mostly of the conotruncal type. Among these, some have severe anomalies, such as tetralogy of Fallot (TOF) or persistent truncus arteriosus (PTA), while others have mild phenotypes such as isolated ventricular septal defects (VSD) or right-sided aortic arch. Haploinsufficiency of *TBX1* (MIM: 602054), encoding a T-box transcription factor, is believed responsible for the etiology of CHD in 22q11DS.[1–3] A likely explanation for phenotypic heterogeneity of CHD would be the presence of mutations in *TBX1* on the remaining 22q11.2 allele. However, causative second-hit mutations in *TBX1* among 1,022 22q11DS-positive subjects were not found,[4] implicating other genes as being responsible, perhaps lying elsewhere in the genome. A common duplication of the glucose transporter gene *SLC2A3* (MIM: 138170) was significantly enriched in 22q11DS-affected individuals with CHD, implicating this copy-number variation as a genetic modifier for some subjects.[5] However, this cannot explain the presence or absence of a heart or aortic arch defect in most other individuals with 22q11DS. Another possibility is that rare coding predicted deleterious single-nucleotide variants (rdSNVs) might act as modifiers of phenotype in 22q11DS.

To search for rare coding variants, we ascertained 186 individuals with the typical 3 Mb, 22q11.2 deletion and extremes of cardiac phenotypes for a whole exome sequencing (WES) study. Echocardiography or cardiology reports were provided on each individual with their informed consent (Internal Review Board, 1999–2001). WES was performed with Nimblegen v.3 libraries (NHLBI, Resequencing Service) on DNA from blood and two samples were removed because they were duplicates, leaving 184 for the analysis. The average read depth for the targeted exome was 61×, with 84% of the target regions covered at greater than 20×. Finally, we compared data from 89 CHD-affected case subjects (mostly TOF) with 22q11DS and 95 deleted individuals with a normal heart and aortic arch (control subjects) to identify genes that harbor rare coding variants that could alter risk for extremes of heart phenotypes (Figure 1A). In addition, we evaluated ethnicity of the cohort via principal-component analysis and found all except two were of European descent (Figure 1B).

Among all 184 DNA samples, 411,618 variants including 370,551 SNVs and 41,067 indels were identified (Figure S1). A total of 279,820 SNVs were retained after filtering on quality controls of genotype and SNVs depending upon the depth of sequencing reads. Most relevant to any phenotype are non-synonymous or splicing variants because they could impact the function of the protein. Of the total SNVs identified, 74,982 were non-synonymous variants including 72,074 non-synonymous SNVs and 2,908 loss-of-function (LOF) SNVs including stop-gain or stop-loss variants, frameshift indels, splice donor or acceptor variants, and initiator codon variants. To identify putative functional variants, we applied five standard computational prediction methods applicable to non-synonymous SNVs (SIFT [RRID: nlx_154618], PolyPhen-2

[1]Department of Genetics, Albert Einstein College of Medicine, Yeshiva University, Bronx, NY 10461, USA; [2]Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Yeshiva University, Bronx, NY 10461, USA; [3]Division of Human Genetics, The Children's Hospital of Philadelphia, and Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA 19104, USA; [4]Department of Psychiatry and Behavioral Sciences, and Program in Neuroscience, SUNY Upstate Medical University, Syracuse, NY 13210, USA; [5]Department of Genetics, Polish Mother's Memorial Hospital - Research Institute, Łódź 93-338, Poland; [6]Marcus Autism Center, Children's Healthcare of Atlanta, Atlanta, GA 30329, USA
*Correspondence: bernice.morrow@einstein.yu.edu
http://dx.doi.org/10.1016/j.ajhg.2015.10.013.

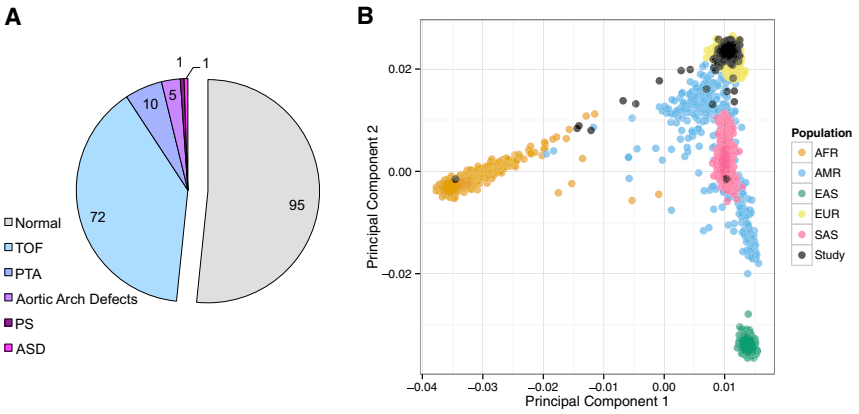

Figure 1. Phenotype and Ethnicity of 184 Individuals with 22q11DS

(A) CHD phenotype in the 22q11DS cohort. Whole exome sequencing was performed on a cohort of 184 individuals with 22q11DS and the 3 Mb deletion, including 89 mostly severe CHD-positive case subjects and 95 with a normal heart and normal aortic arch (controls). Among 89 CHD-positive case subjects, 72 had tetralogy of Fallot (TOF), 10 had a persistent truncus arteriosus (PTA), 5 had aortic arch defects, and 1 each had pulmonary stenosis (PS) or an atrial septal defect (ASD).

(B) Ethnicity of 22q11DS cohort. Principal components (PCs) were calculated for common variants in WES data from the 184 study subjects (black dots) and 1000 Genomes Project population dataset to check the ethnicity of the cohort. The majority of the cohort clustered with the European (EUR, yellow) population, but one had African (AFR, orange) ancestry and one subject had South Asian (SAS, pink) ancestry. Abbreviations are as follows: AMR, Admixed American; EAS, East Asian.

HVAR [RRID: OMICS_00136], MutationTaster [RRID: OMICS_00153], MutationAssessor [RRID: nlx_149228], FATHMM) and CADD (scaled C-scores greater or equal to 10)[6] from the dbNSFP database v.2.6.[7] Non-synonymous SNVs predicted to be damaging by at least two of the six applicable methods were included for further analysis. Then, LOF variants and putatively functional non-synonymous SNVs were combined for a total of 30,532 predicated deleterious variants. We then further selected variants that have minor allele frequency (MAF) of $\leq$0.1% in the database for the 1000 Genomes Project and National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project (ESP) 6500 Exome Project. This filtering resulted in a total of 20,402 rdSNVs in 10,620 genes (Figure S1).

All 184 individuals have a similar size 3 million base pair (Mb) 22q11.2 deletion as determined by molecular testing. We first evaluated the overall genetic architecture of rdSNVs in the entire cohort of case and control subjects together. The rdSNVs were distributed among 10,620 genes, in which at least two individuals had rdSNVs in 5,431 genes. To reduce the noise of the enrichment analysis, we further selected genes based upon the residual variation intolerance score (RVIS)[8] and haploinsufficient score.[9] A total of 671 genes with rdSNVs in at least two 22q11DS-positive individuals were used to investigate the potential biological relevance of the affected genes (Functional Annotation Tool in DAVID, The Database for Annotation, Visualization and Integrated Discovery[10]). We used all the genes in the genome as background to calculate p values via the Fisher exact statistical test. Gene Ontology (GO) terms were considered significantly enriched with a Benjamini multiple test correction p value of < 0.05.

We found significant enrichment in 58 "biological process" terms, 68 terms in "cellular component," and 44 terms in "molecular function" (Table S1). The top three GO terms relate to chromosome organization or chromatin function (Figure 2A). The GO terms related to "organ development" were also enriched (Figure 2A). To reduce the redundancy of GO terms from the enrichment analysis, we used the REVIGO tool, which uses a simple clustering algorithm to summarize GO terms.[11] GO terms were clustered into "tissue development," "chromosome organization," "regulation of transcription," and "RNA transport" (Figure 2B). The genes related to chromosome-associated GO terms might be involved in the etiology of the de novo deletion itself, while those involved in organ development could contribute to the phenotypic heterogeneity of 22q11DS.

To determine whether there were genes that could alter risk for extremes of heart phenotypes in particular, rdSNVs among 5,431 genes in at least two individuals with 22q11DS were collapsed into a single gene burden score. Then the Fisher exact test was used to compare the burden score between 22q11DS-affected subjects with CHD (case subjects with severe heart defects) and 22q11DS subjects without CHD (control subjects). Further, single value permutations, in which case versus control status is permuted 10,000 times, were used to perform an association analysis on the genes. Results were considered statistically significant with a p value of < 0.05. We did not identify any genes on the remaining allele with rdSNVs in two or more case or control subjects (data not shown). Three genes—*JMJD1C* (jumonji domain containing 1C [MIM: 604503]), *RREB1* (Ras responsive element binding protein 1 [MIM: 602209]), and *SEC24C* (SEC24 family member C [MIM: 607185])—were significant in both tests. All three genes had rdSNVs in seven different case subjects but none in control subjects (Figure 3A; Table S2). Among *JMJD1C*, *RREB1*, and *SEC24C*, 17 rdSNVs have been identified and affected 20 case subjects (Figure 3B; Table S2). One case subject has rdSNVs in both *JMJD1C* and *RREB1* (see Figure 3B). If we included all rare non-synonymous variants (MAF $\leq$ 0.1%) irrespective as to whether they are predicted as damaging or not, there were two additional case subjects with variants in *JMJD1C* and one case subject for *RREB1*. *JMJD1C* had the strongest p value for CHD among all the genes ($p = 0.001$; Figures 3A and 3B). Of interest,

A

Category	GO Term	P Value	Benjamini
BP	GO:0051276~chromosome organization	9.65E-12	2.74E-08
BP	GO:0048729~tissue morphogenesis	1.45E-09	2.06E-06
BP	GO:0006325~chromatin organization	1.81E-09	1.71E-06
BP	GO:0016568~chromatin modification	3.04E-09	2.16E-06
BP	GO:0035295~tube development	9.10E-09	5.18E-06
BP	GO:0043009~chordate embryonic development	1.28E-08	6.05E-06
BP	GO:0009792~embryonic development ending in birth or egg hatching	1.63E-08	6.62E-06
BP	GO:0032989~cellular component morphogenesis	6.93E-08	2.19E-05
BP	GO:0000904~cell morphogenesis involved in differentiation	9.87E-08	2.81E-05
BP	GO:0002009~morphogenesis of an epithelium	1.55E-07	3.68E-05
BP	GO:0048598~embryonic morphogenesis	1.81E-07	3.96E-05
BP	GO:0000902~cell morphogenesis	2.62E-07	5.32E-05
BP	GO:0007507~heart development	2.82E-07	5.01E-05
MF	GO:0003682~chromatin binding	2.51E-08	1.73E-06
MF	GO:0004386~helicase activity	1.34E-07	8.42E-06
MF	GO:0003678~DNA helicase activity	2.58E-04	0.006
CC	GO:0044427~chromosomal part	1.09E-06	2.30E-05
CC	GO:0005694~chromosome	9.06E-06	1.55E-04
CC	GO:0000785~chromatin	1.77E-04	0.002
CC	GO:0016585~chromatin remodeling complex	0.001	0.01

B

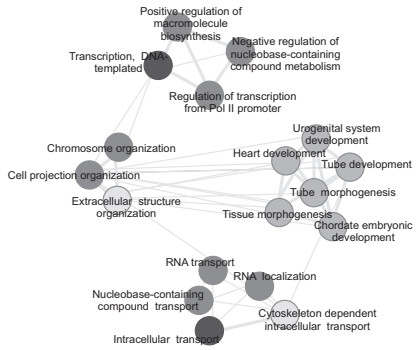

Figure 2. GO Enrichment Analysis of rdSNVs in the 22q11DS Cohort

(A) Top GO terms and their enrichment p value. Table contains categories (abbreviations are as follows: BP, biological process; MF, molecular function; CC, cellular component), p value of GO terms and values for Benjamini multiple test correction for the top categories. Additional GO terms can be found in Table S1.

(B) Interactive network of significant GO terms. A total of 671 genes had rdSNVs in two or more individuals, as well as a 25% high percentile score for residual variation intolerance score (RVIS) and haploinsufficient score combined. These are the top genes that would likely possess inactivating or gain-of-function mutations. The biological GO enrichment of the 671 genes using DAVID (The Database for Annotation, Visualization and Integrated Discovery) was evaluated. The "Interactive graph" of the enriched "Biological Process" GO terms (Benjamini multiple test correction p value < 0.05) was visualized with REViGO software. The shade of gray of the fill in the circles indicates user-provided p value, with the darkest shade representing the highest p value. Highly similar GO terms are linked by edges in the graph, where the gray line width indicates the degree of similarity.

Jmjd1c is expressed in the pharyngeal apparatus and outflow tract region of the heart in mouse embryos at stage E10.5 (Figure S3). All rare non-synonymous variants in *JMJD1C* and *RREB1* were validated by Sanger sequencing (Figure S2).

To assess whether individual rdSNVs might affect protein function, we examined the type of the variant and its location with respect to the known protein domain structure of the gene product (Figure 3B). There were ten rare SNVs in *JMJD1C*, which affected nine CHD-positive case subjects and none of the control subjects (Figures 3B and 3C; Table S3). Thus, one specific SNV affected two CHD-positive case subjects (Figure 3B). Among the ten variants, the p.Ser1429Leu (c.4286C>T) variant has been identified in dbSNP (rs201627592) and NHLBI ESP6500 with a MAF of 0.0005 in 6,132 genomes. None of the other nine have been reported in these databases. The frequency of SNVs were also determined from the Exome Aggregation Consortium (ExAC), which is based on exome sequencing data of 60,706 unrelated individuals from various disease-specific and population genetic studies. The p.Ser1429Leu variant had a MAF of 0.0004 among 60,450 samples, whereas the p.His546Tyr (c.1636C>T) and p.His941Arg (c.2822A>G) variants have been reported in six and one person in ExAC, respectively. The JmjC domain is characteristic of the jumonji family of transcription factors and it might be involved in histone demethylation.[12] The p.His2466Tyr (c.7396C>T) variant is located within the JmjC domain and the p.Leu250fs (c.748_749delTT) variant could generate a truncated protein that would lack the JmjC domain (Figure 3C), suggesting that they might result in a loss-of-function mutation. The rest of the variants are not located within a known domain of the protein.

There were seven rare missense variants in *RREB1*, which affected eight CHD-affected case subjects and none of the control subjects (Figure 3B; Table S3). Thus, one variant was present in two different unrelated case subjects. Among the seven variants, p.Gly344Val (c.1031G>T), p.Ile614Met (c.1842C>G), and p.Pro1064Leu (c.3191C>T) have been identified in dbSNP (rs114551633, rs146678576, and rs143874633) and NHLBI ESP 6500 with a MAF < 0.0006 among 6,132 samples. None of the four other missense variants have been reported in dbSNP, 1000 Genomes Project, or NHLBI ESP 6500 Exome sequence databases. The p.Pro295Leu (c.884C>T) and p.Gly621Arg (c.1861G>C) variants have been reported in one and three case subjects, respectively, among 60,706 individual genomes. The p.Pro1064Leu variant was present in two CHD-affected case subjects. This variant is located in a proline-rich domain,[13] but none of the other variants are located within a known functional domain of the protein (Figure 3C).

SEC24C encodes a protein that is a coat component of COPII vesicles, which transport proteins from the endoplasmic reticulum to the Golgi[14–16] and is required for embryonic development.[17] There were four rdSNVs in *SEC24C*, which affected seven CHD-affected case subjects and none of the control subjects (Table S3). The p.Arg222Gln (c.665G>A) variant was present in four case subjects, and the others were each present in one individual. Among the four variants, three of them—p.Arg222Gln, p.Leu230Phe (c.688C>T), and p.Pro255Ser (c.763C>T)—have been identified in dbSNP (rs147121844, rs141953475, and rs150235476). The variants were identified in the NHLBI ESP 6500 database, with a MAF < 0.0005 among 6,132 samples. Using ExAC, the p.Arg222Gln, p.Leu230Phe, and p.Pro255Ser variants have been reported in 102, 15, and 2 individuals among 60,706 individual genomes, respectively. None of the rdSNVs were located in a known functional domain of the protein (Figure 3C).

Two of the three genes, *JMJD1C*[18,19] and *RREB1*,[20] are involved in histone modification. *TBX1*, the strongest

A

		Case with rSNV		Fisher P-Value	
Symbol	RefSeq	Coding rSNVs	rdSNVs	Coding rSNVs	rdSNVs
JMJD1C	NM_032776.2	9	7	0.001	0.005
RREB1	NM_001168344	8	7	0.002	0.005
SEC24C	NM_004922	7	7	0.005	0.005

B

Gene	HGVS Protein	rdSNVs	Affected Individual ID
RREB1	p.Arg51Gln	Yes	42461
RREB1	p.Pro295Leu	Yes	42321
RREB1	p.Gly344Val		42395
RREB1	p.Ile614Met	Yes	42491
RREB1	p.Gly621Arg	Yes	42336
RREB1	p.Pro1064Leu	Yes	42464
RREB1	p.Asp1229Gly	Yes	42411
JMJD1C	p.His2466Tyr	Yes	42442
JMJD1C	p.Ser1474Pro	Yes	42462
JMJD1C	p.Ser1429Leu		42417,42442
JMJD1C	p.Ile1117Leu		42347
JMJD1C	p.His941Arg	Yes	42466
JMJD1C	p.Pro653Ser	Yes	42321
JMJD1C	p.His546Tyr		42346
JMJD1C	p.Pro287Arg	Yes	42372
JMJD1C	p.Leu250fs	Yes	42313
JMJD1C	p.Asn190Asp	Yes	42347
SEC24C	p.Ala81Ser	Yes	42460
SEC24C	p.Arg222Gln	Yes	42317,42339,42436,42337
SEC24C	p.Leu230Phe	Yes	42318
SEC24C	p.Pro255Ser	Yes	42465

C

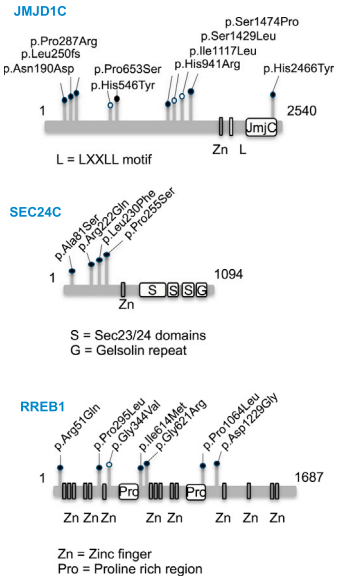

Figure 3. Identification of rdSNVs in *JMJD1C*, *RREB1*, and *SEC24C* in CHD-Affected Case Subject with 22q11DS

(A) Burden test results for *JMJD1C*, *RREB1*, and *SEC24C*. We performed a burden test for genes with at least one rdSNV in two or more individuals. Table lists the p value for Fisher exact and permutation tests for *JMJD1C*, *RREB1*, and *SEC24C*. If we included all rare non-synonymous SNVs (MAF < 0.001) in the burden test irrespective to whether they are predicted to be benign or damaging, then three additional individuals with variants in *JMJD1C* and one in *RREB1* were identified. *JMJD1C* has a stronger p value than any other gene ($p = 0.001$).

(B) Rare coding mutations identified in *JMJD1C*, *RREB1*, and *SEC24C*. Rare SNVs in *JMJD1C*, *RREB1*, and *SEC24C* are indicated with respect to the predicted amino acid change (HVGS Protein). In the rdSNV column, the variants predicted to be damaging are acknowledged as a "yes" while those that are predicted to be benign are left blank. Affected Individual ID is the coded ID of the subject containing the SNV. All 22q11DS-affected individuals have a CHD and are case subjects.

(C) Rare coding mutations with respect to the domain structure of *JMJD1C*, *RREB1*, and *SEC24C*. The domain structure of each protein is shown (RCSB, Protein Data Bank) with respect to the position of non-synonymous rdSNVs in case subjects (MAF ≤ 0.001) as filled lollipops, with additional predicted benign rSNVs, indicated as open lollipops. Abbreviations are as follows: L, LXXLL motif; S, Sec23/24 domains; G, Gelsolin repeat; Zn, zinc finger; Pro, proline-rich region. Detailed information about functional prediction and population frequency can be found in Table S3.

candidate for 22q11DS, encodes a DNA binding transcription factor that interacts with the chromatin machinery.[21,22] This implicates regulation of transcription as being an important consideration for affecting extremes of heart phenotypes in the background of the 22q11.2 deletion. Further, other genes that can influence the phenotype in individuals with 22q11DS include *CRKL* (v-crk avian sarcoma virus CT10 oncogene homolog-like [MIM: 602007]),[23–25] *DGCR8* (DiGeorge syndrome critical region gene 8 [MIM: 609030]),[26–28] and *HIC2* (hypomethylated in cancer 2 [MIM: 617712]).[29] It is possible that expression of these genes could be influenced by chromatin modification.

Of interest, in a previous WES study of non-syndromic or non-deleted syndromic CHD, of unknown diagnoses, of all types including conotruncal defects, de novo mutations were found in histone-modification pathway genes expressed in the heart.[30] Based on this evidence, we then examined other genes in the histone modification pathway (PW:0001338; Pathway Portal of the Rat Genome Database[31]) for rdSNVs. We tested whether rdSNVs could be identified in two or more case subjects or two or more control subjects (not in both case and control subjects). We also included the histone modification-related genes, which were identified in the WES study of non-syndromic CHD-affected subjects.[30]

Histones can be modified by several post-translational mechanisms resulting in either activation or suppression of gene expression.[32–36] There are three major mechanisms responsible for histone modifications: methylation, acetylation, and ubiquitination. There are well-characterized modification sites on specific histone proteins that alter chromatin structure in order to activate or repress transcription. These particular modifications form the epigenetic code, as shown in Figure 4A. We identified rdSNVs of interest in some of these genes, in either case or control subjects. One hypothesis is that genes with mutations in case subjects have the opposite function as genes with mutations in control subjects. Alternatively, genes in case versus control subjects might affect certain modifications, whether activating or repressing transcription, implicating gene dysregulation.

Histone methylation is a complex epigenomic process whereby both the position of the targeted amino acid as well as the degree of methylation (mono-, di-, or trimethylation) affects the overall structure of chromatin and ultimately gene transcription. The histone-modifying genes identified in subjects with non-syndromic CHD, such as *WDR5*, *MLL2* (also known as *KMT2D*), *CHD7*, *KDM5A*, *KDM5B*, and *SMAD2*, affect H3K4 and H3K27 methylation.[30] Methylation of lysine residues is catalyzed by histone methyltransferases, whereas the reverse reaction is catalyzed by histone demethylases (Figure 4A). We examined these genes in our cohort for rdSNVs. Among them, we identified rdSNVs in four genes affecting methylation (H3K4) in control and not case subjects: *MEN1*

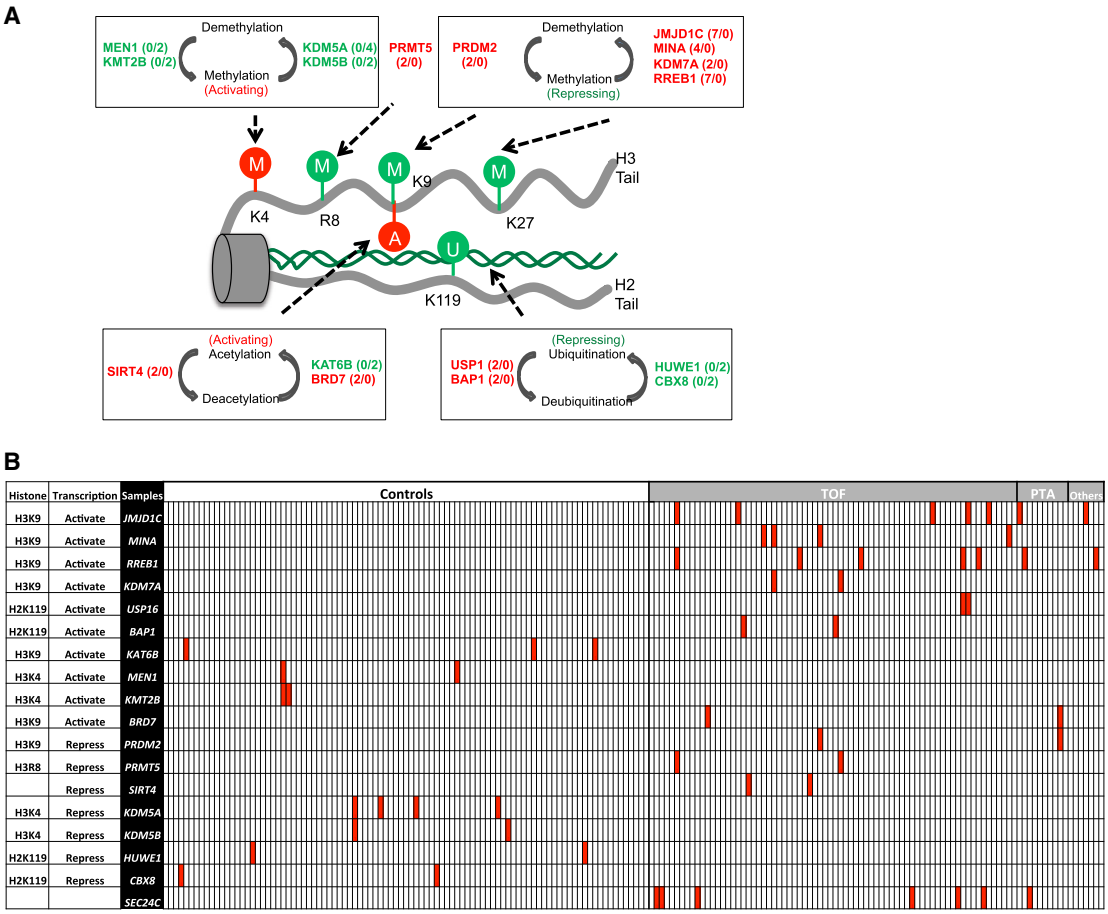

Figure 4. Rare Predicted Deleterious Variants

(A) Model of histone modification proteins with rdSNVs in respective genes in case versus control subjects. Histone marks are represented by circles in the histone tails (abbreviations are as follows: M, methylation; A, acetylation; U, ubiquitination). Both active histone marks (red-filled circles) and repressive marks (green-filled circles) are shown. Individual amino acids and their position with respect to H3 or H3 histone tails (gray) and double-stranded DNA (green) are indicated (K, lysine; R, arginine). The four boxes contain a specific histone-modification process, with methylation-demethylation shown in the boxes above the drawing of the histone tails. Acetylation-deacetylation as well as the ubiquitination-deubiquitination are shown in the two boxes below. Proteins are included that have rdSNVs in their respective genes in two or more case subjects only (red) or control subjects only (green). The number of individuals affected is indicated in the order of number of case subjects/number of control subjects.

(B) Profile of each rdSNV for each individual with 22q11DS. Representative histone marks (Histone column) and their role in transcription (Transcription column) are shown at left. The gene name with rdSNVs is listed in the third column (Samples). The rest of the columns represent rdSNVs (red box) from each individual with 22q11DS, with the control subjects shown first, on the left, followed by the case subjects, indicating the type of heart defect (see details in Figure 1).

(multiple endocrine neoplasia I [MIM: 613733]), *KMT2B* (lysine (K)-specific methyltransferase 2B [MIM: 606834]), *KDM5A* (lysine (K)-specific demethylase 5A [MIM: 180202]), and *KDM5B* (lysine (K)-specific demethylase 5B [MIM: 605393]).[37–43] Although we have not yet validated the rdSNVs in functional assays, we speculate that dysregulation of transcription via modification of histones could influence risk for developmental anomalies with the 22q11.2 deletion. Surprisingly, de novo mutations in *KDM5A* and *KDM5B* have been reported in non-syndromic CHD-positive individuals,[30] which affect different amino acids and the opposite of what we found, suggesting that these genes might play protective roles in those that have the 22q11.2 deletion.

Our most significant finding is that we identified 18 of 89 case subjects with rdSNVs in four genes whose encoded proteins demethylate histones, including H3K9 and possibly H3K27—*JMJD1C*, *MINA* (MYC induced nuclear antigen [MIM: 612049]), *KDM7A* (lysine (K)-specific demethylase 7A), and *RREB1*[44–49]—acting to activate transcription (Figure 4). The rdSNVs occurred in different subjects, implicating them as acting independently. We found that two subjects had rdSNVs in *PRDM2* (PR domain containing 2, with ZNF domain [MIM: 601196]), which has the opposite function[50,51] for the two modifications. We also found rdSNVs in *PRMT5* (protein arginine methyltransferase [MIM: 604045]) in two case subjects, possibly affecting methylation of arginine at position H3R8.[52]

Besides methylation, there are two other well-known types of histone modifications: acetylation and ubiquitination. Histone acetylation is, in general, associated with gene activation. We found four case and three control subjects

with rdSNVs that might affect acetylation, but in the opposite manner (Figure 4): *SIRT4* (Sirtuin 4 [MIM: 604482]),[53] *BRD7* (bromodomain containing 7),[54] and *KAT6B* (K(Lysine) acetyltransferase 6B [MIM: 605880]).[55] Monoubiquitination of nucleosome histone H2A at lysine 119 leads to gene repression or silencing, whereas deubiquitination leads to gene activation.[56–59] Four case subjects were identified with rdSNVs in two genes that catalyze deubiquitination (Figure 4A)—*USP16* (ubiquitin specific peptidase 16 [MIM: 604735])[60] and *BAP1* (BRCA1 associated protein-1 [MIM: 603089])[61]—and four control subjects had rdSNVs in two genes that promote ubiquitination—*HUWE1* (HECT, UBA, and WWE domain containing 1, E3 ubiquitin protein ligase [MIM: 300697])[62] and *CBX8* (chromobox homolog 8),[63]—suggesting that overall dysregulation could occur downstream of the deletion. Interestingly, genes involved in histone modification have been suggested to act as hub genes that could serve as modifiers of disease states or phenotypes in humans.[64] It is thus possible that histone modifications in context of the 22q11.2 deletion might have effects on embryonic development for which the heart is particularly vulnerable. We examined the expression level of the chromatin-modifying genes in existing gene expression profiling data from mouse or human embryos (Table S4). Expression level was compared to housekeeping genes versus genes differentially expressed in various tissues. Among them, *Jmjd1c* has greater than average gene expression in all the relevant tissues of biological interest (Table S4). However, these findings need to be tested with future expression and functional studies of progenitor cells for the heart during embryonic development.

We then performed the burden tests for the entire histone modification pathway. Rare mutations were not significantly enriched in either case or control subjects ($p = 0.82$, OR = 0.83 [0.34, 2.07]). However, histone modifications could activate (*JMJD1C*, *MINA*, *KDM7A*, *RREB1*, *USP16*, *BAP1*, *BRD7*, *MEN1*, and *KMT2B*) or repress (*KDM5A*, *KDM5B*, *PRDM2*, *PRMT5*, *SIRT4*, *HUWE1*, and *CBX8*) gene expression, making it necessary to consider biological functions for burden testing. After collapsing the rdSNVs in either group and then performing a gene burden test, we found that case subjects with rdSNVs are significantly enriched in genes within the group that activates gene transcription (Fisher exact $p = 0.0004$, permutations $p = 0.0003$, OR = 5.16 [1.99, 13.44]). However, controls are not significantly enriched in the group of genes that represses transcription (Fisher exact $p = 0.80$, permutations $p = 0.70$, OR = 0.82 [0.29, 2.29]). These results are shown graphically in Figure 4B, where we identified the variants in genes activating or repressing transcription depending upon which chromatin-modifying gene was likely to be involved. Each variant of interest from each individual case or control subject with 22q11DS is shown. Overall, these results suggest that extremes of heart phenotypes in the 22q11.2 deletion background might be modified by rdSNVs in genes for demethylation and deubiquitination of transcription repression histone marks such as H3K9 and H2Kun119, thereby generating hypotheses to be tested in cardiac progenitor cells or animal models in the future.

Mutations that are de novo events might not be tolerated and have a high likelihood that they could contribute to the etiology of disease. In a previous report,[30] de novo heterozygous mutations in histone-modifying genes were identified in association with risk for CHD. Here, we have a different hypothesis in that variants in genes identified here do not cause disease on their own, but they act as modifiers of phenotype, when combined with the 22q11.2 deletion. The heterozygous variants in histone-modification genes identified here, when combined with the 22q11.2 deletion, could affect embryonic development or heart formation in humans. Future functional studies would be required to determine their effects with or without loss of critical genes on 22q11.2 or in the presence of the deletion itself. The way in which they can affect function might be complicated to dissect, because histone modifications are inter-connected and they act in combination and/or sequentially to regulate transcription.[65–68] Further, chromatin modification works in concert with tissue-specific as well as general transcription factors, which ultimately regulate gene expression. Nonetheless, when taken together, genes important in modifying gene expression could influence developmental fields or cardiac progenitor cells that alter development of the heart and are worthy of further investigation.

Numerous statistical methods have been developed to test aggregate groups of rare variants for gene-based association to disease.[69–71] Association results depend on locus architecture, effect size, and functional variant filters as well as which statistical methods are used.[72] Here we focused on histone-modifier genes because the top genes identified by WES involved histone modification and, in addition, de novo mutations in this class of genes were identified in non-syndromic or undiagnosed syndromic CHD subjects.[30] It is possible that integration of such methods combined with a larger cohort of subjects will be required to fully dissect the genetic architecture of modifiers of 22q11DS.

Accession Numbers

Exome and amplicon sequencing data have been deposited in the dbGAP, hosted by the NCBI, under the accession number phs000987.

Supplemental Data

Supplemental Data include three figures and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2015.10.013.

Acknowledgments

We thank Dr. Jidong Shan, Ms. Yinghui Song, and Ms. Debbie Lewis in the Molecular Cytogenetics core lab at Einstein for

technical support. We thank Emma Mclaughlin and Eugene Park for their outstanding technical support. We also appreciate the services of the Genomics Core at Einstein for Sanger DNA sequencing analysis. WES services were provided through the RS&G Service by the Northwest Genomics Center at the University of Washington, Department of Genome Sciences, under U.S. Federal Government contract number HHSN268201100037C from the National Heart, Lung, and Blood Institute. This work was supported by grant P01 HD070454 from the NICHD at NIH.

Received: September 4, 2015
Accepted: October 23, 2015
Published: November 19, 2015

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
Burrows-Wheeler Aligner, http://bio-bwa.sourceforge.net/
DAVID, http://david.abcc.ncifcrf.gov/
dbGaP, http://www.ncbi.nlm.nih.gov/gap
ExAC Browser, http://exac.broadinstitute.org/
GATK Best Practices, https://www.broadinstitute.org/gatk/guide/best-practices
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://www.omim.org/
Picard, http://picard.sourceforge.net/
RCSB Protein Data Bank, http://www.rcsb.org/pdb/home/home.do
REViGO software, http://revigo.irb.hr/
SAMtools, http://samtools.sourceforge.net/

References

1. Merscher, S., Funke, B., Epstein, J.A., Heyer, J., Puech, A., Lu, M.M., Xavier, R.J., Demay, M.B., Russell, R.G., Factor, S., et al. (2001). TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. Cell *104*, 619–629.
2. Jerome, L.A., and Papaioannou, V.E. (2001). DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. Nat. Genet. *27*, 286–291.
3. Lindsay, E.A., Vitelli, F., Su, H., Morishima, M., Huynh, T., Pramparo, T., Jurecic, V., Ogunrinu, G., Sutherland, H.F., Scambler, P.J., et al. (2001). Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. Nature *410*, 97–101.
4. Guo, T., McDonald-McGinn, D., Blonska, A., Shanske, A., Bassett, A.S., Chow, E., Bowser, M., Sheridan, M., Beemer, F., Devriendt, K., et al.; International Chromosome 22q11.2 Consortium (2011). Genotype and cardiovascular phenotype correlations with TBX1 in 1,022 velo-cardio-facial/DiGeorge/22q11.2 deletion syndrome patients. Hum. Mutat. *32*, 1278–1289.
5. Mlynarski, E.E., Sheridan, M.B., Xie, M., Guo, T., Racedo, S.E., McDonald-McGinn, D.M., Gai, X., Chow, E.W., Vorstman, J., Swillen, A., et al.; International Chromosome 22q11.2 Consortium (2015). Copy-number variation of the glucose transporter gene SLC2A3 and congenital heart defects in the 22q11.2 deletion syndrome. Am. J. Hum. Genet. *96*, 753–764.
6. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310–315.
7. Liu, X., Jian, X., and Boerwinkle, E. (2013). dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum. Mutat. *34*, E2393–E2402.
8. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.
9. Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. *6*, e1001154.
10. Sherman, B.T., Huang, W., Tan, Q., Guo, Y., Bour, S., Liu, D., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics *8*, 426.
11. Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE *6*, e21800.
12. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811–816.
13. Williamson, M.P. (1994). The structure and function of proline-rich regions in proteins. Biochem. J. *297*, 249–260.
14. Wendeler, M.W., Paccaud, J.P., and Hauri, H.P. (2007). Role of Sec24 isoforms in selective export of membrane proteins from the endoplasmic reticulum. EMBO Rep. *8*, 258–264.
15. Siddiqi, S., Siddiqi, S.A., and Mansbach, C.M., 2nd. (2010). Sec24C is required for docking the prechylomicron transport vesicle with the Golgi. J. Lipid Res. *51*, 1093–1100.
16. Mancias, J.D., and Goldberg, J. (2008). Structural basis of cargo membrane protein discrimination by the human COPII coat machinery. EMBO J. *27*, 2918–2928.
17. Adams, E.J., Chen, X.W., O'Shea, K.S., and Ginsburg, D. (2014). Mammalian COPII coat component SEC24C is required for embryonic development in mice. J. Biol. Chem. *289*, 20858–20870.
18. Watanabe, S., Watanabe, K., Akimov, V., Bartkova, J., Blagoev, B., Lukas, J., and Bartek, J. (2013). JMJD1C demethylates MDC1 to regulate the RNF8 and BRCA1-mediated chromatin response to DNA breaks. Nat. Struct. Mol. Biol. *20*, 1425–1433.
19. Lu, J., and Matunis, M.J. (2013). A mediator methylation mystery: JMJD1C demethylates MDC1 to regulate DNA repair. Nat. Struct. Mol. Biol. *20*, 1346–1348.
20. Ray, S.K., Li, H.J., Metzger, E., Schüle, R., and Leiter, A.B. (2014). CtBP and associated LSD1 are required for transcriptional activation by NeuroD1 in gastrointestinal endocrine cells. Mol. Cell. Biol. *34*, 2308–2317.
21. Chen, L., Fulcoli, F.G., Ferrentino, R., Martucciello, S., Illingworth, E.A., and Baldini, A. (2012). Transcriptional control in cardiac progenitors: Tbx1 interacts with the BAF chromatin remodeling complex and regulates Wnt5a. PLoS Genet. *8*, e1002571.
22. Randall, V., McCue, K., Roberts, C., Kyriakopoulou, V., Beddow, S., Barrett, A.N., Vitelli, F., Prescott, K., Shaw-Smith, C., Devriendt, K., et al. (2009). Great vessel development requires biallelic expression of Chd7 and Tbx1 in pharyngeal ectoderm in mice. J. Clin. Invest. *119*, 3301–3310.

23. Guris, D.L., Fantes, J., Tara, D., Druker, B.J., and Imamoto, A. (2001). Mice lacking the homologue of the human 22q11.2 gene CRKL phenocopy neurocristopathies of DiGeorge syndrome. Nat. Genet. *27*, 293–298.
24. Moon, A.M., Guris, D.L., Seo, J.H., Li, L., Hammond, J., Talbot, A., and Imamoto, A. (2006). Crkl deficiency disrupts Fgf8 signaling in a mouse model of 22q11 deletion syndromes. Dev. Cell *10*, 71–80.
25. Racedo, S.E., McDonald-McGinn, D.M., Chung, J.H., Goldmuntz, E., Zackai, E., Emanuel, B.S., Zhou, B., Funke, B., and Morrow, B.E. (2015). Mouse and human CRKL is dosage sensitive for cardiac outflow tract formation. Am. J. Hum. Genet. *96*, 235–244.
26. Chapnik, E., Sasson, V., Blelloch, R., and Hornstein, E. (2012). Dgcr8 controls neural crest cells survival in cardiovascular development. Dev. Biol. *362*, 50–56.
27. Chen, Z., Wu, J., Yang, C., Fan, P., Balazs, L., Jiao, Y., Lu, M., Gu, W., Li, C., Pfeffer, L.M., et al. (2012). DiGeorge syndrome critical region 8 (DGCR8) protein-mediated microRNA biogenesis is essential for vascular smooth muscle cell development in mice. J. Biol. Chem. *287*, 19018–19028.
28. Shiohama, A., Sasaki, T., Noda, S., Minoshima, S., and Shimizu, N. (2003). Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region. Biochem. Biophys. Res. Commun. *304*, 184–190.
29. Dykes, I.M., van Bueren, K.L., Ashmore, R.J., Floss, T., Wurst, W., Szumska, D., Bhattacharya, S., and Scambler, P.J. (2014). HIC2 is a novel dosage-dependent regulator of cardiac development located within the distal 22q11 deletion syndrome region. Circ. Res. *115*, 23–31.
30. Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature *498*, 220–223.
31. Petri, V., Hayman, G.T., Tutaj, M., Smith, J.R., Laulederkind, S.J., Wang, S.J., Nigam, R., De Pons, J., Shimoyama, M., Dwinell, M.R., et al. (2014). Disease pathways at the Rat Genome Database Pathway Portal: genes in context-a network approach to understanding the molecular mechanisms of disease. Hum. Genomics *8*, 17.
32. Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707–719.
33. Klose, R.J., and Zhang, Y. (2007). Regulation of histone methylation by demethylimination and demethylation. Nat. Rev. Mol. Cell Biol. *8*, 307–318.
34. Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.
35. Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.
36. Rando, O.J., and Chang, H.Y. (2009). Genome-wide views of chromatin structure. Annu. Rev. Biochem. *78*, 245–271.
37. Xiang, Y., Zhu, Z., Han, G., Ye, X., Xu, B., Peng, Z., Ma, Y., Yu, Y., Lin, H., Chen, A.P., and Chen, C.D. (2007). JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. Proc. Natl. Acad. Sci. USA *104*, 19226–19231.
38. Christensen, J., Agger, K., Cloos, P.A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell *128*, 1063–1076.
39. Iwase, S., Lan, F., Bayliss, P., de la Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstine, J.R., Bonni, A., Roberts, T.M., and Shi, Y. (2007). The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell *128*, 1077–1088.
40. Klose, R.J., Yan, Q., Tothova, Z., Yamane, K., Erdjument-Bromage, H., Tempst, P., Gilliland, D.G., Zhang, Y., and Kaelin, W.G., Jr. (2007). The retinoblastoma binding protein RBP2 is an H3K4 demethylase. Cell *128*, 889–900.
41. Yamane, K., Tateishi, K., Klose, R.J., Fang, J., Fabrizio, L.A., Erdjument-Bromage, H., Taylor-Papadimitriou, J., Tempst, P., and Zhang, Y. (2007). PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. Mol. Cell *25*, 801–812.
42. Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y. (2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. Nature *447*, 601–605.
43. Lee, M.G., Norman, J., Shilatifard, A., and Shiekhattar, R. (2007). Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. Cell *128*, 877–887.
44. Tsuneoka, M., Koda, Y., Soejima, M., Teye, K., and Kimura, H. (2002). A novel myc target gene, mina53, that is involved in cell proliferation. J. Biol. Chem. *277*, 35450–35459.
45. Zhang, Y., Lu, Y., Yuan, B.Z., Castranova, V., Shi, X., Stauffer, J.L., Demers, L.M., and Chen, F. (2005). The Human mineral dust-induced gene, mdig, is a cell growth regulating gene associated with lung cancer. Oncogene *24*, 4873–4882.
46. Feng, W., Yonezawa, M., Ye, J., Jenuwein, T., and Grummt, I. (2010). PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. Nat. Struct. Mol. Biol. *17*, 445–450.
47. Qiu, J., Shi, G., Jia, Y., Li, J., Wu, M., Li, J., Dong, S., and Wong, J. (2010). The X-linked mental retardation gene PHF8 is a histone demethylase involved in neuronal differentiation. Cell Res. *20*, 908–918.
48. Zhu, Z., Wang, Y., Li, X., Wang, Y., Xu, L., Wang, X., Sun, T., Dong, X., Chen, L., Mao, H., et al. (2010). PHF8 is a histone H3K9me2 demethylase regulating rRNA synthesis. Cell Res. *20*, 794–801.
49. Fortschegger, K., de Graaf, P., Outchkourov, N.S., van Schaik, F.M., Timmers, H.T., and Shiekhattar, R. (2010). PHF8 targets histone methylation and RNA polymerase II to activate transcription. Mol. Cell. Biol. *30*, 3286–3298.
50. Derunes, C., Briknarová, K., Geng, L., Li, S., Gessner, C.R., Hewitt, K., Wu, S., Huang, S., Woods, V.I., Jr., and Ely, K.R. (2005). Characterization of the PR domain of RIZ1 histone methyltransferase. Biochem. Biophys. Res. Commun. *333*, 925–934.
51. Pastural, E., Takahashi, N., Dong, W.F., Bainbridge, M., Hull, A., Pearson, D., Huang, S., Lowsky, R., DeCoteau, J.F., and Geyer, C.R. (2007). RIZ1 repression is associated with insulin-like growth factor-1 signaling activation in chronic myeloid leukemia cell lines. Oncogene *26*, 1586–1594.
52. Stopa, N., Krebs, J.E., and Shechter, D. (2015). The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. Cell. Mol. Life Sci. *72*, 2041–2059.
53. North, B.J., and Verdin, E. (2004). Sirtuins: Sir2-related NAD-dependent protein deacetylases. Genome Biol. *5*, 224.

54. Peng, C., Zhou, J., Liu, H.Y., Zhou, M., Wang, L.L., Zhang, Q.H., Yang, Y.X., Xiong, W., Shen, S.R., Li, X.L., and Li, G.Y. (2006). The transcriptional regulation role of BRD7 by binding to acetylated histone through bromodomain. J. Cell. Biochem. *97*, 882–892.
55. Champagne, N., Bertos, N.R., Pelletier, N., Wang, A.H., Vezmar, M., Yang, Y., Heng, H.H., and Yang, X.J. (1999). Identification of a human histone acetyltransferase related to monocytic leukemia zinc finger protein. J. Biol. Chem. *274*, 28528–28536.
56. Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873–878.
57. de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., et al. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev. Cell *7*, 663–676.
58. Fang, J., Chen, T., Chadwick, B., Li, E., and Zhang, Y. (2004). Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. J. Biol. Chem. *279*, 52812–52815.
59. Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol. Cell *20*, 845–854.
60. Zhang, Z., Yang, H., and Wang, H. (2014). The histone H2A deubiquitinase USP16 interacts with HERC2 and fine-tunes cellular response to DNA damage. J. Biol. Chem. *289*, 32883–32894.
61. Scheuermann, J.C., de Ayala Alonso, A.G., Oktaba, K., Ly-Hartig, N., McGinty, R.K., Fraterman, S., Wilm, M., Muir, T.W., and Müller, J. (2010). Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature *465*, 243–247.
62. Liu, Z., Oughtred, R., and Wing, S.S. (2005). Characterization of E3Histone, a novel testis ubiquitin protein ligase which ubiquitinates histones. Mol. Cell. Biol. *25*, 2819–2831.
63. Malik, B., and Hemenway, C.S. (2013). CBX8, a component of the Polycomb PRC1 complex, modulates DOT1L-mediated gene expression through AF9/MLLT3. FEBS Lett. *587*, 3038–3044.
64. Lehner, B., Crombie, C., Tischler, J., Fortunato, A., and Fraser, A.G. (2006). Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. Nat. Genet. *38*, 896–903.
65. Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J. Biol. Chem. *277*, 28368–28371.
66. Sun, Z.W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104–108.
67. Lee, J.S., Shukla, A., Schneider, J., Swanson, S.K., Washburn, M.P., Florens, L., Bhaumik, S.R., and Shilatifard, A. (2007). Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell *131*, 1084–1096.
68. Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC Genomics *13*, 424.
69. Kiezun, A., Garimella, K., Do, R., Stitziel, N.O., Neale, B.M., McLaren, P.J., Gupta, N., Sklar, P., Sullivan, P.F., Moran, J.L., et al. (2012). Exome sequencing and the genetic basis of complex traits. Nat. Genet. *44*, 623–630.
70. Asimit, J., and Zeggini, E. (2010). Rare variant association analysis methods for complex traits. Annu. Rev. Genet. *44*, 293–308.
71. Stitziel, N.O., Kiezun, A., and Sunyaev, S. (2011). Computational and statistical approaches to analyzing variants identified by exome sequencing. Genome Biol. *12*, 227.
72. Moutsianas, L., Agarwala, V., Fuchsberger, C., Flannick, J., Rivas, M.A., Gaulton, K.J., Albers, P.K., McVean, G., Boehnke, M., Altshuler, D., and McCarthy, M.I.; GoT2D Consortium (2015). The power of gene-based rare variant methods to detect disease-associated variation and test hypotheses about complex disease. PLoS Genet. *11*, e1005165.