A PRDX1 mutant allele causes a MMACHC secondary epimutation in cblC patients
Jean-Louis Guéant, Céline Chery, Abderrahim Ouissalah, Javad Nadaf, David Coelho, Thomas Josse, Justine Flayac, Aurélie Robert, Isabelle Koscinski, Isabelle Gastin, et al.

To cite this version:
Jean-Louis Guéant, Céline Chery, Abderrahim Ouissalah, Javad Nadaf, David Coelho, et al.. A PRDX1 mutant allele causes a MMACHC secondary epimutation in cblC patients. Nature Communications, 2018, 9 (1), pp.67. 10.1038/s41467-017-02306-5. hal-01801484

HAL Id: hal-01801484
https://hal.univ-lorraine.fr/hal-01801484v1
Submitted on 6 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
A PRDX1 mutant allele causes a MMACHC secondary epimutation in cblC patients

Jean-Louis Guéant1, Céline Chéry1, Abderrahim Oussalah1, Javad Nadaf2, David Coelho1, Thomas Josse1, Justine Flayac1, Aurélie Robert1, Isabelle Koscinski1, Isabelle Gastin1, Pierre Filhine-Tresarrieu1, Mihaela Pupavac2, Alison Brebner2, David Watkins2, Tomi Pastinen2, Alexandre Montpetit2, Fadi Hariri2, David Tregouët3, Benjamin A Raby4, Wendy K. Chung5, Pierre-Emmanuel Morange6, D. Sean Froese7, Matthias R. Baumgartner7, Jean-François Benoist8, Can Ficicioglu9, Virginie Marchand10, Yuri Motorin10, Chrystèle Bonnemains1, François Feillet1, Jacek Majewski2 & David S. Rosenblatt2

To date, epimutations reported in man have been somatic and erased in germlines. Here, we identify a cause of the autosomal recessive cblC class of inborn errors of vitamin B12 metabolism that we name “epi-cblC”. The subjects are compound heterozygotes for a genetic mutation and for a promoter epimutation, detected in blood, fibroblasts, and sperm, at the MMACHC locus; 5-azacytidine restores the expression of MMACHC in fibroblasts. MMACHC is flanked by CCDC163P and PRDX1, which are in the opposite orientation. The epimutation is present in three generations and results from PRDX1 mutations that force antisense transcription of MMACHC thereby possibly generating a H3K36me3 mark. The silencing of PRDX1 transcription leads to partial hypomethylation of the epiallele and restores the expression of MMACHC. This example of epi-cblC demonstrates the need to search for compound epigenetic-genetic heterozygosity in patients with typical disease manifestation and genetic heterozygosity in disease-causing genes located in other gene trios.
Epigenetic diseases are caused by stable alterations of DNA methylation, posttranslational histone modification, and/or production of non-coding RNA. These phenotypic traits may be transmitted from parents to offspring. They are termed epimutations when they are directly involved in the underlying molecular mechanisms of the disease. Epimutations can be separated into two types, primary and secondary, the latter occurring secondary to a DNA mutation in a cis- or trans-acting factor. In most cases, epimutations are somatic and are likely to exist as mosaics with tissue-specific effects. To our knowledge, to date there is no evidence of epimutations maintained in germlines in any of these cases. For example, an epimutation reported in three generations with a familial cancer syndrome caused by epigenetic silencing of the MLH1 gene is erased in spermatozoa, but reinstalled in the somatic cells of the next generation.

Here, we report cases with a rare inborn error of metabolism produced by the inheritance of a gene mutation that disrupts the gene function of a flanking gene through epigenetic mechanisms in somatic and germ line cells. The cases were classified as belonging to the autosomal recessive cblC (cobalamin, Cbl) class of inborn errors of vitamin B12 metabolism, usually caused by homozygous or compound heterozygous mutations in the MMACHC gene. MMACHC encodes a protein with both chaperone and enzyme functions, and its inactivation disrupts the synthesis of two Cbl derivatives, methylcobalamin (MeCbl) and adenosylcobalamin (AdoCbl), which serve as cofactors for methionine synthase and methylmalonyl-CoA mutase, respectively. The cblC disorder presents with both severe neurologic and systemic metabolic abnormalities. Here, we report a new cause of the cblC class of inborn errors of vitamin B12 metabolism that we name “epi-cblC”. The epi-cblC cases are compound heterozygous for a genetic mutation and a secondary epimutation at the MMACHC locus. The secondary epimutation was triggered by variants in the neighboring 3’ antisense gene PRDX1 that produces an aberrant antisense transcript by skipping the polyadenylation site.

Results
Identification of a new cause of cblC named epi-cblC. We initially identified the MMACHC secondary epimutation in two infant females with paternal transmission, case CHU-12122 (died at 1 month of age, Caucasian ancestry) and case WG-3838 (died at 2 months of age, Japanese and Korean ancestry) and in a 59-year-old male (case WG-4152, Caucasian ancestry). The cblC diagnosis was based on metabolic and fibroblast complementation analyses. Patient clinical and metabolic data are detailed in the Methods section.

Sanger sequencing of MMACHC in DNA of CHU-12122 revealed a single heterozygous c.270_271insA; p.Arg91LysfsX14 mutation, which was also identified in her unaffected mother and maternal grandfather (Fig. 1a). The maternal grandfather and the father had serum concentrations of homocysteine of 33.6 and 17.0 μmol/L, respectively (reference levels <15 μmol/L). No sequence variant in MMACHC was identified in the father and no additional gene mutations were identified in the patient and her relatives using the TruSight NGS Illumina panel of the genes involved in metabolic diseases. We performed methylation analysis of the MMACHC gene by PCR amplification and cloning, followed by Sanger sequencing of bisulfite-treated DNA. This revealed a heterozygous epimutation consisting of 32 hypermethylated CpG sites in a CpG island encompassing the promoter and first exon of MMACHC (Chr1:45,965,587–45,966,049) (Fig. 1b, c). The epimutation was observed in fibroblast DNA from the patient, blood and sperm DNA from her father, and blood DNA from her paternal grandfather. We detected a c.-302G allele of the rs3748643 bearing the epimutation and a c.-302T allele bearing the non-methylated allele. The epimutation and c.-302G allele of rs3748643 were absent in the DNA from the mother and maternal grandmother (Fig. 1a, c). The MMACHC allele bearing the c.270_271insA mutation was the only one detected by RT-PCR analysis of patient fibroblast transcripts. Treating the fibroblasts with 5-azacytidine (5-AZA), an inhibitor of DNA methyltransferases (DNMTs), produced bi-allelic MMACHC expression (Fig. 1d). Taken together, these results showed an epimutation causing promoter hypermethylation and MMACHC silencing in the paternaly transmitted allele and a c.270_271insA mutation in the allele inherited from the mother. The epimutation was present in three generations; it was transmitted to CHU-12122 by her father, and to her father by her grandfather and was present in the sperm of CHU-12122’s father.

Sanger sequencing of MMACHC in blood DNA of WG-3838 revealed the presence of a previously reported heterozygous c.81G>A splice variant in the patient, her mother, and a deceased brother (Fig. 2a). RT-PCR of fibroblast transcripts found no detectable MMACHC mRNA. No paternal MMACHC mutation was identified. Sequencing of bisulfite-converted DNA identified the same mono-allelic epimutation as that identified in CHU-12122 in blood DNA samples from her and her father and in sperm from her father; the patient was compound heterozygous for the c.81G>A mutation and the MMACHC epimutation (Fig. 2b). In contrast to CHU-12122, the c.-302T allele of the rs3748643 c.-302 G>T polymorphism was detected with the epimutation and the c.-302G allele tagged the non-methylated allele. Taken together, these data showed a presentation similar to case CHU-12122. The epimutation silenced MMACHC gene expression in the fibroblasts.

Sequencing of the MMACHC gene in blood DNA of WG-4152 identified a previously described heterozygous c.158T>C (p.Leu53Pro) mutation. Sequencing of bisulfite-converted DNA identified the same mono-allelic epimutation found in CHU-12122 and WG-3838. The patient was compound heterozygous for the c.158T>C mutation and the MMACHC epimutation. The c.-302G allele of rs3748643 was detected in the allele bearing the epimutation and the c.-302T allele in the non-methylated allele.

Confirmation of the epimutation by epigenome-wide analysis. We performed an epigenome-wide study of DNA from the 3 epi-cblC cases and their relatives using the Infinium HumanMethylation450 BeadChip array (HM450K). The methylation variation positions (MVPs) of the three cases were compared to those observed in 690 DNA samples from a Caucasian population from the Asthma BRIDGE consortium (USA)31, 350 DNA samples from a Caucasian population recruited in the South of France (the MARTHA cohort)32, and one typical cblC (CHU-09011) sample with compound heterozygosity for the same mutation as CHU-12122 in exon 2 (c.270_271insA, p.Arg91 > LysfsX14) and an exon 4 mutation (c.616C>T, p.Arg206Trp). The HM450K array included eight probes covering the CpG sites of the epimutation (Fig. 1b). HM450K methylome profiling in the three epi-cblC cases and their relatives confirmed the results obtained by Sanger sequencing of bisulfite-treated DNA (Fig. 2c, Fig. 2d). The epimutation was also confirmed in sperm from the fathers of the CHU-12122 and WG-3838 cases. The HM450K hypomethylation in sperm DNA from a control population from Utah33 was similar to that observed in the genomic DNA of the control subjects (Fig. 2e). We observed only one subject in the control population with the same hypermethylation of the MMACHC sites as was reported in the epi-cblC cases and their relatives. The allele...
Fig. 1 The CHU-12122 epi-cblC case is compound heterozygous for a coding mutation and an epimutation of MMACHC detected in three generations. a Pedigree of the family of case CHU-12122; red bar, mutation (heterozygous c.270_271insA, p.Arg91LysfsTer14) found in the proband, her mother, and her maternal grandfather; green circle, epimutation encompassing the MMACHC promotor/exon 1 found in the proband. b Map of the MMACHC gene and expanded view of the MMACHC CpG island. CpG sites are numbered according to their position upstream and downstream of ATG. Asterisks indicate the CpG sites of the CpG island which were probed in the HM450K array. c Epigrams of MMACHC methylation analyzed by PCR amplification/cloning/Sanger sequencing of bisulfite-treated DNA. Epimutations were detected in the proband, her father, and her paternal grandfather. d RT-PCR of fibroblasts from case CHU-12122 before and after the 3-day treatment with 10 µM 5-azacytidine (5-AZA). Silencing of the wild-type allele (mono-allelic expression in upper panel, red arrow) was reversed after treatment (bi-allelic expression after treatment, lower panel, red arrow)
frequency of the epimutation was therefore not higher than $5 \times 10^{-4}$ in the control population. We also confirmed the absence of the MMACHC epimutation in the CHU-09011 cblC case.

To compare the epigenome-wide changes between the three cases and their relatives with those from the control population, we plotted the difference between the beta methylation coefficients of the MVPs in cases and relatives and those reported in the MARTHA control population for each CpG site probed in the HM450 array. The 'epi-Manhattan' plot generated by this analysis showed that the MMACHC epimutation was the single change greater than 0.3 that reached a high level of statistical significance (Fig. 3a). In the control population, the CpG island...
that encompassed the promoter and the first exon of MMACHC was hypomethylated\textsuperscript{14} and surrounded by hypermethylated CpG sites (Fig. 3b, c). This corresponded to an ‘epi-haplotype block’ composed of the eight CpG sites probed by the HM450K array (Fig. 3b, c). Expression of MMACHC is regulated by a transcription regulatory complex of proteins that involves HCFC1, THAP11, and ZNF143\textsuperscript{15}. Analysis of the HM450K methylome failed to find any aberrant hypermethylation related to HCFC1-ZNF143-THAP11 in our epi-cblC cases.

We performed ChIP-Seq analyses on the chromatin of case and control fibroblasts to further characterize the epigenetic changes at MMACHC and neighboring genes (Fig. 4a). We observed a significant accumulation of trimethylated lysine 36 on histone H3 (H3K36me3) in the MMACHC promoter region and the proximal part of the 5’ adjacent CCDC163P gene only in patients harboring the epimutation (Fig. 4a). In contrast, all control and patient samples show the absence of the H3K36me3 mark in the promoter region of PRDX1, the gene 3’ to MMACHC (which was associated with the absence of DNA hypermethylation).

The MMACHC epimutation occurs secondary to a PRDX1 mutation. We performed whole-genome sequencing (WGS) of DNA from cases CHU-12122 and WG-3838 and their relatives to determine whether the promoter hypermethylation of MMACHC was a secondary epimutation that resulted from a genetically heritable variant\textsuperscript{1, 2, 16}. The only sequence variant that segregated with the phenotype across all samples was in PRDX1 (Fig. 4b and Fig. 5a). PRDX1 is located adjacent to MMACHC and on the opposite strand. We found a c.515-1G>T mutation of PRDX1 in cases CHU-12122 and WG-4152 and a c.515-2A>T mutation of the same gene in WG-3838 (Fig. 4b). These PRDX1 mutations were also found in the relatives who carried the secondary epimutation, i.e., the father and paternal grandfather of CHU-12122 and the father of WG-3838 (Fig. 5a). The PRDX1 variants and the polymorphism rs3748643 associated with the secondary epimutation were present in the same allele as evidenced by DNA sequencing and transmission in the heterozygous relatives (Fig. 1c and Fig. 2b).

Presence of the secondary epimutation in additional cases. We replicated our results in five cases from two other centers. We found the secondary epimutation and the c.515-1G>T PRDX1 mutation in all five cases, and confirmed the maternal transmission of the PRDX1 mutation with the secondary epimutation in one case (Supplementary Table 1). The c.515-1G>T variant was present in seven individuals of European ancestry out of 66,000 individuals reported in the Exome Aggregation Consortium (ExAC) database (estimated frequency in the general population 1.052 × 10\textsuperscript{-4}).

The epimutation results from PRDX1 antisense transcription. MMACHC belongs to a trio of reverse (R1)–forward (F2)–reverse (R3) genes with dual partial overlaps of MMACHC by CCDC163P and PRDX1 at its 5’ and 3’ ends, respectively. The two mutations in PRDX1 affect the canonical splice acceptor site of intron 5, 1 and 2 bp apart, respectively. They are predicted to disrupt normal

---

**Fig. 3** Epigenome-wide analyses confirm the presence of the epimutation in genomic DNA and sperm. a Epi-Manhattan plot for the differences between β values in the epi-MMACHC subjects and controls (\(\beta_{\text{Epi-MMACHC}}-\beta_{\text{control}}\)). The horizontal line indicates a difference of 0.3; the volcano plot shows the relationship between (\(\beta_{\text{Epi-MMACHC}}-\beta_{\text{control}}\)) and \(-\log_{10}(P\text{-value})\) for the t-test. b Enlarged view of the MMACHC locus: β values of the CpG probes in the MMACHC locus and the epi-linkage disequilibrium plot reporting the matrix relationship between the CpG probes based on the blood samples. The dotted line corresponds to a β value threshold of 0.2, below which the CpG probe was considered to be fully unmethylated. c Genomic position of the MMACHC locus according to RefSeq Genes 105v2, NCBI (GRCh37)
splicing and to produce skipping of exon 6 and the polyA transcription termination signal of PRDX1. We confirmed this prediction by single-strand RNA-Seq and RT-PCR analyses of case fibroblasts, and found that the loss of transcription termination of PRDX1 produced an aberrant extension of antisense transcription through the MMACHC and CCDC163P genes. The level of antisense transcripts was high in case fibroblasts and undetectable in control fibroblasts. The activation of cryptic acceptor sites and the resulting skipping of exons and introns of MMACHC generated several transcripts identified in both RNA-Seq and RT-PCR. Among them, we identified a predominant 1.0 kb transcript that encompassed the exons 4 and 5 of PRDX1, the CpG island and the bidirectional promoter, and the first exon of CCDC163P (Fig. 5b, Fig. 6a, b, c). This transcript resulted from the activation of a cryptic antisense splicing site located in the middle of MMACHC exon 1 (Fig. 6d, e).

We further demonstrated that the PRDX1 mutation produced the silencing of MMACHC expression through the methylation of its promoter. To address this issue, we silenced the expression of PRDX1 with siRNA transfection in WG-3838 fibroblasts and the cell line MeWo-LC1. We used WG-3838 fibroblasts because of the complete absence of detectable sense transcription of MMACHC exons. We showed previously that the MeWo-LC1 melanoma-derived cell line had a cblC phenotype and no
transcription of MMACHC. We found recently that this cell line had the same antisense transcription of PRDX1 found in patient fibroblasts. We observed that the silencing of PRDX1 restored the expression of MMACHC in both cell lines and that it produced 10–15% hypomethylation of the allele initially fully methylated in the clones generated from bisulfite-treated DNA. No effect was observed in control fibroblasts transfected with the PRDX1 siRNA and in WG-3838 fibroblasts and MeWo-LC1 cells transfected with a control siRNA (Fig. 6f, g).

**Discussion**

Our epi-cblC cases represent a cause of the autosomal recessive cblC disorder, resulting from a secondary epimutation in the bidirectional promoter of MMACHC on one allele and a coding mutation on the other. The secondary epimutation is caused by an inherited mutation of the adjacent PRDX1 gene. It is directly involved in the mechanism of the disease through the forced antisense transcription of the adjacent mutated PRDX1 and is present in three generations. Previous reports have described DNA methylation at genes that mimics recessive mutations by creating a transcriptional haploinsufficiency with tissue-specific epigenetic silencing. In contrast, the secondary epimutation seen in our cases is present in the DNA from blood cells, fibroblasts, and sperm. The high expression of PRDX1 in spermatic cells may explain the presence of the epimutation in sperm, in contrast to the spermatozoa erasure previously observed in families with epigenetic silencing of the MLH1 gene.

The secondary epimutation of our epi-cblC cases is located in a trio of genes (reverse CCDC163P–forward MMACHC–reverse PRDX1) (Fig. 6d, e), in a bidirectional promoter in the head-to-
head sense–antisense gene pair (SAGP) constituted by CCDC163P and MMACHC. The forced antisense transcription of MMACHC resulted from the skipping of the last exon of PRDX1. This skipping was the causative defect that produced the epimutation since the silencing of PRDX1 decreased the methylation of exon 1 of MMACHC and the promoter, and restored the transcription of MMACHC, in WG-3838 and MeWo-LC1 cells. However, the silencing of PRDX1 produced only 10–15% hypomethylation of the allele initially fully methylated, suggesting a limited reversibility of the epimutation (Fig. 6g). Antisense

Fig. 6 The PRDX1 mutation causes aberrant antisense transcription through the MMACHC and CCDC163P genes. a qRT-PCR analyses of antisense transcripts encompassing exons 5 and 6 of PRDX1, PRDX1 exon 5—MMACHC exon 1 and MMACHC exon 1—CCDC163P exon 1 showing the aberrant extension of PRDX1 antisense transcription through sense MMACHC and antisense CCDC163P and no aberrant transcripts in HDF control fibroblasts. b RT-PCR detection of an aberrant 1 kb transcript produced with PRDX1 sense and CCDC163P antisense primers. Lanes 2, 4, and 6 correspond to RT-PCR of RNA from fibroblasts CHU-12122, WG-3838, and HDF (control fibroblasts). Lanes 3 and 5 correspond to control experiments without fibroblast RNA in the reaction mixture. They show no amplification artifacts. c, d Sanger sequencing of cDNA shows that the antisense aberrant transcript encompasses PRDX1 exons 4 (red) and 5 (blue), part of exon 1 and promoter of MMACHC, and part of exon 1 of CCDC163P (green). e Proposed mechanisms of epigenetic silencing of the non-mutated F2 allele in patients with a heterozygous mutation of a causal gene that belong to a trio of reverse (R1)–forward (F2)–reverse (R3) genes. f Expression of PRDX1 and MMACHC in control fibroblasts, fibroblasts of patient WG3838, and MeWo-L1 cell line transfected with control and PRDX1 siRNA. g MMACHC methylation epigrams of MeWo-LC1 cells transfected with control and PRDX1 siRNA.
transcription could lead to RNA polymerase collision and the formation of triplexes in the bidirectional promoter. The formation of triplexes can generate CpG methylation, as shown for MHC. Recent data also strongly support a role of antisense transcription in the regulation of H3K36me3. The high expression of R1 in germ cells from sperm, fibroblasts, and blood may be explained by the ubiquitously high expression of PRDX1 in germ cells, stem cells, and somatic cells. The high expression of PRDX1 in germ cells may explain why the secondary epimutation escaped eraise in spermatozoa, in contrast to previous reports of other diseases. The high expression of PRDX1 could also maintain this epimutation during early embryonic development. The transcription of PRDX1 has been studied in bovine embryos from the two-cell to the blastocyst stage. PRDX1 transcripts were detected throughout development, including in oocytes before and after maturation and in 2c, 5–8c, 9–16c embryos, morulae, and blastocysts. PRDX1 was also ubiquitously expressed in E7–E10 mouse embryos and adult humans. The methylation of the promoter in sperm is also consistent with previous results in a mouse model, in which genes with high H3K36me3 levels showed elevated gene-body DNA methylation in sorted germ cells of pooled E13.5 testes.

The cases of epi-CblC provide the first example of a rare disease produced by compound epigenetic/genetic heterozygosity in a reverse1 (R1)–forward (F2)–reverse (R3) trio of genes (Fig. 6d, e). This suggests consideration of this mechanism in cases of recessive diseases bearing a single heterozygous variant in the causal gene in the presence of typical disease manifestations. The number of R1–F2–R3 trios of genes with a configuration similar to the CCDC163P–MMACHC–PRDX1 trio is limited in the human genome. We have listed examples of such trios with a recessive transmission of inherited diseases produced by F2 gene mutations (Supplementary Table 2). In most of them, the CpG island is hypomethylated and the RNA-Seq analysis of control fibroblasts shows the absence of aberrant transcription of the adjacent genes. Our data suggest that the presence of an epimutation in patients with a single heterozygous mutation and a severe phenotype for any of these diseases should be systematically investigated. In contrast to coding mutations, pharmacological inhibitors of DNMTs such as 5-AZA can reverse epigenetic changes.

Consistently, treatment of fibroblasts from one of our epi-CblC cases with 5-AZA restored bi-allelic MMACHC expression with detection of the wild-type allele (Fig. 2a). This suggests that further evaluation of the use of this compound in the treatment of potentially fatal, severe decompensation in these ‘epi-diseases’, as observed in our two pediatric cases, is warranted.

In conclusion, we report cases of a rare metabolic disease caused by compound heterozygosity of a secondary epimutation detected in somatic cells and sperm in one allele and a genetic mutation in the other. These cases represent an entity that we named “epi-CblC“. This research reveals perspectives on the diagnosis and treatment of inherited metabolic diseases in general and “epi-CblC” patients in particular. It suggests to search for epimutations in rare diseases with typical severe presentation despite heterozygous gene mutation.

Methods

Patients. Informed written, signed consent for performing the analyses was obtained from all subjects and from the parents, when appropriate. We followed the rules of the French National Reference Center of Inborn Metabolism Diseases of Nancy that were reviewed and approved by The Haute Autorité de Sénté (HAS, Sain Denis, France, https://www.has-sante.fr/) for the study of the patients and their relatives. Case 1 (CHU-12122) was the first child of an unrelated French Caucasian couple, with a healthy paternal half-sibling and a healthy maternal half-sibling. A history of Sudden Infant Death Syndrome was reported for the father’s sister. The mother had three episodes of spontaneous abortion of unknown origin. The proband was born at term with a low birth weight of 2490 g. After 1 month of breastfeeding, she had marked mucocutaneous pallor, tachycardia, and peripheral hypotonia. Further investigations during hospitalization showed megaloblastic anemia with 3.5 g/dL hemoglobin, discrete concentric myocardial hypertrophy, pulmonary hypertension, and hepatomegaly. Serum vitamin B12 and folate levels were within the routine ranges. Impaired activities of methionine synthase and methylenetetrahydrofolate reductase were reflected by decreased methionine at 5 μmol/L (0.3–3.9) and increased homocysteine and methylenetetrahydrofolate acid at 174 μmol/L (reference range <0.15 μmol/L) and 325 μmol/L (reference range <0.35 μmol/L), respectively. She was diagnosed with CblC by fibroblast complementation analysis. The patient died at 1 month of age after multiple organ failure with hypotension, desaturation, and severe bradycardia.

Patient WG-3838 was born after an uneventful full-term pregnancy and weighed 3400 g. Her older brother died at 1 month of age from cardiac arrest, and a cardiomyopathy was noted upon autopsy. At week 5, the patient was taken to the emergency room. Serum vitamin B12 was within the reference range, serum methionine was low, and homocysteine and MMA were increased to 110 and 20 μmol/L respectively. The patient was hypertensive and developed tachycardia. She had a sudden cardiac arrest and developed acute renal failure. She had a second cardiac arrest and died at age 2 months. She was diagnosed as a CblC case by fibroblast complementation analysis.

Control populations. The Asthma BioRepository for Integrative Genomic Exploration consortium (Asthma BRIDGE consortium, USA) is an open-access collection of cDNA and DNA from more than 1450 well-characterized subjects participating in ongoing genetic studies of asthma and an accompanying database of phenotype, genome-wide SNP genotype, expression, and methylation data. The Caucasian population of the MARThA MARSeille (FRombosis Association) cohort was recruited in the Thrombophila Centre of La Timone University Hospital (Marseille, France) between January 1994 and October 2002. The MH450K profiling data of the sperm DNA from the control population in Utah are publicly available at the Gene Expression Omnibus under GSE64096, Utah cohort.

Studies in skin fibroblasts. We measured the cellular incorporation of the label from [1-14C]propionate and 5-14C]methylTHF into cellular macromolecules that were metabolized in 5% trichloroacetic acid at 5°C to evaluate the metabolic functions of methylmalonyl-CoA mutase (MUT) and methionine synthase (MS), respectively. If incorporation was significantly reduced, complementation analysis was performed. Patient cells were fused with fibroblasts from patients with different inborn errors of cobalamin metabolism, and the incorporation of labeled [1-14C]propionate and 5-14C]methylTHF was measured in parallel in fused and unfused cells. Cell membrane incorporation was complemented by fibroblasts from all classes of inborn errors except for the class to which the patient belonged. These tests were used to assign
patients to known complementation classes. Treatment with 5-aza-2'-deoxycytidine (5-aza-dC) was performed in fibroblasts from case CHU-12122. Fibroblasts were seeded at 3 × 10^5 cells/100-mm dish 24 h prior to treatment in DMEM supplemented with 10% fetal calf serum at 37°C and 5% CO₂. Cells were exposed for 72 h to 10 μM 5-aza-dC (Sigma-Aldrich Chimie S.A.R.L, Saint-Quentin, France). The culture media was replaced every 24 h with fresh media containing 5-aza-dC. At day 3, TRIzol® reagent (Thermo Scientific™, Fontenay-sous-Bois, France) was used to extract RNA from the cultured fibroblasts. RT-PCR was performed with RNA from treated and non-treated cells, using the primers listed in the Supplementary Table 3. First-strand cDNA was synthesized at 2 μg of total RNA using the SuperScript First-Strand Synthesis System (Invitrogen). PCRs were performed with Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Fontenay-sous-Bois, France) under standard conditions. The primers used for MMACHC_cDNA synthesis were: forward, 5'-AGCGTTGAGAAGCACATCCA-3' and reverse, 5'-CATAATACGAGGTCAGCACA-3'. The RT-PCR was sequenced by direct sequencing using the BigDye Terminator v.3.1 kit (Thermo Scientific™, Fontenay-sous-Bois, France) and the same primers.

For RNA-Seq and RT-qPCR study of aberrant transcription, RNA depletion of rRNA transcripts was extracted from case and control fibroblasts using RiboMinus™ (Thermo Scientific, Villebon, France). cDNA libraries were prepared using the TrueSeqTM RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). RNA sequencing was performed using Illumina HiSeq 2000, HiSeq 3000, and HiSeq X Ten Illumina System sequencers with paired-end 150-bp read length and the mean average of 35 billion reads per sample. Sequence reads were aligned as explained for the whole exome analysis. The Burrows-Wheeler Aligner (BWA-MEM, version 0.7.7) was used to map all reads to UCSC hg19. PCR duplicates were removed from the alignments using Picard version 1.96 (http://picard.sourceforge.net). Indels were realigned using the Genome Analysis Toolkit (GATK). SNVs and short indels were called using GATK haplotyper caller (version 3.4–46). All variants were annotated with ANNOVAR (http:// Annovar.openbioinformatics.org) and filtered on MAF > 0.001 and in any of the aforementioned databases were removed from further analysis.

Whole-genome ChiP-Seq. ChiP-Seq was performed in patient’s fibroblasts and control fibroblasts lines without inherited metabolic defects in the cobalamin metabolism. MMACHC fibroblasts were obtained from a 5-year-old female, and WG-4301 fibroblasts were from a 3-month-old female, and WG-3679 fibroblasts were from a 5-year-old female with methylmalonic aciduria (methylmalonic aciduria). For crosslinking, cells were directly fixed with formaldehyde (final concentration of 1%) for 10 min with gentle shaking at room temperature. The fixed cells were then quenched with glycine at a final concentration of 138 mM for 5 min at room temperature. The crosslinked cells were subsequently washed with 10 mL ice-cold PBS and scrapped. The cells were pelleted by centrifugation at 5000g for 5 min at 4°C. For cell and nuclear lysis, the crosslinked cell pellets were resuspended in ice-cold Cell-Lysis buffer (5 mM Pipes pH 8.5, 85 mM KC1, 1% Igepal CA-630, 50 mM NaF, 1 mM PMSE, 1 mM phenylarsine oxide, 5 mM sodium orthovanadate, 1 mM EDTA free protease inhibitors) at 50 μL/10 million cells and incubated for 15 min on ice. The nuclei were pelleted by centrifugation at 5000g for 10 min at 4°C. Nuclei were lysed in Nuclear Lysis buffer (50 mM tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, 50 mM NaF, 1 mM PMSE, 1 mM phenylarsine oxide, 5 mM sodium orthovanadate, 1 mM EDTA free protease inhibitors) at 50 μL for every excess of 10 million additional cells. The nuclei were incubated for 30 min on ice and then transferred to 1.5 mL TPX tubes (Diagenode) at 100–250 μL aliquots. The lysed nuclei were sonicated with a UCD-300 Diagenode Bioruptor on high power at 4°C with repeated cycles. The chromatin size distribution was evaluated on 2% pre-cast agarose gel. The chromatin size distribution >500 bp were further sonicated. For chromatin concentration and pre-clearing, the material was first spun down at 12,000g for 10 min at 4°C. The resulting pellets were washed with 10 mL tris-HCl pH 8.0 and pooled with collected supernatant for each sample. The chromatin material was then resuspended to a desired concentration equivalent to 5–10 μg/100 μL and <0.1% SDS using Nonase 10k omega centrifugal devices (PALL life sciences). The chromatin was then pre-cleared with 50 μL Protein A Dynabeads for 1 h at 4°C with rotation. The Immuno precipitation and reverse crosslinking were performed with the IP-Star SX-8G (Diagenode) according to the manufacturer’s instructions. Input material was prepared by following the reverse crosslinking procedure above, but without immunoprecipitation. The eluted material was collected by placing the samples on a magnetic rack. DNA was extracted using Qiagen’s MinElute PCR purification columns. ChiP-Seq libraries were prepared using the Lambda Library Preparation Kit Illumina Platforms (Kappa Biosystems) and Illumina TruSeq adapters (Illumina) according to the manufacturer’s instructions. DNA clean-up and size selection was performed using Ampure XP Beads. Samples were then sequenced using the HiSeq 2000 with selected read length of 50 bp. Sequence reads were aligned as explained for the WGS bioinformatic analysis. Duplicate reads were removed using PICARD. Peaks were called using Model-based analysis of ChiP-Seq data (MACS) as control and using the broad peak mode (–broad –broad-cutoff 0.1)25. In order to generate Wiggle (WIG) and Tiled Data Format (TDF) files, HOMER (4.7)26 and igvtools (2.3.6; http://software.broadinstitute.org/software/igv/igvtools) were used. Tags were normalized to 10 million reads to generate tracks for visualization.

Effect of PRDX1 silencing on MMACHC methylation. The human melanoma cell line MeWo-LC1 derived from MeWo cells (ATCC® HTB-65®). It was obtained from Dr. Robert Lipeto (University of Ottawa, Ottawa, Ontario, Canada). The human melanoma cell line MeWo-LC1 variant of MeWo cells, fibroblasts from case WG-3838, and fibroblasts HDF for control were cultured in DMEM medium supplemented with 4.5 g/L glucose, 10% v/v heat-inactivated fetal bovine serum, 1% v/v pyruvate in a humidified atmosphere with 5% CO₂ at 37°C. The cell lines were tested negative for mycoplasma contamination. Input material was prepared by sonication to two well plates at a concentration of 0.5 × 10⁶ cells/well the day before transfection; 10 nM of siRNA against PRDX1 gene (GGUCUCAUACCUCAAGAAG, GGCGCA-GAAGAAUUAAGAA, GAUGUGCUAGUAAAGAUA, and CUU-CACAGCCACACGUGU) was transfected with Lipofectamine RNAiMAX transfection reagent (Invitrogen) into MeWo-LC1 for 72 h at 37°C. Non-silencing siRNA (AllStars Negative Controls—Qiagen) was used as a negative control. After a 72-h period of incubation, the cells were collected and used for mRNA and DNA isolation. Two-hundred nanograms of RNA was reverse-transcribed with SuperScript First strand synthesis system (Invitrogen). Primers for the MMACHC gene (forward 5′-ACCTGGCTGCTTGTAAGGAGA-3′, reverse 5′-CTTACTCATCCTGTGGTGG-3′, PRDX1 forward 5′-CGCGGAGTATCGTCTTCTA-3′, reverse 5′-AGGTTGATTTACCCATGCTAGAT-3′). Quantification was

artificially-depleted with 4.5 g/L glucose, 10% v/v heat-inactivated fetal bovine serum, 1% v/v pyruvate in a humidified atmosphere with 5% CO₂ at 37°C. The cell lines were tested negative for mycoplasma contamination. Input material was prepared by sonication to two well plates at a concentration of 0.5 × 10⁶ cells/well the day before transfection; 10 nM of siRNA against PRDX1 gene (GGUCUCAUACCUCAAGAAG, GGCGCA-GAAGAAUUAAGAA, GAUGUGCUAGUAAAGAUA, and CUU-CACAGCCACACGUGU) was transfected with Lipofectamine RNAiMAX transfection reagent (Invitrogen) into MeWo-LC1 for 72 h at 37°C. Non-silencing siRNA (AllStars Negative Controls—Qiagen) was used as a negative control. After a 72-h period of incubation, the cells were collected and used for mRNA and DNA isolation. Two-hundred nanograms of RNA was reverse-transcribed with SuperScript First strand synthesis system (Invitrogen). Primers for the MMACHC gene (forward 5′-ACCTGGCTGCTTGTAAGGAGA-3′, reverse 5′-CTTACTCATCCTGTGGTGG-3′, PRDX1 forward 5′-CGCGGAGTATCGTCTTCTA-3′, reverse 5′-AGGTTGATTTACCCATGCTAGAT-3′). Quantification was

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-017-02306-5 | www.nature.com/naturecommunications
performed using the housekeeping gene RNA polymerase II as an internal standard with the following primers: forward 5'-CACAGCCTAAAGTGTA-3' and reverse 5'-GGTAGAAGATGGGAGA-3'. The temperature cycling for MMACHC and PRDX1 was 30 s at 95°C followed by 40 cycles consisting of 95°C for 5 s and 61°C for 30 s. The PCR program for RNA polymerase II was 30 s at 95°C followed by 40 cycles consisting of 95°C for 5 s and 60°C for 30 s. Results were expressed as arbitrary units by calculating the ratio of crossing points of amplification curves of MMACHC or PRDX1 and the internal standard by using the 2^(-ΔΔCt) method. The bisulfite-treated DNA from MeWo-LC1 was used for methylation analysis of the MMACHC promoter by PCR amplification, cloning, and Sanger sequencing.

List of sense-antisense trio pairs of genes. We established a list of tail-to-tail SAGPs using a systematic search through the United Sense-Antisense Gene Pairs Database and retrieved 385 gene pairs. Each gene pair was visually inspected (RefSeq Genes) to confirm expression of both members of the gene pair and retrieved 385 gene pairs. Among the 385 manually inspected gene pairs, 99 had a triplet gene configuration according to a reverse(R1)-forward(F2)-reverse(R3) pattern (n=2) or a forward(F1)-reverse(R2)-forward(F1) pattern (n=17), with a CpG island in the promoter of R1/F2 and R2/F1, respectively.

Statistical analyses. The comparison of CpG beta values between the cases and controls was carried out using a t-test, and multiple testing corrections were performed using the Bonferroni adjustment. To assess the methylation architecture (epi-haptoype blocks, epi-LD analysis), linkage disequilibrium (LD) pairwise analysis was performed on all adjacent CpG pairs in a chromosome or within a haplotype block using a matrix output for both the expectation–maximization (EM) algorithm and the composite haplotype method (CHM). R^2 values were used in the epi-LD plots. All epi-LD analyses were performed after transforming the CpG beta values to categorical variables according to the ENCODE project (9). Any beta value ≥0.6 was considered fully methylated (MM). Any beta value ≤0.2 was considered to correspond to a fully unmethylated DNA. Beta values strictly greater than 0.2 and strictly less than 0.6 were considered to correspond to a partially methylated DNA (one of the two alleles). To assess the population stratification according to their whole methylene profile, numeric principal component analysis was performed using normalized beta values of each CpG probe across the Infinium HumanMethylation450 BeadChip array. The top 10 principal components (eigenvectors) were calculated with their respective eigenvalue (EV). Population stratification was assessed both in 2-dimensional and 3-dimensional visualization based on the eigenvalues of the top two and three principal components, respectively. All analyses were performed using the SNP & Variation Suite (SVS) 8.4.2 (Golden Helix, Inc. Bozeman, MT, USA).

Data availability. The RNA-Seq data that support the findings of this study have been deposited in the European Genome-phenome Archive (https://www.ebi.ac.uk/ega/) with the accession code EGAD00001000370. The other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Received: 5 May 2017 Accepted: 17 November 2017
Published online: 04 January 2018

References
1. Oey, H. & Whitelaw, E. On the meaning of the word ‘epimutation’. Trends Genet. 30, 519–520 (2014).
2. Horstemke, B. Epimutations in human disease. Curr. Top. Microbiol. Immunol. 310, 45–59 (2006).
3. Marti, C. D., Cropley, J. E. & Suter, C. M. Epigenetics in disease: leader or follower? Epigenetics 6, 843–848 (2011).
4. Heard, E. & Martienssen, R. A. Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157, 95–109 (2014).
5. Hinchliffe, S. P. Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 3′UTR. Cancer Cell 20, 209–213 (2011).
6. Watkins, D. & Rosenblatt, D. S. Inborn errors of cobalamin absorption and metabolism. Am. J. Med. Genet. C Semin. Med. Genet. 157C, 33–44 (2011).
7. Lerner-Ellis, J. P. et al. Identification of the gene responsible for methylmalonic aciduria and homocystinuria, cblc type. Nat. Genet. 38, 93–100 (2006).
8. Lerner-Ellis, J. P. et al. Spectrum of mutations in MMACHC, allelic expression, and evidence for genotype-phenotype correlations. Hum. Mutat. 30, 1072–1081 (2009).
9. Giaiczi, R. et al. Long-term visual outcome of methylmalonic aciduria and homocystinuria, cobalamin C type. Ophthalmology 112, 381–386 (2014).
10. Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–72 (2012).
11. McGeachie, M. J. et al. The metabolomics of asthma control: a promising link between genetics and disease. Immun. Inflamm. Dis. 3, 224–238 (2015).
12. Dick, K. J. et al. DNA methylation and body-mass index: a genome-wide analysis. Lancet 383, 1990–1998 (2014).
13. Jenkins, T. G. et al. Intra-sample heterogeneity of sperm DNA methylation. Mol. Hum. Reprod. 21, 313–319 (2015).
14. Lee, H. J., Hore, T. A. & Reik, W. Reprogramming the methylome: erasing memory and creating diversity. Cell Stem Cell 14, 710–719 (2014).
15. Michaud, J. et al. HCFC1 is a common component of active human CpG-island promoters and coincides with ZNF143, THAP11, YY1, and GABP transcription factor occupancy. Genome Res. 23, 907–916 (2013).
16. Siklenka, K. et al. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. Science 350, aab2006 (2015).
17. Loewy, A. D. et al. Epigenetic modification of the gene for the vitamin B(12) chaperone MMACHC can result in increased tumorigenicity and methionine dependency. Mol. Genet. Metab. 96, 261–267 (2009).
18. Zhou, H. et al. Epigenetic allele silencing unveils recessive RYR1 mutations in core myopathies. Am. J. Hum. Genet. 79, 859–868 (2006).
19. Plechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. Nat. Rev. Genet. 14, 880–893 (2013).
20. O’Leary, V. B. et al. PARTICLE, a triplex-forming long ncRNA, regulates locus-specific methylation in response to low-dose irradiation. Cell Rep. 11, 474–485 (2015).
21. Loos, F., Loda, A., van Wijk, L., Grootegoed, J. A. & Grimbau, J. Chromatin-mediated reversible silencing of sense-antisense gene pairs in embryonic stem cells is consolidated upon differentiation. Mol. Cell. Biol. 35, 2436–2447 (2015).
22. Krogan, N. J. et al. Methylation of histone H3 by Setz in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol. Cell. Biol. 23, 4207–4214 (2003).
23. Baubec, T. et al. Genomic profiling of DNA methyltransferases reveals a role for DNNM3 in genic methylation. Nature 520, 243–247 (2015).
24. Ladd, P. D. et al. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. Hum. Mol. Genet. 16, 3174–3187 (2007).
25. Barbour, V. M. et al. Alpha-thalassemia resulting from a negative chromosomal rearrangement in an adolescent. N. Engl. J. Med. 350, 2418 (2004).
26. Grinchuk, O. V., Tenenouropou, P., Orlov, L. Y., Zhou, J. & Kuznetsov, V. A. Integrative analysis of the human cis-antisense gene pairs, miRNAs and their transcription regulation patterns. Nucleic Acids Res. 38, 534–547 (2010).
27. Gardner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. J. Mol. Biol. 196, 261–282 (1987).
28. Rhead, B. et al. The UCSC Genome Browser database: update 2010. Nucleic Acids Res. 38, D613–D619 (2010).
29. Witkowski, L. et al. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nat. Genet. 46, 438–443 (2014).
30. Schwartingruber, J. et al. DRIVER mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482, 226–231 (2012).
31. Zhang, Y. et al. Model-based analysis of CHIP-Seq (MCs). Genome Biol. 9, R137 (2008).
32. Heinz, S. et al. Simple combinations of lineage-determination transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell. Biol. 25, 5206–5215 (2015).
33. Remington, D. L. et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. USA 98, 11479–11484 (2001).
34. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. PLoS Genet. 2, e190 (2006).
35. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38, 904–909 (2006).
valuable contribution and help in the collection of data and samples used to study the entire family, and the scientists of the McGill University and Genome Quebec and the Genomic Platform of the FR3209 CNRS-Inserm, University of Lorraine, for expert advice and performing high throughput sequencing. W.K.C received support from the Simons Foundation and the JPB Foundation.

Author contributions
J.L.G. recruited case CHU-12122 and her relatives, designed and coordinated the study, interpreted the data, and wrote the paper. C.C. participated in the study design, coordinated and performed experiments, interpreted the data, and participated in writing the paper. J.L.G. and C.C. shared equal contribution. A.O. participated in the design of the study, performed bio-info analyses, interpreted the data, and participated in writing the paper. D.S.R. and D.W. coordinated the metabolic evaluation of WG-3838 and WG-4152 cases, participated in the study design, interpreted the data, and participated in writing the paper. J.M. and J.N. performed analyses of WGS, RNA-Seq and genome-wide ChIP-Seq, interpreted the data, and participated in writing the paper. W.K.C., J.F.B., M.B., and C.C. provided clinical and metabolic data and DNA and cells of the replication study, analyzed and interpreted the related data, and revised the manuscript. M.P. and A.B. performed cell experiments, analyzed and interpreted the related data, and revised the manuscript. M.P. and A.B. and T.T., J.F., A.R., and S.F. performed experiments and analyzed and interpreted the related data. A.M. performed bioinformatic analyses of WGS and genome-wide ChIP-Seq. T.P. and F.H. helped in designing and carried out the ChIP-sequencing experiments. L.M. and V.M. performed RNA-Seq analyses. D.T., B.R., and P.E.M. provided the control populations, analyzed and interpreted the related data, and participated in the preparation of the paper. D.C. interpreted the data and participated in the preparation of the paper. F.F., C.B., I.K., and I.G. performed clinical and metabolic reporting and sample collection, interpreted the data, and revised the manuscript. All authors discussed the results and commented on the manuscript.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-02306-5.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017