HHS Public Access

Author manuscript

Mol Reprod Dev. Author manuscript; available in PMC 2016 May 01.

Published in final edited form as:

Mol Reprod Dev. 2015 May ; 82(5): 377–386. doi:10.1002/mrd.22483.

Genome-wide site-specific differential methylation in the blood of individuals with Klinefelter Syndrome

Emily S. Wan[1,2,*], **Weiliang Qiu**[1], **Jarrett Morrow**[1], **Terri H. Beaty**[3], **Jacqueline Hetmanski**[3], **Barry J. Make**[4], **David A. Lomas**[5], **Edwin K. Silverman**[1,2], and **Dawn L. DeMeo**[1,2]

[1] Channing Division of Network Medicine, Brigham & Women's Hospital, Boston, MA, 02115, USA

[2] Division of Pulmonary and Critical Care, Brigham & Women's Hospital, Boston, MA, 02115, USA

[3] Department of Biostatistics and Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, 21205, USA

[4] Division of Pulmonary, Critical Care and Sleep Medicine, National Jewish Health, Denver, CO, 80206, USA

[5] Wolfson Institute for Biomedical Research, University College London, London, WC1E 6BT, United Kingdom

Abstract

Klinefelter syndrome (KS) (47 XXY) is a common sex-chromosome aneuploidy with an estimated prevalence of 1 in every 660 male births. Investigations into the associations between DNA methylation and the highly variable clinical manifestations of KS have largely focused on the supernumerary X chromosome; systematic investigations of the epigenome have been limited. We obtained genome-wide DNA methylation data from peripheral blood using the Illumina HumanMethylation450K platform in 5 KS (47 XXY), 102 male (46 XY), and 113 female (46 XX) control subjects participating in the chronic obstructive pulmonary disease (COPD) Gene Study. Empirical Bayes-mediated models were used to test for differential methylation by KS status. CpG sites with a false-discovery rate <0.05 from the first-generation HumanMethylation27K platform were further examined in an independent replication cohort of 2 KS subjects, 590 male, and 495 female controls drawn from the International COPD Genetics Network (ICGN). Differential methylation at sites throughout the genome were identified, including 86 CpG sites that were differentially methylated in KS subjects relative to *both* male and female controls. CpG sites annotated to the HEN1 methyltransferase homolog 1 (*HENMT1*), calcyclin-binding protein

[*] Correspondence to: 181 Longwood Avenue, Boston, MA 02115, United State of America, (617) 525-2112 (phone), (617) 525-0958 (facsimile), wan@channing.harvard.edu.

Work for this manuscript was performed at the Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA.

CONFLICT OF INTEREST STATEMENT

In the past three years, E.K.S. received honoraria and consulting fees from Merck and grant support and consulting fees from GlaxoSmithKline. D.A.L. has received grant support and honoraria from, and Chairs the Respiratory Therapy Area Board for, GlaxoSmithKline.

E.S.W., W.Q., J.M., T.H.B, J.H., B.J.M., and D.L.D. have no conflicts to disclose.

(*CACYBP*), and GTPase-activating protein (SH3 domain)-binding protein 1 (*G3BP1*) genes were among the "KS-specific" loci that were replicated in ICGN. We therefore conclude that site-specific differential methylation exists throughout the genome in KS. The functional impact and clinical relevance of these differentially methylated loci should be explored in future studies.

Keywords

[MeSH]: Klinefelter syndrome; DNA methylation; epigenomics; XXY syndrome

INTRODUCTION

Klinefelter syndrome (KS), the most common viable sex-chromosome aneuploidy among males, is characterized by a 47 XXY karyotype. Population-based studies estimate the prevalence of KS to be between 1 in every 450-to-660 male births (Bojesen et al. 2003; Herlihy et al. 2011), although a significant proportion of KS cases remain undiagnosed or are diagnosed later in life due to a lack of obvious physical stigmata at birth (Lahlou et al. 2011) and a high degree of variability in the clinical phenotype (Bojesen et al. 2003).

The classic characteristics of KS – such as tall stature with eunuchoid proportions, gynecomastia, and infertility, as well as a number of other disorders that have an increased prevalence in KS, including osteoporosis (van den Bergh et al. 2001), metabolic syndrome, and type 2 diabetes mellitus (Bojesen et al. 2006) – have been attributed at least in part to the low testosterone levels observed in these patients (Groth et al. 2013). On the other hand, additional disorders associated with KS are less readily explained by perturbations of hypothalamic-pituitary axis, such as an increased incidence of mediastinal tumors among children (Hasle et al. 1995) as well as a range of neurocognitive and psychiatric disturbances (Groth et al. 2013). The variable prevalence of these characteristics and severity among non-mosaic KS subjects, as well as reports of phenotypic discordance among monozygotic twins (Davis and Wade 1972; Michalski et al. 1978; Schlegelberger et al. 1984), suggest a role for epigenetic mediators in their presentation.

DNA methylation studies in KS have largely focused on the supernumerary X chromosome to assess the impact of parent-of-origin effects and skewed X-inactivation on physical characteristics, cognitive functioning, and psychiatric diseases (Boks et al. 2007; Bruining et al. 2009; Ross et al. 2008; Stemkens et al. 2006; Vawter et al. 2007). Consequently, DNA methylation studies examining autosomal loci in KS have been limited. A study of long interspersed nuclear elements (LINE-1), a class of repetitive sequences with retrotransposon potential that account for a significant proportion of the human genome, revealed decreased methylation of these elements at autosomal sites; the authors of this study postulated that this could be related to the increased genome size in KS (Singer et al. 2012). A recent study examining the genome-wide DNA methylation patterns of a single KS subject at 2 sites in the brain demonstrated many differentially methylated regions throughout the autosomal genome relative to males and females (Viana et al. 2014). Here, we examined the global genome methylation patterns of KS subjects in peripheral blood.

RESULTS

Six subjects with KS had been previously identified through analysis of genome-wide genotyping intensity data in the chronic obstructive pulmonary disease (COPD) Gene database, the discovery cohort (Supplementary Figure S1). Of these KS subjects, 5 had sufficient DNA isolated from whole blood for genome-wide methylation analysis on the Illumina HumanMethylation450K platform. Two additional KS subjects were identified in the replication cohort, the International COPD Genetics Network (ICGN), through analysis of DNA methylation patterns on the X and Y chromosome obtained from blood using the first generation HumanMethylation27K platform (Supplementary Figures S2 and S3). Characteristics of the discovery and replication cohorts are presented in Table 1. There were no significant differences in ancestry, current smoking rates, or cumulative-smoke exposure between KS and control subjects in each group. In the discovery cohort, KS subjects had a significantly higher mean body-mass index than male controls, and were significantly younger than both male and female controls. There were no significant differences by age or body-mass index in the replication cohort.

KS subjects versus male controls

Using empirical Bayes-mediated models adjusting for age, race, and body-mass index, 176 autosomal cytosine-phosphate-guanine (CpG) sites in the Infinium I analysis (Supplementary Table S1) and 223 sites in the Infinium II analysis (Supplementary Table S2) were found to be significantly associated with KS status, at a false-discovery rate (FDR) <0.05 in the discovery cohort. Significantly associated sites were identified throughout the genome, and are illustrated qualitatively in a Manhattan plot (Supplementary Figure S4A). Adjustment for current smoking status did not significantly change the results: 166 Infinium I and 218 Infinium II sites (of which 165 and 212 sites, respectively, overlapped with CpG sites identified in the primary analysis) met a FDR <0.05 (data not shown). Although probes with a high degree of cross-hybridization with X-chromosome sequences were removed during data cleaning (Chen et al. 2013), we examined the source sequence of each of the significantly associated sites and identified an additional 15 CpG sites in the Infinium I analysis and 21 CpG sites in the Infinium II analysis with moderate degrees of homology to sequences located on the X chromosome.

Twenty-one sites with a FDR<0.05 in the discovery cohort were available on the HumanMethylation27K platform used in the replication cohort. Of these, seven sites demonstrated a significant *(P*<0.05) association with KS status in the replication cohort; in each case, the direction of effect was consistent with that observed in the discovery cohort (Table 2).

KS subjects versus female controls

When we examined the methylation patterns of KS subjects relative to females, 48 CpG sites in the Infinium I analysis (Supplementary Table S3) and 87 CpG sites in the Infinium II analysis (Supplementary Table S4) were significant, at a FDR<0.05. Differentially methylated sites were located throughout the autosomal genome (Supplementary Figure S4B). Again, adjustment for current smoking status did not significantly change the results:

48 Infinium I and 98 Infinium II sites (of which 47 and 87 of the sites, respectively, overlapped with sites identified in the primary analysis) had a FDR<0.05 (data not shown). Ten of the significant CpG sites were available on the HumanMethylation27K chip used in the replication cohort; of these, three demonstrated significant differential methylation between KS subjects and female controls with a consistent direction of effect (Table 3).

Overlap between KS versus male and KS versus female analyses

The overlap between differentially methylated sites in the KS versus male controls and KS versus female controls analyses is illustrated qualitatively in Figure 1. Notably, the direction of effect at all 86 overlapping sites was consistent relative to both males and females (Supplementary Table S5). All 3 sites that successfully replicated in the KS versus female analyses were among the 10 replicated loci in the KS versus male controls analysis. Boxplots of the mean methylation among males, females, and KS subjects at these 3 overlapping sites in COPDGene are illustrated in Figure 2.

Overlap between sites previously reported as differentially methylated in KS

We examined the overlap between differentially methylated sites identified in our analysis in blood with autosomal sites from brain tissue recently reported by Viana et al (2014). One differentially methylated region in the prefrontal cortex annotated to the germinal center-associated, signaling and motility-like (*GCSAML*) gene and one region in the cerebellum annotated to *C9orf64* overlapped with loci identified in our analysis (Supplementary Table S6).

Analyses stratified by race

Although our primary analysis in the discovery cohort contained a covariate adjustment for race, we performed a subgroup analysis in the discovery cohort by first examining non-Hispanic-white and African-American subjects separately, followed by a meta-analysis of non-Hispanic-white and African-American results. In the analysis of non-Hispanic-white KS (n = 2) versus non-Hispanic-white males (n = 53), 537 Type I and 5258 Type II CpG sites had an FDR<0.05. In the analysis of non-Hispanic-white KS versus non-Hispanic-white females (n = 55), 116 Type I and 1951 Type II CpG sites met an FDR <0.05. In the African American-only analysis, 281 Infinium I and 347 Infinium II sites were differentially methylated in the African-American KS (n = 3) versus African-American male controls (n = 49) while 96 Infinium I and 190 Infinium II sites were differentially methylated relative to African-American females (n = 58). The overlap between CpG sites significant in the non-Hispanic-white-only and African-American-only analyses with CpG sites identified in the primary analyses is illustrated qualitatively in Supplementary Figure S5. The overlap between CpG sites significant in the non-Hispanic-white-only and sites in the African-American-only analyses was modest (Supplementary Figure S6).

We performed a sample-size-weighted meta-analysis of the non-Hispanic-white-only and African-American-only analyses. Six hundred and eleven CpG sites (309 of which overlapped with the 399 sites identified in the original analysis) met an FDR<0.05 in the KS vs. male controls analysis, while 270 CpG sites (96 of which overlapped with the 135 sites identified in the original analysis) were significant in the KS vs. female controls analysis

(Supplementary Figure S7). One hundred and seventy three CpG sites were significant in *both* the KS vs. male and KS vs. female meta-analyses; 68 of these sites overlapped with the 86 sites identified in the primary analysis.

Assessment of confounders due to cell-type heterogeneity and stochastic effects

Because whole blood is comprised of multiple circulating cell types, each with potentially distinct epigenetic profiles, confounding due to cell-type heterogeneity is a concern often raised in epigenome-wide association studies (Jaffe and Irizarry 2014). We therefore determined the cell-type composition of each sample in the discovery cohort using the raw methylation data and the "estimateCellCount" function in the *minfi* package (Aryee et al. 2014), finding no significant differences by KS status relative to either male or female controls (Supplementary Table S7).

Because the number of KS cases in each cohort was modest, causing the results to be susceptible to outlier effects, we performed one-way permutation testing based on 10,000 Monte-Carlo resamplings at the replicated loci listed in Tables 2 (KS vs. male) and 3 (KS vs. female). Permutation *P* values for each of the loci are listed in Supplementary Table S8; these data support the premise that our findings at these loci are unlikely to be due to random outlier effects.

Exploratory analysis of site-specific differential methylation in the sex chromosomes

CpG sites located on the X and Y chromosomes were initially excluded from the cleaning and normalization that the autosomal data underwent because of expected differences resulting from X-chromosome inactivation. Given that the sex chromosomes encode for genes known to impact biological processes relevant to the clinical manifestations observed in KS, we extracted and examined CpG sites on the X and Y chromosomes in a separate analysis. Using empirical Bayes-mediated models adjusting for age, race, and body-mass index, we identified 23 CpG sites on the Y chromosome that were nominally (unadjusted $P<0.05$) differentially methylated in KS subjects relative to male controls (Supplementary Table S9). The six most-highly associated sites were annotated to the testis-specific transcript, Y-linked 14 (*TTTY14*) region, and demonstrated relative hypermethylation in KS. We identified 260 nominally (unadjusted $P<0.05$) differentially methylated sites on the X chromosome in KS subjects relative to female controls (Supplementary Table S10). While some of the differentially methylated sites mapped to regions with known homology to sequences located on the Y chromosome or to the pseudoautosomal regions, the majority of the differentially methylated sites on the X chromosome were *not* located within these regions; therefore, differential methylation beyond that expected due to X-inactivation may exist.

DISCUSSION

KS is attributable to a discreet genetic cause, yet our understanding of the mechanisms underlying its often highly variable clinical phenotypic presentation in KS subjects remains incomplete. Investigations into the etiology of the protean manifestations associated with this syndrome have largely focused on the supranumery X; consequently, systematic

interrogations of loci throughout the genome that may contribute to the pathology associated with the syndrome have been limited.

We examined the global DNA methylation patterns of KS subjects relative to both male and female controls, and present evidence supportive of site-specific differences in methylation throughout the genome – the vast majority of which have not previously been associated with KS. The number of differentially methylated autosomal sites between KS subjects and males was greater than the number of differentially methylated sites identified when KS subjects were compared to females, suggesting that KS subjects may have an autosomal epigenetic landscape more similar to females. A second striking feature is the large difference in mean methylation observed at the majority of the differentially methylated sites (Supplementary Figure S8): the mean absolute difference in methylation at autosomal sites was 12.8% in the KS vs. male controls and 14.3% in the KS vs. female controls. Finally, an enrichment of differentially methylated sites located within CpG islands was observed. The functional significance of these observations is not currently known, and should be explored in future studies.

Perhaps most intriguing of all was the identification of 86 autosomal sites where methylation among KS subjects appeared distinct relative to *both* males and females. Genes annotated to this “KS-specific” subset include both biologically plausible loci as well as novel loci for which the pathobiological relationship to KS remains unknown. Several CpG sites annotated to the nuclear receptor binding SET domain protein 1 (*NSD1*) were significantly hypermethylated in KS subjects. The *NSD1* gene product is a putative nuclear transcription factor/regulator and histone methyltransferase that has been shown to enhance trans-activation of the androgen receptor. Mutations in *NSD1* have been linked to Sotos syndrome, a condition characterized by excessive growth in early childhood as well as neurocognitive and social deficits (Kurotaki et al. 2001; Lucio-Eterovic et al. 2010). CpG sites annotated to the GTPase-activating protein (SH3 domain)-binding protein 1 (*G3BP1*) as well as sites annotated to the calcyclin-binding protein (*CACYBP*) are among the “KS-specific” sites that were replicated in the ICGN cohort. The gene product of *G3BP1* is an RNA-binding protein involved in Ras signal transduction, RNA metabolism, and stress-granule formation (Irvine et al. 2004; Matsuki et al. 2013). The product of *CACYBP* is involved in ubiquitin-mediated proteolysis, and has been investigated in the context of tumorigenesis and metastasis (Zhai et al. 2008). Cg00328227, the third replicated “KS-specific” site, is annotated to the HEN1 methyltransferase homolog 1 (*HENMT1*, also known as *C1orf59*), and demonstrated moderate homology to sequences on the X chromosome. We noted hypermethylation among KS samples at the *HENMT1* site relative to *both* male and female controls – a phenomenon that could not readily be explained by cross-hybridization with sequences subject to X inactivation. Additional sites annotated to *HENMT1*, including cg24737783, that did not have probe homology to X chromosome sequences, were also noted to be relative hypermethylated in KS subjects. The significance and cause of the relative hypermethylation observed in KS subjects at these sites remains to be elucidated.

Because the impact of ancestry on epigenome-wide methylation patterns is incompletely understood (Barfield et al. 2014; Moen et al. 2013), we undertook several subgroup analyses stratified by race. Surprisingly, despite decreased numbers of subjects in each subgroup

analysis, the number of sites that met an FDR threshold <0.05 was *increased* relative to the number of sites identified in the original analysis. This was noted in both the non-Hispanic-white- and African-American-only analyses, and may reflect increased susceptibility to outlier effects due to small sample sizes. Interestingly, the number of overlapping sites between the non-Hispanic-white-only and African-American-only analyses was modest, which may be indicative of ancestry-specific methylation patterns in KS; additional studies are needed to explore this possibility. Reassuringly, the majority of sites identified in the primary analysis were robust when examined using race-stratified meta-analysis.

The overlap between the differentially methylated sites reported in our manuscript and those reported in Viana et al (2014) is modest. Possible reasons for the lack of overlap include tissue-specific methylation patterns between brain and peripheral blood as well differences in disease processes present in the independent KS populations. The KS subjects examined by Viana and colleageues were known to have schizophrenia, a history of alcohol abuse, and physical stigmata of advanced liver disease; the impact of each of these disease processes on the methylome may have contributed to the differential sites identified in each of our studies.

The strengths of the current study include the use of an unbiased approach for the identification of differentially methylated loci; rigorous statistical methods and thresholds for significance; and replication in an independent population. We acknowledge that the number of KS subjects included in our study is modest, possibly limiting our ability to detect moderate differences in methylation at additional loci. A second potential limitation to our study is the absence of comprehensive, simultaneous gene-expression data to directly assess the functional impact of differential methylation on expression. Although we did perform an exploratory *in silico* analysis of publicly available expression data from blood in 5 KS subjects relative to 5 male controls (Huang et al. 2015) (accession GSE47584), and identified 13 nominally differentially expressed transcripts within ±50kb of the 86 “KS-specific” loci described in our study (data not shown), additional studies using simultaneously collected methylation and expression data are needed for more robust validation. A third potential limitation of this study is the lack of subjects who have never smoked cigarettes. since the parent-study populations were determined based on smoking and smoking-related lung diseases. Therefore, testing if our findings can be extended to never-smoked populations should be explored in future studies. Finally, none of the KS subjects examined in our analyses reported a history of KS symptoms on the questionnaire, implying that the majority of our subjects were subclinical cases. It is not known how the methylation patterns of our KS subjects would compare to clinically diagnosed (and presumably, more phenotypically affected) cases.

In conclusion, we provide provocative evidence that site-specific differential methylation throughout the genome exists in KS individuals. The functional impact and clinical relevance of these differentially methylated loci warrants exploration in future studies.

MATERIALS AND METHODS

Cohorts and Samples

The discovery cohort consisted of 5 KS subjects and 102 male and 113 female control subjects drawn from current and former smokers enrolled in the COPDGene Study (clinicaltrials.gov identifier: NCT 000608764) (Regan et al. 2010). COPDGene is a cross-sectional observational study that enrolled self-identified non-Hispanic-white or African-American individuals between the ages of 45-80 years with at least 10 pack-years of smoking exposure from 21 clinical centers throughout the United States. Subjects completed a structured questionnaire, pre- and post- bronchodilator spirometry, 6-min walk test, and volumetric-computed chest tomography. Peripheral blood for DNA analysis was collected at the time of enrollment.

The replication cohort consisted of 2 KS subjects and 590 male and 495 female controls drawn from the ICGN (Patel et al. 2008; Zhu et al. 2007), a family-based study that recruited probands between the ages of 45-65 with ≥5 pack-years of smoking and COPD (defined as an forced expiratory volume in the first sec (FEV_1)-to-forced vital capacity (FVC) ratio (FEV_1/FVC)<90% with a predicted FEV_1<60%), and eligible siblings with ≥5 pack-years of smoking exposure. All subjects included in this subset were of self-reported European descent. Subjects completed questionnaires and spirometry, and provided whole-blood for DNA analysis.

Internal Review Board approval was obtained at each of the participating recruitment sites in both COPDGene and ICGN. All subjects provided written informed consent.

DNA Methylation

DNA from peripheral blood was bisulfite-converted using the EZ DNA Methylation Gold Kit (Zymo Research, Irvine, CA). In the discovery cohort, quantitative DNA methylation at 485,512 individual CpG sites was obtained using the Illumina HumanMethylation450K array. Data pre-processing and quality control for autosomal CpG sites were performed with the R programming language (release 2.15.0) using packages *limma (Smyth 2005)*, *minfi (Aryee et al. 2014)*, and *sva (Leek et al.)*. Because the HumanMethylation450K chip employs 2 distinct probe types (Infinium I and II), quantile normalization, adjustment for batch effects using the ComBat function (v 3.4.0), and subsequent analyses were performed separately by probe chemistry in the discovery cohort. Additionally, non-specific probes with a high degree of homology to more than one site in the genome –as reported by Chen et al (2013)– were removed prior to analysis. A summary of the data cleaning is provided as Supplementary Material. Raw beta values annotated to the X and Y chromosomes were extracted and analyzed in a separate exploratory analysis.

In the replication cohort, quantitative DNA methylation data were generated using bisulfite-converted DNA and the Illumina HumanMethylation27K array. Raw beta values for the 26 CpG sites that were significantly associated in the discovery cohort and assayed on the HumanMethylation27K chip were also extracted and analyzed.

Statistical analysis

In the discovery cohort, empirical Bayes-mediated models were used to test for differential methylation by KS status at 442,059 autosomal CpG sites; models were adjusted for age, race, and body-mass index. A second model including a covariate for current smoking status was also tested. A FDR<0.05 was used to denote statistical significance. In the replication cohort, generalized-linear-mixed-effect models were employed to test for differential methylation by binary KS status while accounting for familial correlations. An unadjusted $P<0.05$ with a consistent direction of effect was considered significant. Combined *P* values were calculated using the Liptak method (Liptak 1958).

For the race-stratified analyses in COPDGene, models were separately adjusted for age and body-mass index in non-Hispanic-white and African-America cohorts. Meta-analysis weighted by sample size was performed using METAL software (Willer et al. 2010), and FDRs were calculated from meta-analysis Z-scores using the tail-area-based FDR as implemented in the package *fdrtool* (v 1.2.12) (Klaus and Strimmer 2014). Permutation testing was implemented using the *coin* (Hothorn et al. 2008) package for the R programming language.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

International COPD Genetics Network (ICGN) investigators:

Edwin K. Silverman, Brigham & Women's Hospital, Boston, MA, USA; David A. Lomas, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK; Barry J. Make, National Jewish Medical and Research Center, Denver, CO, USA; Alvar Agusti and Jaume Sauleda, Hospital Universitari Son Dureta, Fundación Caubet-Cimera and Ciber Enfermedades Respiratorias, Spain; Peter M.A. Calverley, University of Liverpool, UK; Claudio F. Donner, Division of Pulmonary Disease, S. Maugeri Foundation, Veruno (NO), Italy; Robert D. Levy, University of British Columbia, Vancouver, Canada; Peter D. Paré, University of British Columbia, Vancouver, Canada; Stephen Rennard, Section of Pulmonary & Critical Care, University of Nebraska Medical Center, Omaha, NE, USA; Jørgen Vestbo, Department of Cardiology and Respiratory Medicine, Hvidovre Hospital, Copenhagen, Denmark; Emiel F.M. Wouters, University Hospital Maastricht, The Netherlands.

COPDGene Core Teams:

Administrative Core: James Crapo, MD (PI), Edwin Silverman, MD, PhD (PI), Barry Make, MD, Elizabeth Regan, MD, PhD, Rochelle Lantz, Lori Stepp, Sandra Melanson,

Genetic Analysis Core: Terri Beaty, PhD, Barbara Klanderman, PhD, Nan Laird, PhD, Christoph Lange, PhD, Michael Cho, MD, Stephanie Santorico, PhD, John Hokanson, MPH, PhD, Dawn DeMeo, MD, MPH, Nadia Hansel, MD, MPH, Craig Hersh, MD, MPH, Peter Castaldi, MD, MSc, Merry-Lynn McDonald, PhD, Jing Zhou, MD, PhD, Manuel Mattheissen, MD, PhD, Emily Wan, MD, Megan Hardin, MD, Jacqueline Hetmanski, MS, Margaret Parker, MS, Tanda Murray, MS

Imaging Core: David Lynch, MB, Joyce Schroeder, MD, John Newell, Jr., MD, John Reilly, MD, Harvey Coxson, PhD, Philip Judy, PhD, Eric Hoffman, PhD, George Washko, MD, Raul San Jose Estepar, PhD, James Ross, MSc, Mustafa Al Qaisi, MD, Jordan Zach, Alex Kluiber, Jered Sieren, Tanya Mann, Deanna Richert, Alexander McKenzie, Jaleh Akhavan, Douglas Stinson

PFT QA Core, LDS Hospital, Salt Lake City, UT: Robert Jensen, PhD

Biological Repository, Johns Hopkins University, Baltimore, MD: Homayoon Farzadegan, PhD, Stacey Meyerer, Shivam Chandan, Samantha Bragan

Data Coordinating Center and Biostatistics, National Jewish Health, Denver, CO: Douglas Everett, PhD, Andre Williams, PhD, Carla Wilson, MS, Anna Forssen, MS, Amber Powell, Joe Piccoli

Epidemiology Core, University of Colorado School of Public Health, Denver, CO: John Hokanson, MPH, PhD, Marci Sontag, PhD, Jennifer Black-Shinn, MPH, Gregory Kinney, MPH, PhDc, Sharon Lutz, MPH, PhD.

We further wish to acknowledge the COPDGene Investigators from the participating Clinical Centers:

Ann Arbor VA: Jeffrey Curtis, MD, Ella Kazerooni, MD

Baylor College of Medicine, Houston, TX: Nicola Hanania, MD, MS, Philip Alapat, MD, Venkata Bandi, MD, Kalpalatha Guntupalli, MD, Elizabeth Guy, MD, Antara Mallampalli, MD, Charles Trinh, MD, Mustafa Atik, MD, Hasan Al-Azzawi, MD, Marc Willis, DO, Susan Pinero, MD, Linda Fahr, MD, Arun Nachiappan, MD, Collin Bray, MD, L. Alexander Frigini, MD, Carlos Farinas, MD, David Katz, MD, Jose Freytes, MD, Anne Marie Marciel, MD

Brigham and Women's Hospital, Boston, MA: Dawn DeMeo, MD, MPH, Craig Hersh, MD, MPH, George Washko, MD, Francine Jacobson, MD, MPH, Hiroto Hatabu, MD, PhD, Peter Clarke, MD, Ritu Gill, MD, Andetta Hunsaker, MD, Beatrice Trotman-Dickenson, MBBS, Rachna Madan, MD

Columbia University, New York, NY: R. Graham Barr, MD, DrPH, Byron Thomashow, MD, John Austin, MD, Belinda D'Souza, MD

Duke University Medical Center, Durham, NC: Neil MacIntyre, Jr., MD, Lacey Washington, MD, H Page McAdams, MD

Fallon Clinic, Worcester, MA: Richard Rosiello, MD, Timothy Bresnahan, MD, Joseph Bradley, MD, Sharon Kuong, MD, Steven Meller, MD, Suzanne Roland, MD

Health Partners Research Foundation, Minneapolis, MN: Charlene McEvoy, MD, MPH, Joseph Tashjian, MD

Johns Hopkins University, Baltimore, MD: Robert Wise, MD, Nadia Hansel, MD, MPH, Robert Brown, MD, Gregory Diette, MD, Karen Horton, MD

Los Angeles Biomedical Research Institute at Harbor UCLA Medical Center, Los Angeles, CA: Richard Casaburi, MD, Janos Porszasz, MD, PhD, Hans Fischer, MD, PhD, Matt Budoff, MD, Mehdi Rambod, MD

Michael E. DeBakey VAMC, Houston, TX: Amir Sharafkhaneh, MD, Charles Trinh, MD, Hirani Kamal, MD, Roham Darvishi, MD, Marc Willis, DO, Susan Pinero, MD, Linda Fahr, MD, Arun Nachiappan, MD, Collin Bray, MD, L. Alexander Frigini, MD, Carlos Farinas, MD, David Katz, MD, Jose Freytes, MD, Anne Marie Marciel, MD

Minneapolis VA: Dennis Niewoehner, MD, Quentin Anderson, MD, Kathryn Rice, MD, Audrey Caine, MD

Morehouse School of Medicine, Atlanta, GA: Marilyn Foreman, MD, MS, Gloria Westney, MD, MS, Eugene Berkowitz, MD, PhD

National Jewish Health, Denver, CO: Russell Bowler, MD, PhD, David Lynch, MB, Joyce Schroeder, MD, Valerie Hale, MD, John Armstrong, II, MD, Debra Dyer, MD, Jonathan Chung, MD, Christian Cox, MD

Temple University, Philadelphia, PA: Gerard Criner, MD, Victor Kim, MD, Nathaniel Marchetti, DO, Aditi Satti, MD, A. James Mamary, MD, Robert Steiner, MD, Chandra Dass, MD, Libby Cone, MD

University of Alabama, Birmingham, AL: William Bailey, MD, Mark Dransfield, MD, Michael Wells, MD, Surya Bhatt, MD, Hrudaya Nath, MD, Satinder Singh, MD

University of California, San Diego, CA: Joe Ramsdell, MD, Paul Friedman, MD

University of Iowa, Iowa City, IA: Alejandro Cornellas, MD, John Newell, Jr., MD, Edwin JR van Beek, MD, PhD

University of Michigan, Ann Arbor, MI: Fernando Martinez, MD, MeiLan Han, MD, Ella Kazerooni, MD

University of Minnesota, Minneapolis, MN: Christine Wendt, MD, Tadashi Allen, MD

University of Pittsburgh, Pittsburgh, PA: Frank Sciurba, MD, Joel Weissfeld, MD, MPH, Carl Fuhrman, MD, Jessica Bon, MD, Danielle Hooper, MD

University of Texas Health Science Center at San Antonio, San Antonio, TX: Antonio Anzueto, MD, Sandra Adams, MD, Carlos Orozco, MD, Mario Ruiz, MD, Amy Mumbower, MD, Ariel Kruger, MD, Carlos Restrepo, MD, Michael Lane, MD

FUNDING

This work was supported by: National Institutes of Health [R01 HL089856 to E.K.S., R01 HL089897 to J.D.C. and E.K.S., RC1 HL100076 to E.K.S., P01 HL105339 to E.K.S., R01 HL089438 to D.L.D]; and a Parker B. Francis Foundation Fellowship to E.S.W. The COPDGene® project is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer Ingelheim, Novartis, Pfizer, Siemens and Sunovion.

ABBREVIATIONS

COPD	chronic obstructive pulmonary disease
CpG	cytosine-phosphate-guanine
FDR	false-discovery rate
ICGN	International COPD Genetics Network
KS	Klinefelter syndrome

REFERENCES

Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014; 30(10):1363–1369. [PubMed: 24478339]

Barfield RT, Almli LM, Kilaru V, Smith AK, Mercer KB, Duncan R, Klengel T, Mehta D, Binder EB, Epstein MP, Ressler KJ, Conneely KN. Accounting for population stratification in DNA methylation studies. Genet Epidemiol. 2014; 38(3):231–241. [PubMed: 24478250]

Bojesen A, Juul S, Gravholt CH. Prenatal and postnatal prevalence of Klinefelter syndrome: a national registry study. J Clin Endocrinol Metab. 2003; 88(2):622–626. [PubMed: 12574191]

Bojesen A, Kristensen K, Birkebaek NH, Fedder J, Mosekilde L, Bennett P, Laurberg P, Frystyk J, Flyvbjerg A, Christiansen JS, Gravholt CH. The metabolic syndrome is frequent in Klinefelter's syndrome and is associated with abdominal obesity and hypogonadism. Diabetes Care. 2006; 29(7): 1591–1598. [PubMed: 16801584]

Boks MP, de Vette MH, Sommer IE, van Rijn S, Giltay JC, Swaab H, Kahn RS. Psychiatric morbidity and X-chromosomal origin in a Klinefelter sample. Schizophr Res. 2007; 93(1-3):399–402. [PubMed: 17459660]

Bruining H, Swaab H, Kas M, van Engeland H. Psychiatric characteristics in a self-selected sample of boys with Klinefelter syndrome. Pediatrics. 2009; 123(5):e865–870. [PubMed: 19364768]

Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013; 8(2):203–209. [PubMed: 23314698]

Davis JC, Wade AP. Discordant identical twins. 3. Klinefelter's syndrome discordant for carcinoma of the breast. Practitioner. 1972; 209(251):380–383. [PubMed: 5081320]

Groth KA, Skakkebaek A, Host C, Gravholt CH, Bojesen A. Clinical review: Klinefelter syndrome--a clinical update. J Clin Endocrinol Metab. 2013; 98(1):20–30. [PubMed: 23118429]

Hasle H, Mellemgaard A, Nielsen J, Hansen J. Cancer incidence in men with Klinefelter syndrome. Br J Cancer. 1995; 71(2):416–420. [PubMed: 7841064]

Herlihy AS, Halliday JL, Cock ML, McLachlan RI. The prevalence and diagnosis rates of Klinefelter syndrome: an Australian comparison. Med J Aust. 2011; 194(1):24–28. [PubMed: 21449864]

Hothorn T, Hornik K, van de Wiel MA, Zeileis A. Implementing a Class of Permutation Tests: The coin Package. Journal of Stastical Software. 2008; 28(8):1–23.

Huang J, Zhang L, Deng H, Chang L, Liu Q, Liu P. Global transcriptome analysis of peripheral blood identifies the most significantly down-regulated genes associated with metabolism regulation in Klinefelter syndrome. Molecular reproduction and development. 2015; 82(1):17–25. [PubMed: 25581374]

Irvine K, Stirling R, Hume D, Kennedy D. Rasputin, more promiscuous than ever: a review of G3BP. Int J Dev Biol. 2004; 48(10):1065–1077. [PubMed: 15602692]

Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014; 15(2):R31. [PubMed: 24495553]

Klaus B, Strimmer K. fdrtool: Estimation of (Local) False Discovery Rates (version 1.2.12). 2014

Kurotaki N, Harada N, Yoshiura K, Sugano S, Niikawa N, Matsumoto N. Molecular characterization of NSD1, a human homologue of the mouse Nsd1 gene. Gene. 2001; 279(2):197–204. [PubMed: 11733144]

Lahlou N, Fennoy I, Ross JL, Bouvattier C, Roger M. Clinical and hormonal status of infants with nonmosaic XXY karyotype. Acta Paediatr. 2011; 100(6):824–829. [PubMed: 21429009]

Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. sva: Surrogate Variabel Analysis, R package version 3.4.0.

Liptak T. On the combination of independent tests. Magyar Tud Akad Mat Kutato Int. 1958; 3:171–197.

Lucio-Eterovic AK, Singh MM, Gardner JE, Veerappan CS, Rice JC, Carpenter PB. Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. Proc Natl Acad Sci U S A. 2010; 107(39):16952–16957. [PubMed: 20837538]

Matsuki H, Takahashi M, Higuchi M, Makokha GN, Oie M, Fujii M. Both G3BP1 and G3BP2 contribute to stress granule formation. Genes Cells. 2013; 18(2):135–146. [PubMed: 23279204]

Michalski JP, Snyder SM, McLeod RL, Talal N. Monozygotic twins with Klinefelter's syndrome discordant for systemic lupus erythematosus and symptomatic myasthenia gravis. Arthritis Rheum. 1978; 21(3):306–309. [PubMed: 565638]

Moen EL, Zhang X, Mu W, Delaney SM, Wing C, McQuade J, Myers J, Godley LA, Dolan ME, Zhang W. Genome-wide variation of cytosine modifications between European and African populations and the implications for complex traits. Genetics. 2013; 194(4):987–996. [PubMed: 23792949]

Patel BD, Coxson HO, Pillai SG, Agusti AG, Calverley PM, Donner CF, Make BJ, Muller NL, Rennard SI, Vestbo J, Wouters EF, Hiorns MP, Nakano Y, Camp PG, Nasute Fauerbach PV, Screaton NJ, Campbell EJ, Anderson WH, Pare PD, Levy RD, Lake SL, Silverman EK, Lomas DA. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2008; 178(5):500–505. [PubMed: 18565956]

Regan EA, Hokanson JE, Murphy JR, Make B, Lynch DA, Beaty TH, Curran-Everett D, Silverman EK, Crapo JD. Genetic epidemiology of COPD (COPDGene) study design. COPD. 2010; 7(1):32–43. [PubMed: 20214461]

Ross JL, Roeltgen DP, Stefanatos G, Benecke R, Zeger MP, Kushner H, Ramos P, Elder FF, Zinn AR. Cognitive and motor development during childhood in boys with Klinefelter syndrome. Am J Med Genet A. 2008; 146A(6):708–719. [PubMed: 18266239]

Schlegelberger B, Schlegelberger T, Kekow J, Gross WL, Grote W. [Monozygotic twins with Klinefelter syndrome (concordant) and systemic lupus erythematodes (discordant)]. Klin Wochenschr. 1984; 62(19):906–910. [PubMed: 6542161]

Singer H, Walier M, Nusgen N, Meesters C, Schreiner F, Woelfle J, Fimmers R, Wienker T, Kalscheuer VM, Becker T, Schwaab R, Oldenburg J, El-Maarri O. Methylation of L1Hs promoters is lower on the inactive X, has a tendency of being higher on autosomes in smaller genomes and shows inter-individual variability at some loci. Hum Mol Genet. 2012; 21(1):219–235. [PubMed: 21972244]

Smyth, GK. Limma: linear models for microarray data.. In: Gentleman, R.; Carey, V.; Dudoit, S.; Irizarry, RA.; Huber, W., editors. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer; New York: 2005. p. 397-420.

Stemkens D, Roza T, Verrij L, Swaab H, van Werkhoven MK, Alizadeh BZ, Sinke RJ, Giltay JC. Is there an influence of X-chromosomal imprinting on the phenotype in Klinefelter syndrome? A clinical and molecular genetic study of 61 cases. Clin Genet. 2006; 70(1):43–48. [PubMed: 16813603]

van den Bergh JP, Hermus AR, Spruyt AI, Sweep CG, Corstens FH, Smals AG. Bone mineral density and quantitative ultrasound parameters in patients with Klinefelter's syndrome after long-term testosterone substitution. Osteoporos Int. 2001; 12(1):55–62. [PubMed: 11305084]

Vawter MP, Harvey PD, DeLisi LE. Dysregulation of X-linked gene expression in Klinefelter's syndrome and association with verbal cognition. Am J Med Genet B Neuropsychiatr Genet. 2007; 144B(6):728–734. [PubMed: 17347996]

Viana J, Pidsley R, Troakes C, Spiers H, Wong CC, Al-Sarraj S, Craig I, Schalkwyk L, Mill J. Epigenomic and transcriptomic signatures of a Klinefelter syndrome (47,XXY) karyotype in the brain. Epigenetics. 2014; 9(4):587–599. [PubMed: 24476718]

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26(17):2190–2191. [PubMed: 20616382]

Zhai H, Shi Y, Jin H, Li Y, Lu Y, Chen X, Wang J, Ding L, Wang X, Fan D. Expression of calcyclin-binding protein/Siah-1 interacting protein in normal and malignant human tissues: an immunohistochemical survey. J Histochem Cytochem. 2008; 56(8):765–772. [PubMed: 18443365]

Zhu G, Warren L, Aponte J, Gulsvik A, Bakke P, Anderson WH, Lomas DA, Silverman EK, Pillai SG. The SERPINE2 gene is associated with chronic obstructive pulmonary disease in two large populations. Am J Respir Crit Care Med. 2007; 176(2):167–173. [PubMed: 17446335]

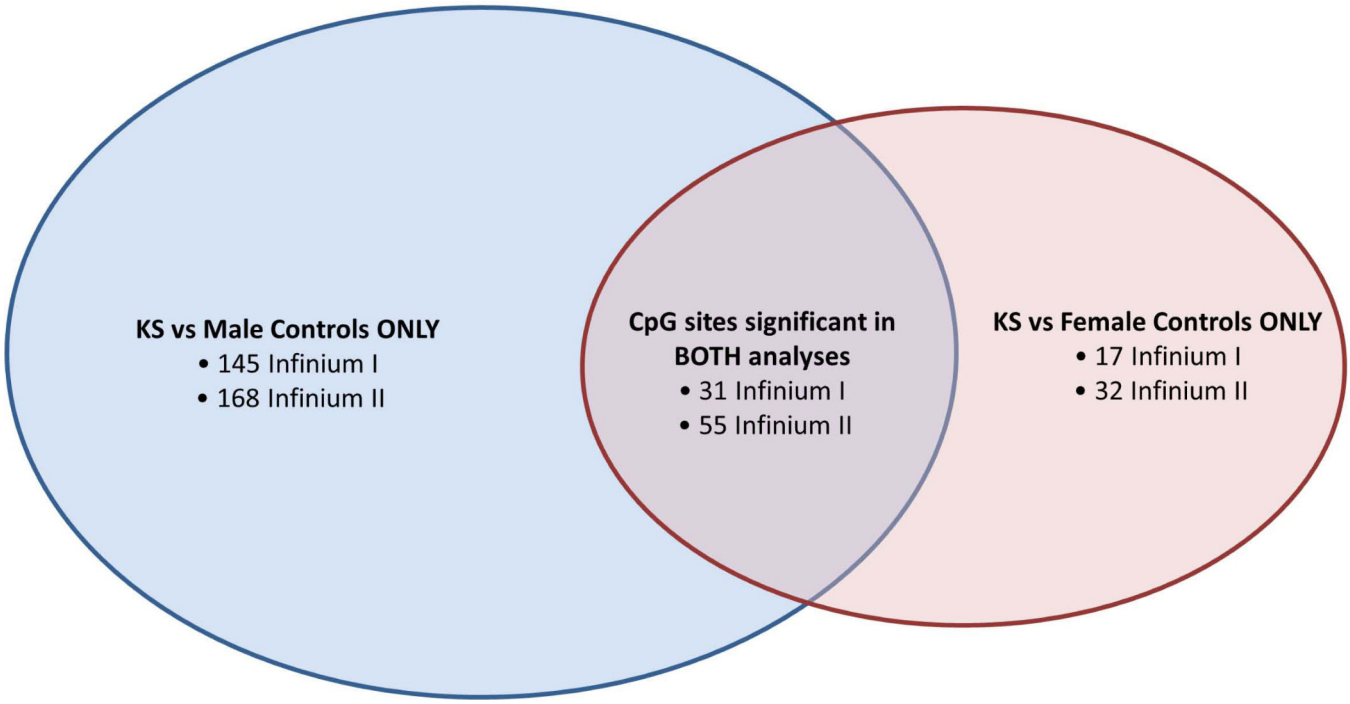
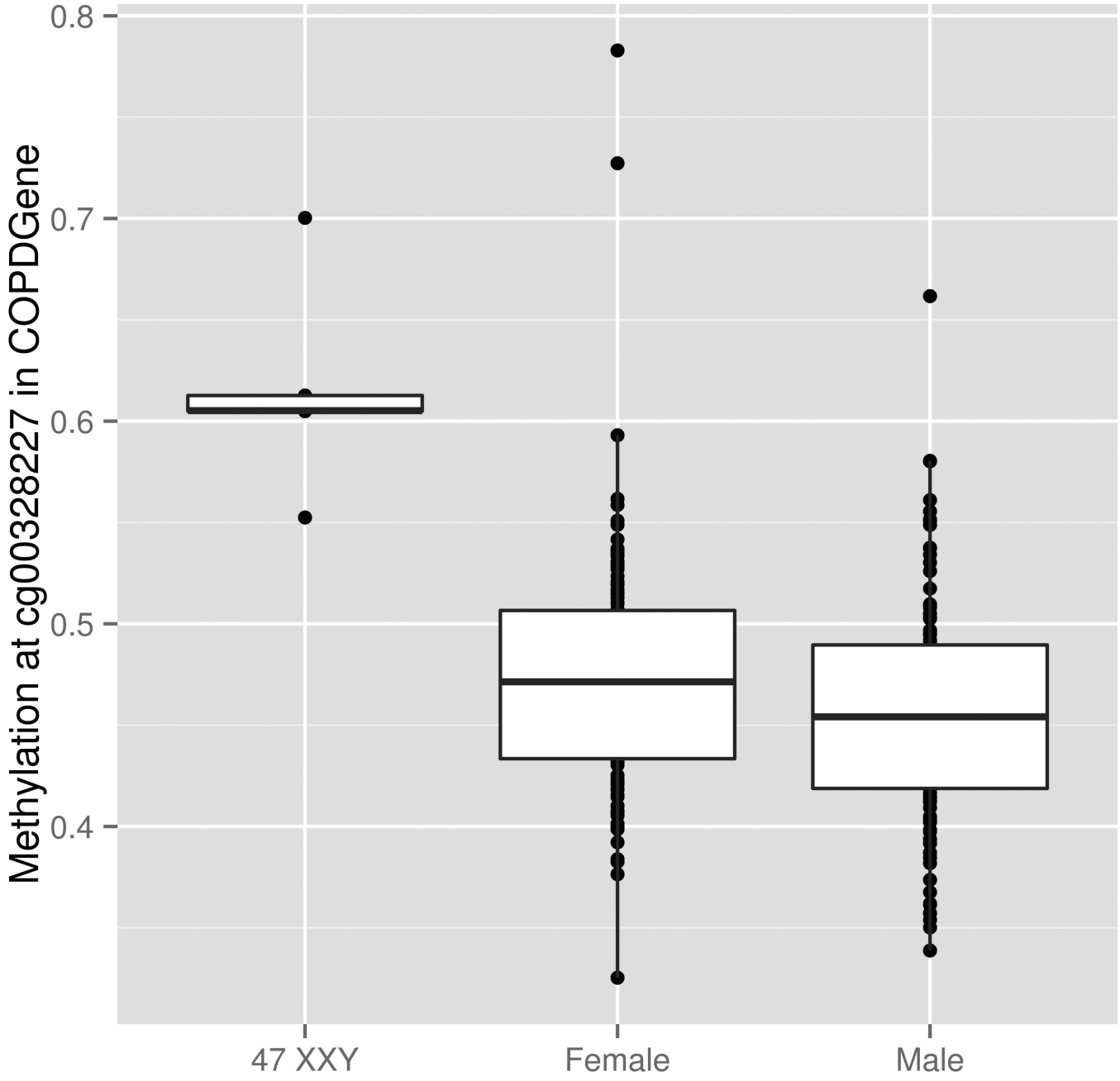
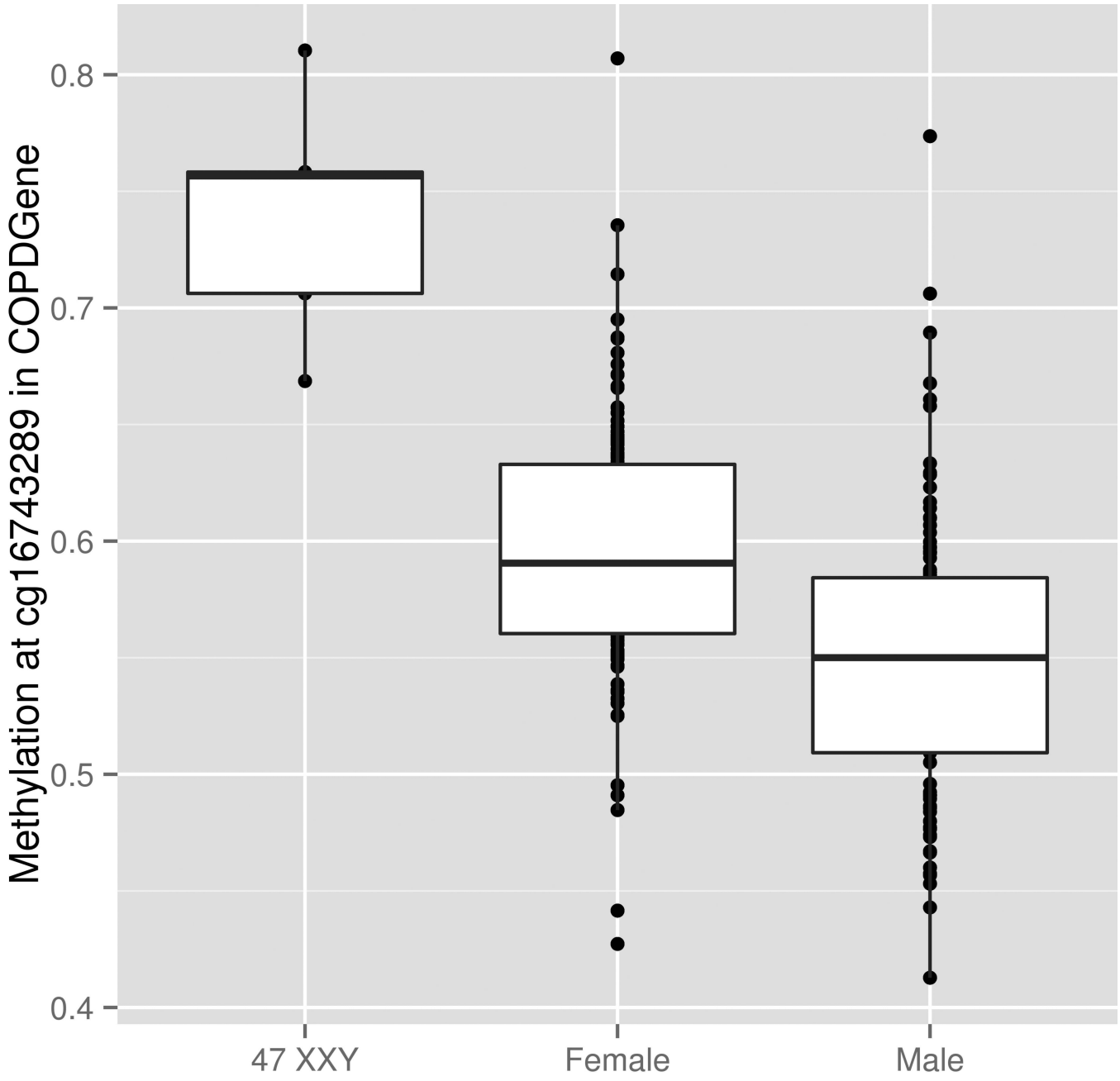


Figure 1.

Venn diagram of overlap between CpG sites significant in the KS versus male (blue) and female (red) controls from COPDGene. CpG sites with an FDR<0.05 are shown.





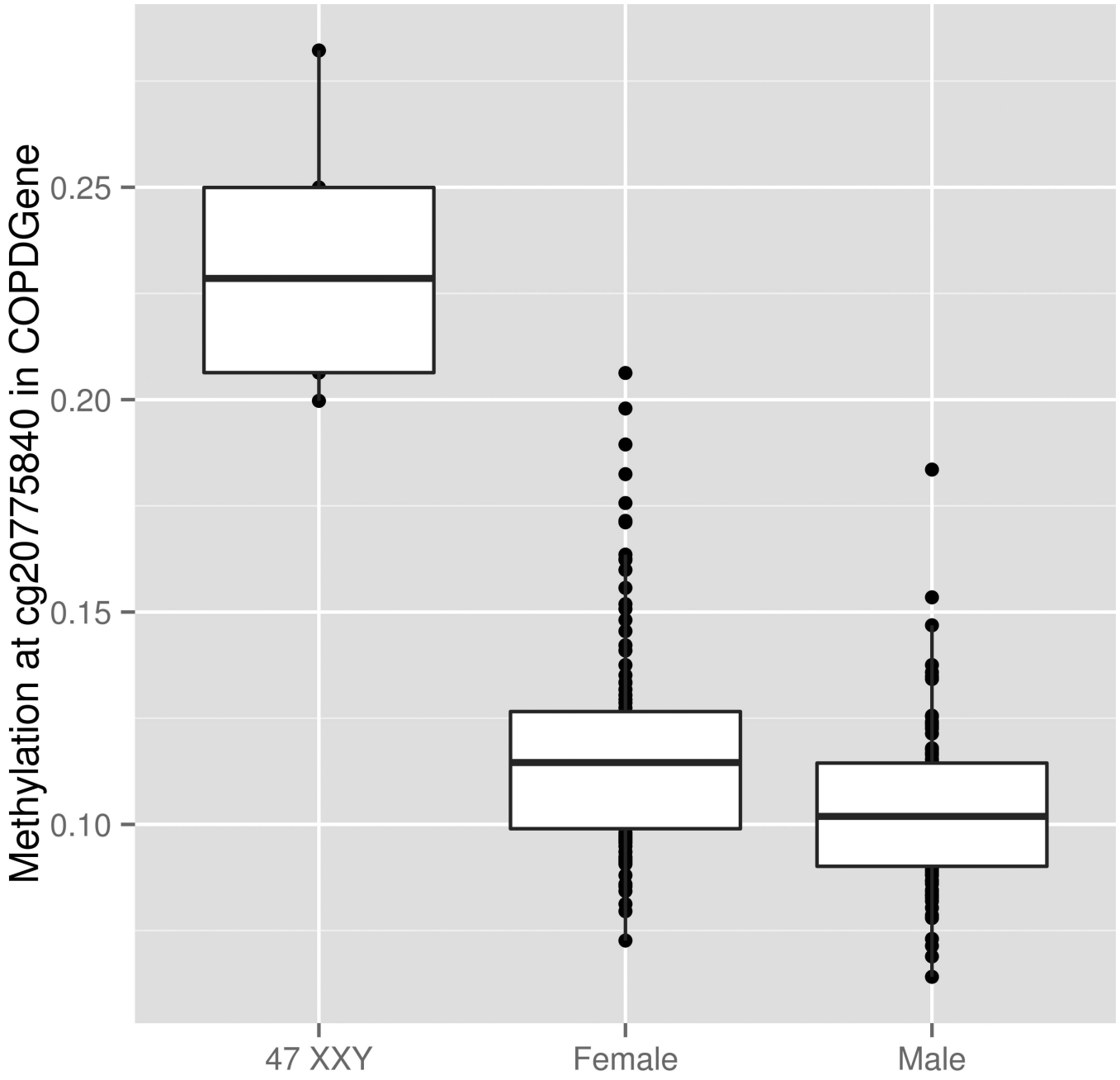


Figure 2.

Mean methylation at replicated CpG sites significant in both KS versus male and KS versus female controls analyses in COPDGene. Mean methylation at (A) cg00328227 (*HENMT1*), (B) cg16743289 (*CACYBP*), and (C) cg20775840 (*G3BP1*) loci in males, females, and KS subjects from COPDGene. At each of the sites, hypermethylation was observed in KS subjects relative to male and female controls.