Received: 21 February 2020 | Revised: 30 April 2020 | Accepted: 1 May 2020

DOI: 10.1002/ajmg.c.31799

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART C Seminars in Medical Genetics | WILEY

Epigenetic and transcriptomic consequences of excess X-chromosome material in 47,XXX syndrome—A comparison with Turner syndrome and 46,XX females

Morten Muhlig Nielsen[1] | Christian Trolle[1,2] | Søren Vang[1] | Henrik Hornshøj[1] | Anne Skakkebæk[1,3] | Jakob Hedegaard[1] | Iver Nordentoft[1] | Jakob Skou Pedersen[1,4] | Claus Højbjerg Gravholt[1,2]

[1]Department of Molecular Medicine, Aarhus University Hospital, Aarhus, Denmark

[2]Department of Endocrinology and Internal Medicine and Medical Research Laboratories, Aarhus University Hospital, Aarhus, Denmark

[3]Department of Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark

[4]Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark

Correspondence
Claus Højbjerg Gravholt, Department of Endocrinology and Internal Medicine and Medical Research Laboratories, Aarhus University Hospital, 8000 Aarhus C, Denmark.
Email: ch.gravholt@dadlnet.dk

Funding information
Aarhus Universitet; Fonden til Lægevidenskabens Fremme; Helga og Peter Kornings Fond; Lundbeckfonden; Novo Nordisk Fonden, Grant/Award Numbers: NNF13OC0003234, NNF15OC0016474

Abstract

47,XXX (triple X) and Turner syndrome (45,X) are sex chromosomal abnormalities with detrimental effects on health with increased mortality and morbidity. In karyotypical normal females, X-chromosome inactivation balances gene expression between sexes and upregulation of the X chromosome in both sexes maintain stoichiometry with the autosomes. In 47,XXX and Turner syndrome a gene dosage imbalance may ensue from increased or decreased expression from the genes that escape X inactivation, as well as from incomplete X chromosome inactivation in 47,XXX. We aim to study genome-wide DNA-methylation and RNA-expression changes can explain phenotypic traits in 47,XXX syndrome. We compare DNA-methylation and RNA-expression data derived from white blood cells of seven women with 47,XXX syndrome, with data from seven female controls, as well as with seven women with Turner syndrome (45,X). To address these questions, we explored genome-wide DNA-methylation and transcriptome data in blood from seven females with 47,XXX syndrome, seven females with Turner syndrome, and seven karyotypically normal females (46,XX). Based on promoter methylation, we describe a demethylation of six X-chromosomal genes (*AMOT*, *HTR2C*, *IL1RAPL2*, *STAG2*, *TCEANC*, *ZNF673*), increased methylation for *GEMIN8*, and four differentially methylated autosomal regions related to four genes (*SPEG*, *MUC4*, *SP6*, and *ZNF492*). We illustrate how these changes seem compensated at the transcriptome level although several genes show differential exon usage. In conclusion, our results suggest an impact of the supernumerary X chromosome in 47,XXX syndrome on the methylation status of selected genes despite an overall comparable expression profile.

KEYWORDS

differential gene expression, DNA-methylation, triple-X, Turner syndrome, X chromosome inactivation, X-chromosome aneuploidies

Abbreviations: DMP, differentially methylated position; DMR, differentially methylated region; TS, Turner syndrome; XCI, X-chromosome inactivation.

1 | INTRODUCTION

47,XXX (triple X) syndrome is a common sex chromosomal abnormality with an incidence of 1 in 1,000 live born girls (Stochholm, Juul, & Gravholt, 2010; Viuff et al., 2015). The syndrome has detrimental effects on health with increased morbidity and mortality (hazard ratio 2.5 and median reduced lifespan of 7.7 years) (Stochholm et al., 2010; Swerdlow, Schoemaker, Higgins, Wright, & Jacobs, 2005). The syndrome may be undiagnosed for years with a median age at diagnosis of 18.2 years due to the lack of pathognomonic traits (Stochholm et al., 2010) and only about 10% of the expected number of 47,XXX are ever diagnosed (Berglund et al., 2019), which means that if more were diagnosed, the age at diagnosis could change. The age at diagnosis in clinic-based studies, typically stemming from pediatric research groups, is of course lower because of ascertainment bias (Bishop et al., 2019; van Rijn & Swaab, 2015; Wigby et al., 2016). 47,XXX newborns are comparable to normal females except for lower birth weight and smaller head circumference (Robinson, Lubs, Nielsen, & Sorensen, 1979). Later, the syndrome manifests with accelerated growth, a lower IQ (70% have an IQ below 90), and an increased incidence of psychiatric (Otter, Schrander-Stumpel, Didden, & Curfs, 2012), language and autism-like disorders (Bishop et al., 2011; van Rijn et al., 2014; van Rijn, Stockmann, van Buggenhout, van Ravenswaaij-Arts, & Swaab, 2014). Evidence suggest an increased prevalence of kyphosis/scoliosis (Buchan et al., 2014; Otter, Schrander-Stumpel, & Curfs, 2010), epilepsy, physical disorders in general (Tartaglia, Howell, Sutherland, Wilson, & Wilson, 2010), premature ovarian failure, and few achieving motherhood (Goswami et al., 2003; Otter et al., 2010; Stochholm et al., 2010). Klinefelter syndrome (47,XXY), another sex chromosome aneuploidy, is also associated with a high degree of comorbidity and psychiatric disorders (Gravholt et al., 2018). In addition, the overrepresentation of X-linked genes as well as the pathogenic role of microduplications in intellectual disability, suggest a possible incomplete inactivation of the supernumerary chromosome X as a causal mechanism. An early study found methylation of the X chromosomes in 45,X and 47,XXX individuals to be profoundly different. They used a method combining specific monoclonal antibodies against 5-methyl cytosine and then hybridization using a microarray to assess genome methylation to 533 genes on the X chromosome (Kelkar & Deobagkar, 2010). This method allowed them to study the entire X chromosome, or at least the genes that were present on their microarray. However, they could not identify single genes, but only genomic regions. Interestingly, they identified motifs with significant similarity to microRNA sequences, to be enriched in methylated regions specific to the inactive X. Interestingly, previous studies using the XIST/FMR-1 methylation ratio as a surrogate measure of X chromosome inactivation (XCI; Mehta et al., 2012) as well as low-resolution DNA-methylation analysis on Klinefelter syndrome (Sharma et al., 2015) support incomplete XCI in X chromosome aneuploidies. Recently, we described genome wide hypomethylation and RNA expression changes in Turner syndrome (45,X) (Trolle et al., 2016), and found a number of genes that may be implicated in the pathogenesis of Turner syndrome, for example associated with congenital urinary malformations (*PRKX*), premature ovarian failure (*KDM6A*), and aortic aneurysm formation (*ZFYVE9* and *TIMP1*) (Trolle et al., 2016), and implicated haploinsufficiency of *TIMP1* (normally expressed from both X chromosomes) combined with the presence of a specific single nucleotide variant in *TIMP1*'s paralog *TIMP3* (on Chromosome 22), with the presence of bicuspid valves and aortopathy (Corbitt et al., 2018).

No studies have so far investigated whether genome-wide analysis of DNA-methylation and RNA-expression changes can explain phenotypic traits in 47,XXX syndrome. To this end, we compared DNA-methylation and RNA-expression data derived from white blood cells of seven women with 47,XXX syndrome, with data from seven female controls, as well as with seven Turner syndrome (45,X). We examine whether DNA methylation and RNA-expression changes also are present in 47,XXX syndrome and discuss how these could relate to the phenotype. In addition, we relate changes in DNA methylation to transcriptomic changes in the context of protein interaction networks.

2 | METHODS

2.1 | Subjects

Seven women with 47,XXX syndrome were recruited from the outpatient clinic at Department of Endocrinology (median age with range: 28.4 [19.0; 43.2]). Seven women with karyotypically proven Turner syndrome (34.0 [27.6; 42.19]) and their age-matched female controls (32.6 [22.0; 43.5]) previously participating in a comprehensive study of the cardiovascular phenotype of Turner syndrome served as controls (Mortensen, Erlandsen, Andersen, & Gravholt, 2013) and data concerning methylation and RNA expression has been published for a larger group of Turner syndrome females before (Trolle et al., 2016). All participants underwent DNA-methylation profiling and gene expression profiling.

2.2 | Sample preparation

2.2.1 | Illumina 450K methylation assay

EDTA-treated peripheral blood samples were stored immediately and until use at −80°C. Genomic DNA was extracted using QIAmp Mini Kit (Qiagen, Germany). For each sample, 1 μg of genomic DNA was bisulfite converted using Zymo EZ DNA methylation Kit. The methylation level was measured using the Infinium HumanMethylation450 Beadchip Kit (Illumina, Inc., San Diego, CA) at Aros Applied Biotechnology A/S (Aarhus, Denmark).

2.2.2 | Illumina 450K microarray data preprocessing

Data were analyzed using the R package *minfi* (Aryee et al., 2014). Detection *p* values were calculated to identify failed positions with a

p-value cut-off >.01. According to these criteria, probes were removed if they failed in more than 20% of the samples (n = 361). Individual positions were removed at the specified cut-off. No samples had a proportion of failed probes exceeding 1% or a median intensity below 11. Raw data were normalized implementing the preprocessing choices of Genome Studio with background normalization and control normalization. Next, we applied subset-quantile-within-array-normalization correcting for technical differences between Infinium Type I and II assay designs allowing both within-array and between-sample normalization (Wu & Aryee, 2010). Cross reactive probes (n = 29,541), probes with SNPs documented in the C or G of the target (n = 18,286), and probes on the sex chromosomes (n = 11,352) were excluded, leaving 415,015 probes. Subsequently, methylation values where calculated as *M* values (logit[beta]) (Du et al., 2010)

$$M-\text{value} = \log 2\left(\frac{\text{Beta}}{1-\text{Beta}}\right)$$

Multidimensional scaling (MDS) plots were evaluated to identify clusters of samples behaving differently than expected. Finally, probes were annotated to the human genome version 19 using the enhanced Illumina annotation developed by Price et al. (2013).

2.2.3 | Estimate differential cell counts

To account for differences in cell composition *Minfi's* estimateCellCounts was used returning the relative proportions of CD4+ and CD8+ T-cells, natural killer cells, monocytes, granulocytes, and B-cells in each sample (Houseman et al., 2012).

2.2.4 | Identifying differentially methylated positions

To identify positions where methylation is associated with the karyotype, we fitted a multiple linear model that utilizes a generalized least squares model and an *F*-test (lmFit of R-package Limma (Smyth, 2005) allowing for missing values. The sample variances were squeezed by computing empirical Bayes posterior means. A Bonferroni adjusted family wise error rate below 0.05 was considered significant. The model was applied without and second with adjustment for the estimated relative cell proportions (CD4+ and CD8+ T-cells, natural killer cells, monocytes, granulocytes, and B-cells) as well as age. Since small changes in *M* values might be of spurious biological significance, we added a delta *M*-value-threshold excluding all differentially methylated positions (DMPs) with a |delta-*M*-value| < 0.3.

2.2.5 | Identifying differentially methylated regions

DMRcate was used to identify differentially methylated regions (DMRs; Peters et al., 2015). DMRcate identifies and ranks the most DMRs across the genome based on kernel smoothing of the differential methylation signal. The model performs well on small sample sizes and builds on the well-established Limma package (Smyth, 2005) allowing us to incorporate estimated cell proportions as covariates. A Benjamini–Hochberg corrected false discovery rate (FDR) <0.05 with a |delta-*M*-value| > 0.3 was considered significant. Following FDR correction, regions were agglomerated from groups of significant probes with a distance of less than 1,000 nucleotides to the next significant probe. Only DMRs with two or more probes are reported.

2.2.6 | X chromosome analysis with respect to DNA methylation

Quality control and normalization were done as for the autosomes. In addition, cross reactive probes (n = 1,201), probes containing SNPs (n = 1,078), and known male hypermethylated genes (of the MAGE, GAGE, and LAGE gene family) were excluded (n = 193) leaving 8,757 probes to map uniquely to the X chromosome. The number of inactivated X chromosomes differs between our three groups while all retain one active X chromosome. Hence, the promoter methylation of the supernumerary inactivated X chromosome in 47,XXX syndrome could be assessed by the formula (beta value of triple-X × 3) – (beta value of female controls × 2) and in female controls as formula (beta value of female controls × 2) – (beta value of Turner syndrome × 1) corresponding to the method previously reported by Cotton et al. (2015)). The beta differences were calculated applying the function pairwiseCI using the package pairwiseCI returning the median Beta value of the supernumerary X chromosome along with a bootstrap (n = 999) based 95% confidence interval. We then used the methylation status of the promoter region to categorize each gene as inactivated (median beta value >30%) or activated (<18%) according to the criteria reported by Cotton et al. (2015)). A significant difference in activation status was determined by nonoverlapping confidence intervals.

2.2.7 | RNA-seq sample preparation, library construction, and sequencing

Blood samples were drawn using RNApax gene tubes and placed 2 hr at room temperature, sequentially stored overnight at −21° before storage at −80°. Whole transcriptome, strand-specific RNA-Seq libraries facilitating multiplexed paired-end sequencing were prepared from total RNA using the Ribo-Zero Globin technology (Epicenter, an Illumina Company) for depletion of rRNA and globin mRNA followed by library preparation using the ScriptSeq technology (Epicenter, an Illumina Company). Depletion and library preparation were automated on a Sciclone NGS (Caliper, PerkinElmer) liquid handling robot. The total RNA (1.7 μg per sample) was subjected to Baseline-ZERO DNase prior to depletion. Total RNA was purified using Agencourt RNAClean XP Beads before and after DNase treatment followed by on-chip electrophoresis on a LabChip GX (Caliper, PerkinElmer) and by UV

measurements on a NanoQuant (Tecan). Cytoplasmic and mitochondrial rRNA as well as globin mRNA were removed from 400 ng DNase treated total RNA using the Ribo-Zero Globin Gold Kit (Human/Mouse/Rat, Epicenter, an Illumina Company) following the manufacturer's instructions and the quality of the depleted RNA was estimated on a LabChip GX (Caliper, PerkinElmer). Synthesis of directional, paired-end, and indexed RNA-Seq libraries were conducted using the ScriptSeq v2 kit (Epicenter, an Illumina Company) following the recommended procedure and the qualities of the RNA-Seq libraries were estimated by on-chip electrophoresis (HS Chip, LabChip GX, Caliper, PerkinElmer) of a 1 μl sample. The DNA concentrations of the libraries were estimated using the KAPA Library Quantification Kit (Kapa Biosystems). The KAPA qPCR reactions were prepared on a Zephyr NGS liquid handling robot (Caliper, PerkinElmer). The RNA-Seq libraries were multiplexed paired-end sequenced on an Illumina NextSeq (75 bp reads).

2.2.8 | RNA-seq analysis

Paired demultiplexed fastq files were generated using bcl2fastq (Illumina) and initial quality control was performed using FastQC. Adapter trimming was conducted using the GATK ReadAdaptorTrimmer tool followed by mapping to the Human genome (hg19). The gene expression was estimated relative to the known set of genes, as provided from the GENCODE project, using Tophat, Bowtie, and Cufflinks (Trapnell et al., 2012). The HTSeq-count software (in union method) was applied to produce raw read counts which were then submitted for analysis in R using edgeR (Robinson, McCarthy, & Smyth, 2010). All non-informative features were removed (n = 5). Filtering was done removing features with less than one counts per million (cpm) in seven samples leaving 16,334 (10,040 coding genes) for downstream analysis (Anders et al., 2013). A generalized linear model was fitted yielding an overall p value. Second, p values and log fold changes (logFC) were retrieved from the individual comparisons of 47,XXX versus 46,XX, 47,XXX versus Turner syndrome, and 46,XX versus Turner syndrome. Genes mapping to the autosomes were retrieved yielding a matrix of 9,820 genes. MDS plots were used to explore the count tables and to assess if samples clustered according to karyotype (Figure 2b,c).

To address the hypothesis that loss of sex-chromosome material manifests itself in differential expression of X-linked genes, crude p values and logFCs were retrieved for genes mapping to the sex chromosomes and FDR p values calculated using Benjamini–Hochberg's equation.

2.2.9 | Differential exon usage

DEXSeq (Anders, Reyes, & Huber, 2012) was used to evaluate differential exon usage and results from the female contrast returned. Exon-usage fold changes were calculated by fitting for each gene, a generalized linear model from the joint data of all its exons. The autosomes and X chromosome were considered as two independent hypotheses. An FDR < 0.05 was considered significant.

2.2.10 | Integration of methylation and expression, correlation, and association with network hotspots

The Functional Epigenetic Modules (FEM) R-package was used to integrate methylation data with expression data, perform differential analysis needed to detect association with protein network hotspots (Jiao, Widschwendter, & Teschendorff, 2014). Functional enrichment test of genes in identified network hotspots was carried out using the topGO R-package (Adrian Alexa and Jorg Rahnenfuhrer [2016], topGO: Enrichment Analysis for Gene Ontology). For methylation, this was done at the gene level using GenStatM and at the gene level with the most significant differentially methylated CpGs among TSS200 and first Exon regions using GenStatMsp. For GenStatMsp, if the criteria were not met, we selected the most significant differentially methylated CpGs among TSS1500.

2.2.11 | Analysis software

Statistical computations were performed using R 3.1.3 (R Foundation for Statistical Computing, Vienna, Austria) with Bioconductor 3.0 (Gentleman et al., 2004). Graphics were made using the basic R functions, RCircos (Zhang, Meltzer, & Davis, 2013), ggbio, Gviz, DEseq (Anders et al., 2013), DEXSeq, and ggplot2 (Wickham, 2009). The package knitr was used for data documentation.

3 | RESULTS

3.1 | Methylation and expression profiling of 47 XXX syndrome, normal females, and Turner syndrome

To identify epigenetic and transcriptomic differences between 47,XXX, control females (46,XX), and females with only one X chromosome (Turner syndrome; 45,X) we performed DNA methylation and gene expression profiling using Illumina 450K microarray and RNA-seq, respectively. The experimental design consisted of seven females in each group (Figure 1a), which were profiled for genome-wide DNA methylation (Figure 1b) and gene expression levels (Figure 1c). The profiling data were analyzed by clustering (Figure 1d), with significant differences between the groups was found (Figure 1e) and detection of network hotspots using integrated methylation and expression data (Figure 1f).

3.2 | Incomplete XCI in 47,XXX syndrome traceable in autosomal epigenome and transcriptome

With previous studies indicating a genome wide regulatory role of sex chromosomal linked genes (Raznahan et al., 2018), we speculated that

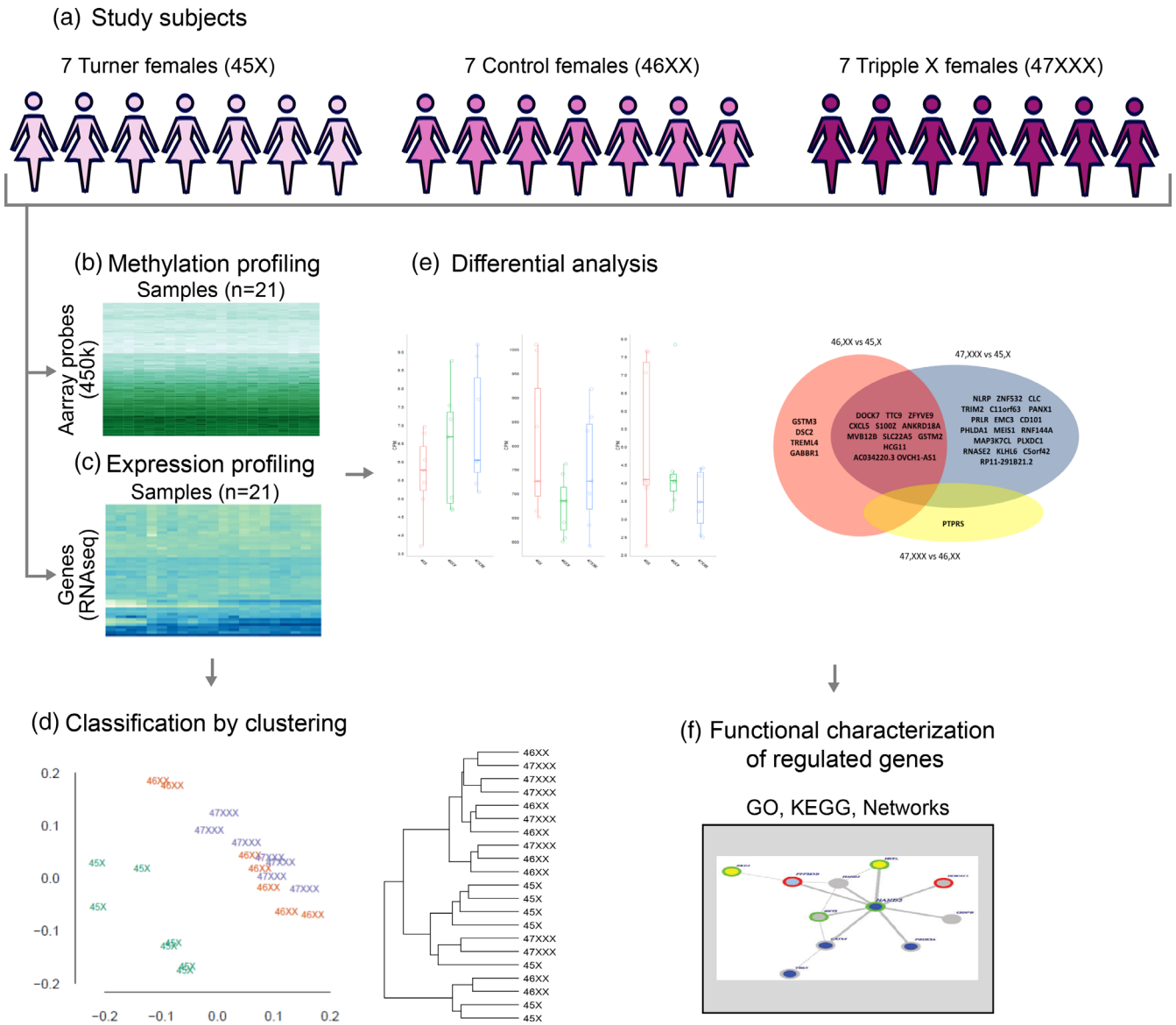

FIGURE 1 An overview of experimental setup and analysis workflow of the presented studies. (a) Twenty-one study subjects; seven Turner females (45,X); seven control females (46,XX); and seven 47,XXX syndrome females. Samples from subjects were profiled for methylation (b) and expression (c) levels. (d) Classification of group samples by clustering. (e) Differential analysis of methylation and expression. (f) Identification of protein network hotspots for genes with methylation-expression correlation

gain or loss of an X chromosome was traceable in the autosomal epigenome and transcriptome of individuals with 47,XXX syndrome. In support, we found segregation of the three groups, 47,XXX, controls, and TS, based on autosomal DNA methylation data (Figure 2a), but not autosomal expression (Figure 2b,c and Supplementary Figure S1). X chromosome expression showed better separation than autosome expression data, with 45,X clearly separated from the two other groups, but with an overlap between controls and 47,XXX.

The X chromosome is enriched for genes involved in intellectual disability and affective disorders (Ji, Higa, Kelsoe, & Zhou, 2015). These findings suggest a link between phenotypes such as lower IQ and psychiatric comorbidity and the supernumerary X chromosome in 47,XXX syndrome. In karyotypically normal females, XCI balances gene expression between sexes and upregulation of the X chromosome in both sexes, of at least some genes, and maintain stoichiometry with the autosomes (Pessia, Engelstadter, & Marais, 2014; Pessia, Makino, Bailly-Bechet, McLysaght, & Marais, 2012). In 47,XXX syndrome, a gene dosage imbalance is possible through at least two mechanisms. First, disruption of gene dosage balance may ensue from an increase in expression from supernumerary escape genes, especially those encoding components of large protein complexes, with disruption of stoichiometry possibly resulting in nonfunctional protein complexes (Pessia et al., 2012). Second, incomplete XCI may result in increased X chromosome gene expression. To address these questions, we explored differences in methylation as a measure of the completeness of XCI. We then compared the X chromosome

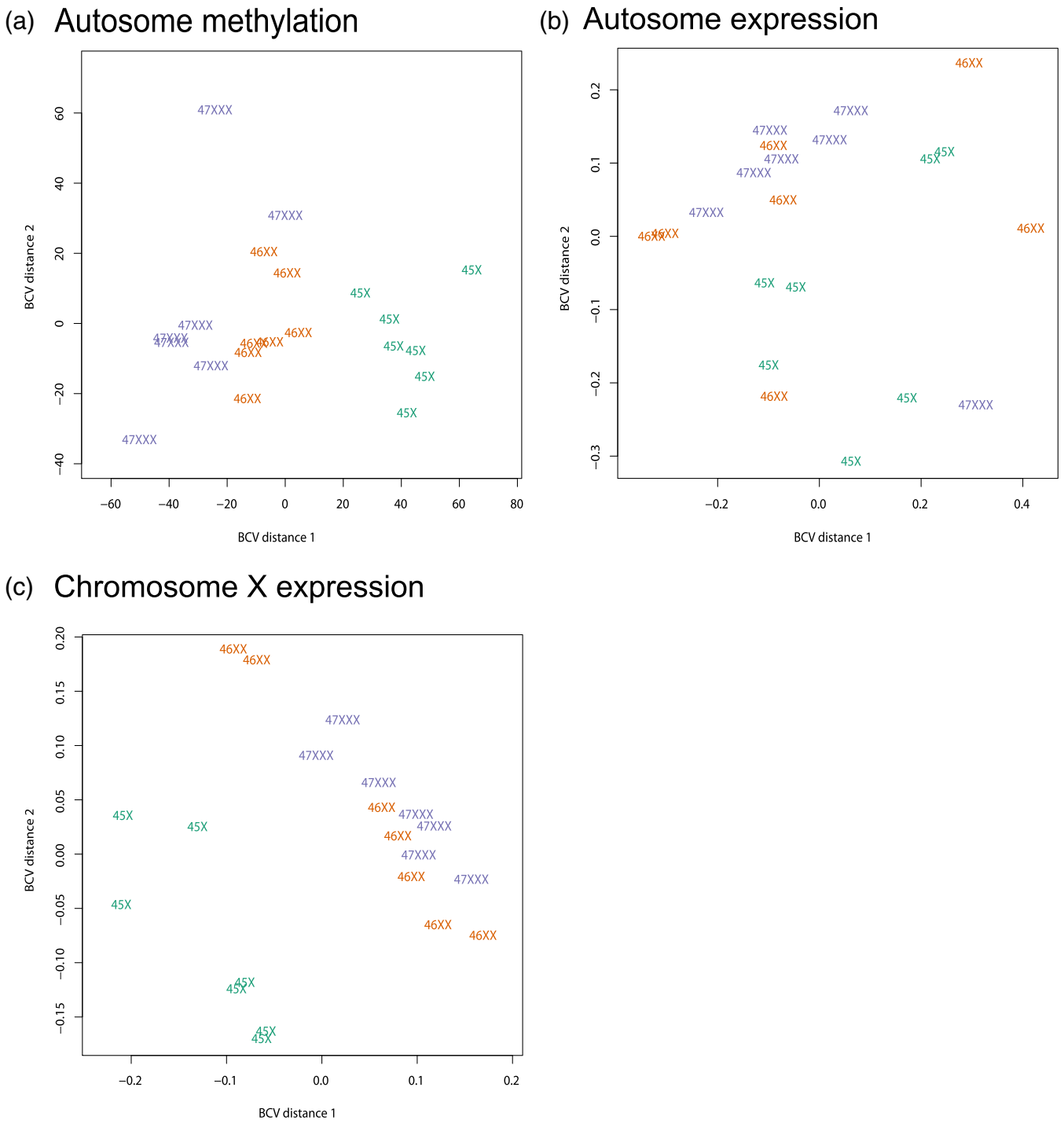

FIGURE 2 Illustrative segregation according to karyotype by multidimensional scaling plots. For DNA methylation results, we chose the 5,000 most variable positions as this yielded the greatest separation, and for RNAseq data, we chose the 500 most variable positions. (a) Five thousand most variable positions for autosomal DNA methylation values. (b) Five hundred most variable autosomal RNAseq expression values. (c) Five hundred most variable X chromosome RNAseq expression values. BCV, biological coefficient of variation

expression profile of 47,XXX syndrome with that of TS as well as karyotypically normal females.

We found six X chromosome genes (*AMOT*, *HTR2C*, *IL1RAPL2*, *STAG2*, *TCEANC*, and *ZNF673*) with a significant change in promoter methylation profile indicating a shift from normally being methylated (and inactivated) in karyotypically normal females (Cotton et al., 2015) (beta value >30%), to much less methylated in 47,XXX syndrome (beta value <18%) (Figure 3). Only one gene (*GEMIN8*) shifted from being demethylated to methylated, with two other genes showing a borderline change (*CDK16* and *SMC1A*) (Supplementary Figure S2). Additionally, six other genes showed an insignificant change in promoter methylation profile (Figure 3). Despite the methylation changes, only

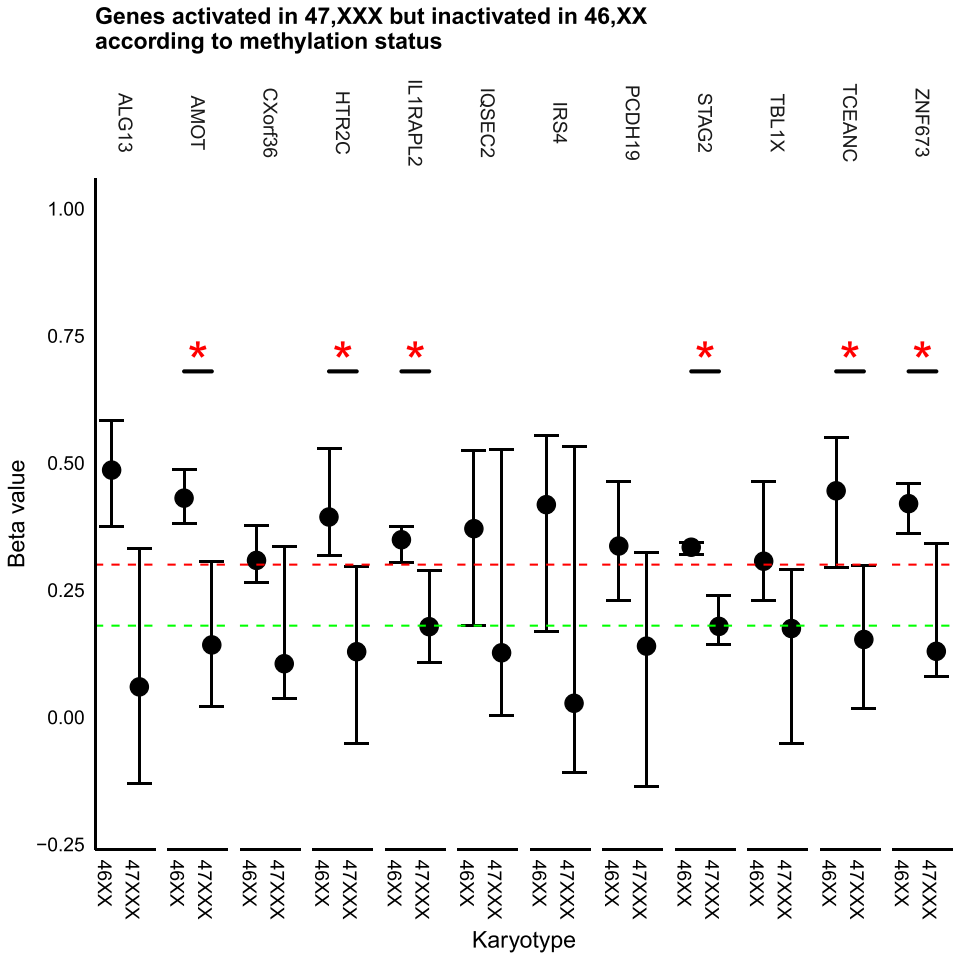

FIGURE 3 Plot of genes differentially methylated in 47,XXX syndrome versus in female controls (46,XX). The plot illustrates the reduced methylation of the supernumerary X chromosome in 47,XXX syndrome. The plot is based on beta values (Y-axis) of CpGs associated with a transcription start site according to Illumina's definition. Only genes with a beta value above 30% (red dashed line) in female controls accompanied by a beta value below 18% (green dashed line) in 47,XXX are plotted. Transcription start sites with a beta value above 30% are known to be associated with inactivated genes and transcription start sites with beta values below 18% to be associated with active genes (Cotton et al., 2015). Dots are medians and spikes are the 95% confidence interval (bootstrap based on 999 resamplings), red asterisks indicate significance. Expression values are provided in Supplementary Figure S2

GEMIN8, *AMOT*, and *STAG2* were expressed in the studied tissue (blood). Interestingly, the change in methylation status of these three genes did not manifest itself at the mRNA level (Supplementary Figure S3) and the genes did not show differential exon usage, implying a gene dosage buffering at a post-transcriptional level, for example, mRNA stability.

As another proxy for XCI, we evaluated the expression of the X-inactive specific transcript (*XIST*) and its activator *JPX*. We found a clear correlation between number of X chromosomes and *XIST* expression, as expected (Figure 4a). Further characterization of expression differences of genes on X chromosome between the three genotypic classes revealed 24 coding, including 6 pseudoautosomal genes and 13 noncoding genes (Figure 4b). When comparing 47,XXX with 46,XX and 45,X females the noncoding genes *XIST*, *JPX*, and *TSIX* showed differential expression with an incremental effect of additional X chromosomes (Figure 4a). *TSIX* is antisense to *XIST* and was only with low, but significant different expression across groups. In mice, *Tsix* prevent accumulation of *Xist* on the future active female X chromosome (Maclary et al., 2014). In humans, the role of *TSIX* is subject to debate and at present, it has no proven role in XCI. Interestingly, of two other genes known to escape X inactivation, one (*IQSEC2*) was not expressed, and the other (*TCEANC*) showed increased methylation, but no effect was seen on expression.

There is enrichment of X-Y gene pairs encoding regulators of transcription, translation, as well as protein stability (Bellott et al., 2014), and aneuploidy is associated with DNA methylation changes possibly changing the status of cryptic splice sites (Carrozza et al., 2005). Hence, we wanted to assess the impact of the supernumerary X chromosome on differential exon usage. We found five genes showing differential exon usage (*TKTL1*, *PIN4*, *VSIG4*, *MAP7D2*, and *DDX3X*). No differential expression or methylation was seen for the four remaining genes. We have previously found differential exon usage of *VSIG4* in Turner syndrome (Trolle et al., 2016), and *VSIG4* seems to be involved in general fetal development and pulmonary function with highest expression in placenta, lung, adrenal gland, heart, and liver. It is also expressed in macrophages and has been suggested to be an inhibitory ligand maintaining T-cell unresponsiveness in healthy tissues (Helmy et al., 2006; Vogt et al., 2006). *DDX3X* has a paralog on the Y chromosome, is ubiquitously expressed and involved in numerous nuclear processes, such as transcriptional regulation, pre-mRNA splicing, and mRNA export, and cytoplasmic processes such as translation, cellular signaling, and viral replication, and mutations in *DDX3X* is also a frequent cause of X-linked mental retardation, almost always affecting females (Snell & Turner, 2018; Snijders et al., 2015).

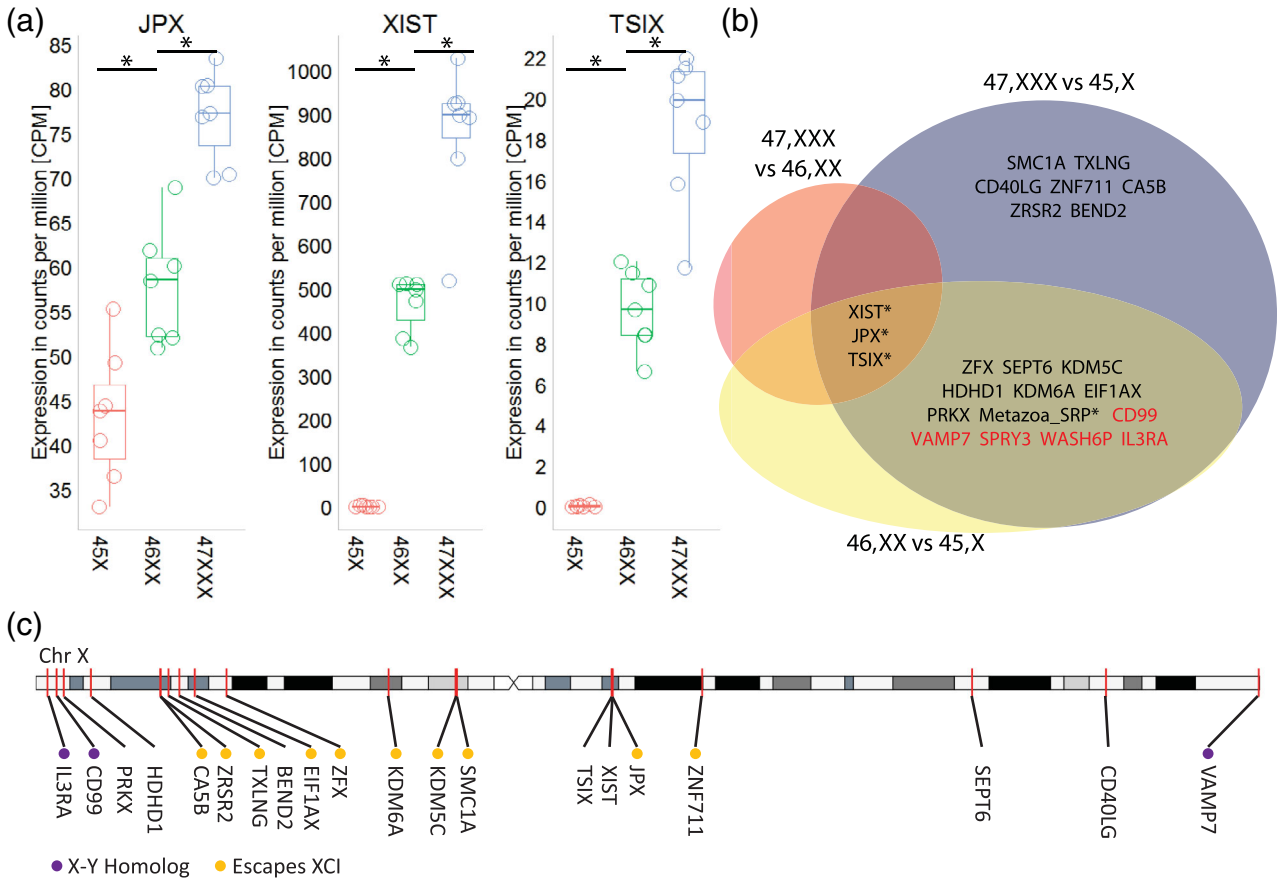

FIGURE 4 X chromosome RNA expression in Turner syndrome (45,X), female controls (46,XX) and 47,XXX syndrome. (a) Boxplot with overlayed dotplot of expression of the *XIST* gene and it's regulator *JPX* as well as it's antisense RNA, *TSIX*. Plot illustrates the increased *XIST*, *JPX*, and *TSIX* expression with additional X chromosomes. (*) indicates *p* value <.05. (b) Venn diagram of the X chromosome illustrating the RNA expression comparing 47,XXX with female controls and the more pronounced differences in the other two comparisons, that is, 45,X versus 46,XXX, and 45,X versus 47,XXX. Genes of the pseudoautosomal region are given in red and * denotes noncoding RNAs. (c) X chromosome ideogram with color coding indicating status of X chromosome genes: X-Y homolog (purple), and Escape gene (yellow) corresponding to the genes differentially expressed only when comparing 47,XXX with 45,X

Thus, 12 genes had a change in status as assessed by promoter methylation, with seven X chromosome genes significantly different in methylation status between 47,XXX and female controls, and in addition, a number of genes on the X chromosome were differentially expressed, including *XIST*, *JPX*, and *TSIX*. In conclusion, we find an overall similar X chromosome expression profile in 47,XXX syndrome compared with female controls, but a change in exon usage, and a shift in the level of methylation for several genes.

3.3 | Impact of X-chromosome aberrations on autosomes

Homologous X-Y gene pairs involved in housekeeping functions has roles in genome-wide control of chromatin modification, transcription, splicing, and translation, and this raises the question if the presence of an extra X chromosome is traceable in the transcriptome of the autosomes (Bellott et al., 2014; Papp, Pal, & Hurst, 2003). To this end, we sought to assess if deviation from the normal amount of X chromosome material changes autosomal coding and noncoding gene expression or result in differential exon usage. Second, we hypothesized that X chromosome aneuploidy would have an effect on epigenetic patterns, possibly traceable in the methylation profile of the autosomes. Similar to our finding of a nearly identical X chromosome expression profile between groups, only two autosomal genes (*PTPRS* and *ANKRD18A*) were differentially expressed between the karyotypic groups. *PTPRS* expression was significantly decreased in 47,XXX syndrome compared to normal females (Figure 5) and *ANKRD18A* expression increased in 47,XXX syndrome compared to 46,XX and 45,X females, despite a comparable methylation profile across the two genes. We next investigated differential exon usage between 47,XXX and 46,XX. Interestingly, we found seven genes (*DOCK7*, *CBLB*, *EGF*, *RP11-2H8.2*, *ALDH1A2*, *CBFA2T3*, and *ULK2*; FDR < 0.05) showing differential exon usage, all on different chromosomes (Supplementary Table S1). The overall similar autosomal expression profile as well as the discrete changes in exon usage may

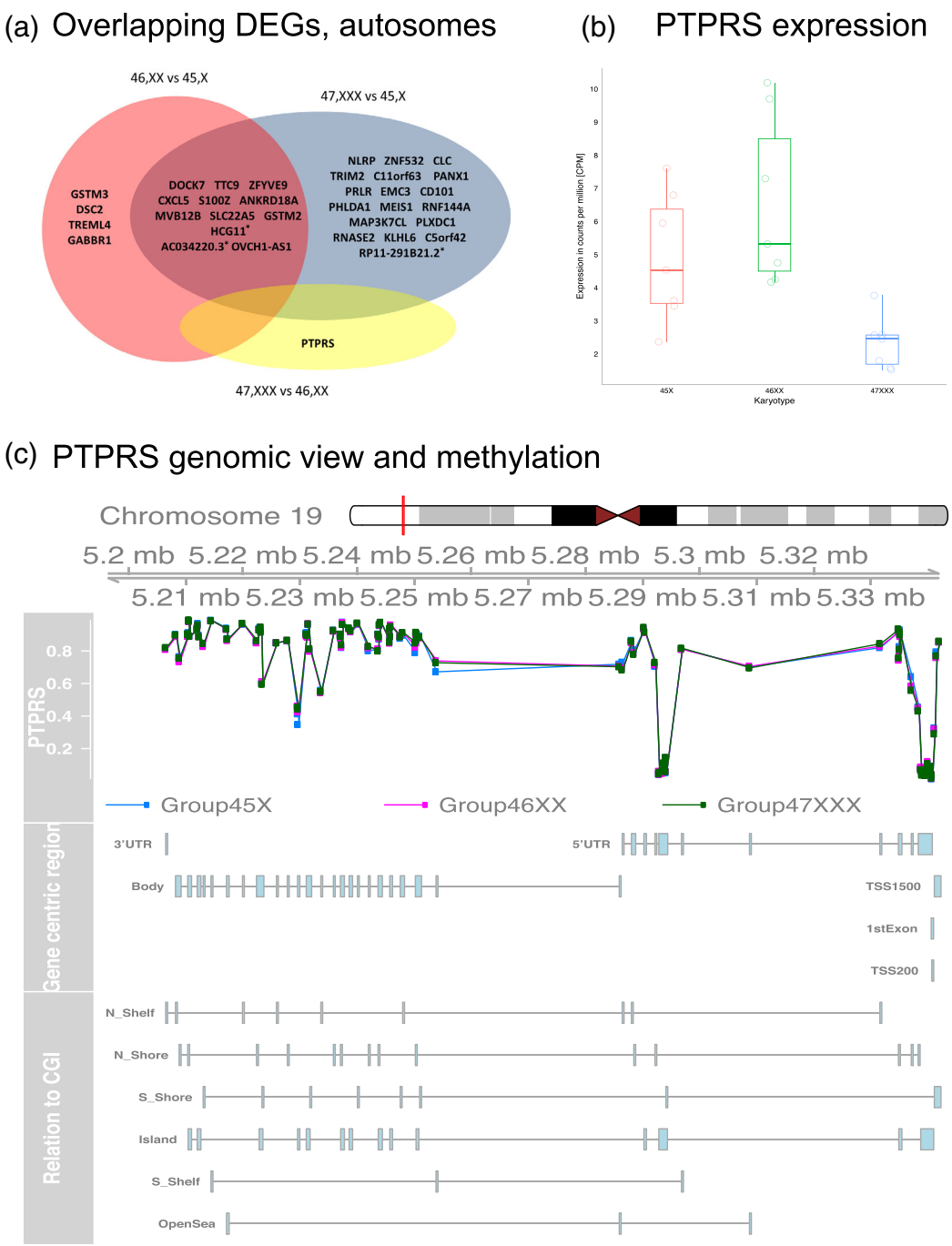

FIGURE 5 Venn diagram of RNA-seq results from the autosomal analysis (* denotes ncRNAs) (a) with only the *PTPRS* gene showing differential expression between female controls (46,XX) and 47,XXX syndrome (b) this despite a comparable methylation profile across the gene (c)

both be the consequence of an adequate inactivation of the supernumerary X chromosome, as well as the result of autosomal epigenetic changes dampening the effect of the extra X chromosome. We then assessed the methylation profile taking a regional perspective by clustering significant probes into regions containing two or more DMPs less than a thousand base pairs apart. We identified four DMRs related to four genes (*SPEG*, *MUC4*, *SP6*, and *ZNF492*) (Supplementary Figure S4). None of these genes were found to be expressed in blood. We note that the methylation and expression changes in 47,XXX syndrome was not as pronounced as methylation and expression changes affecting TS. In TS, the autosome analysis revealed differential expression of 12 coding autosomal genes and 5 noncoding autosomal genes and on the X chromosome, 13 noncoding genes as well as 24 coding genes.

3.4 | Are differentially expressed noncoding RNAs positioned nearby differentially expressed protein coding genes?

Taking into account all three groups, we found differential expression of 13 X chromosome ncRNAs as well as 14 autosomal ncRNAs. The expression of ncRNAs was particular abnormal in TS compared to the two other groups (Figures 4 and 5, and Supplementary Figure S5). The smallest difference was seen between 47,XXX syndrome and 46,XX females, where only *XIST*, *TSIX*, and *JPX* was differentially expressed (Figure 4), while more genes were differentially expressed when only comparing 47,XXX and 45,X females (Figure 4c). Out of the 27 ncRNAs, 18 have no known function. The location of eight of the ncRNAs genes were close to differentially expressed genes (Supplementary Table S2). One ncRNA with unknown function (*G089798*) correlated significantly (Rho = 0.8, *p* value = .048) with a neighboring differentially expressed gene (*SEPT6*). Interestingly, we found differential expression of two pseudogenes—the X-linked translation initiation factor (*EIF1AXP1*) and the X-linked ribosomal protein S4 (*RPS4XP6*) and their corresponding genes (*EIF1AX* and *RPS4X*) when comparing TS with 47,XXX syndrome.

3.5 | Integration of methylation with expression and association with protein network hotspots

To enable an integrated analysis of epigenetic and expression differences between the three analyzed groups, we first matched gene methylation and expression statistics from the three differential analyses (45,X vs. 46,XX; 45,X vs. 47,XXX; and 46,XX vs. 47,XXX) by shared gene IDs. This allowed us to search for interactome hotspots, which are protein–protein interaction (PPI) subnetworks in which a significant number of genes are associated with a phenotype of interest, in this case represented by differential changes among the three groups. The FEM R-package was specifically developed to perform such analysis using integrated 450K methylation and gene expression data (Jiao et al., 2014). We first performed differential analysis of the three contrasts. The described differential methylation statistics was integrated with the differential expression statistics from GenStatR analysis similarly carried out on the three pairwise comparisons. The integrated differential statistics was applied to the FEM tool.

In total, we identified 12 network hotspots corresponding to eight unique seed genes (*AXIN1*, *CD2*, *CXCL5*, *GADD45A*, *NRP1*, signal transducer and activator of transcription 4 [*STAT4*], *ZBTB16*, and *ZFYVE9*; Supplementary Table S4). *STAT4* was found for the 45,X versus 47,XXX contrast with 13 network members ($p = 7.0 \times 10^{-3}$; Figure 6a). The gene members are functionally enriched for T-helper 1 type immune response ($p = 6.0 \times 10^{-9}$; GO-BP; Supplementary Table S5) and interleukin-18 binding ($p = 1.2 \times 10^{-7}$; GO-MF; Supplementary Table S5). The *STAT4* gene show reduced expression in 45,X which correlate with a high methylation level in 3′UTR, for four out of seven 45,X females. *AXIN1* was found for the 46,XX versus 47,XXX contrast with 32 network members ($p = .004$). The gene members are functionally enriched for cell–cell signaling by wnt ($p = 1.8 \times 10^{-16}$; GO-BP; Supplementary Table S5), cell surface receptor signaling

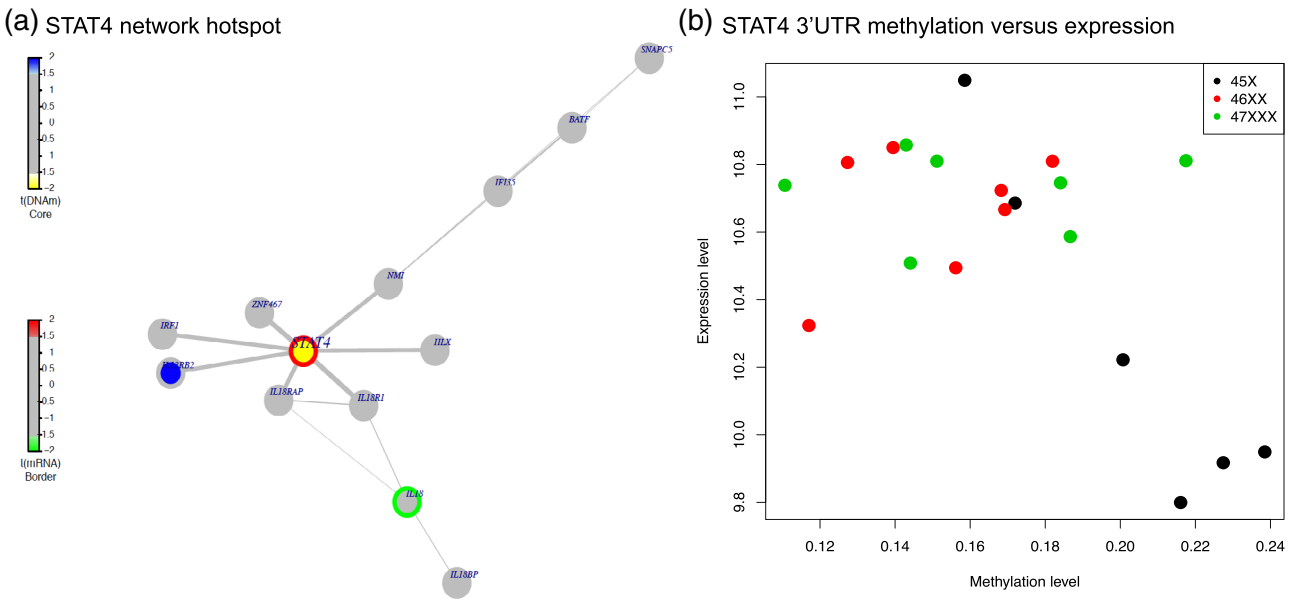

FIGURE 6 Network hotspot and methylation-expression correlation for *STAT4*. (a) Depicted is the functional epigenetic module centered around seed *STAT4*. Edge widths are proportional to the average statistic of the genes making up the edge. Node colors denote the differential DNA methylation statistics as indicated. Border colors denote the differential expression statistics. Note that *STAT4* exhibits the expected anticorrelation with hypomethylation (yellow) leading to overexpression (red). (b) Scatterplot of the methylation levels in 3′ UTR versus expression levels for *STAT4*

pathway ($p = 4.3 \times 10^{-9}$; GO-BP; Supplementary Table S5) and other members of the wnt signaling family.

4 | DISCUSSION

The present study illustrate the more subtle effects of the supernumerary X chromosome in 47,XXX syndrome compared to the effect of the missing X chromosome in TS (Sharma et al., 2015; Trolle et al., 2016). Sex chromosome aneuploidies have been suggested as human models for the study of the sex chromosomes and numerous lines of evidence support a relationship to morbidity and mortality through X-linked genes (Raznahan et al., 2018). The present study explores the impact of an extra X chromosome on the genome-wide DNA methylation and expression profile and investigates the completeness of XCI in 47,XXX syndrome in comparison with karyotypically normal females.

The supernumerary X chromosome is associated with an expected significant increase in *XIST*, *TSIX* and *JPX* expression, indicating that *XIST* inactivate the supernumerary X chromosome. However, methylation data show that despite this increased expression of *XIST*, the inactivation process is incomplete at the single gene level, leading to seven genes (*AMOT*, *HTR2C*, *IL1RAPL2*, *STAG2*, *TCEANC*, *ZNF673*, and *GEMIN8*) with significant promoter methylation differences. This indicates a shift in methylation status from 47,XXX to female controls. *AMOT*, *STAG2*, and *HTR2C* are of particular interest since they participate in multiprotein complexes with the possibility to disrupt stoichiometry resulting in for example improper folding and functioning (Supplementary Table S3). Hence, this shift in promoter methylation status may result in abnormal multiprotein complexes. Interestingly, *STAG2* regulate the separation of sister chromatids during cell division and is involved in the regulation of gene expression. In addition, *STAG2* has previously been implicated in familial skewed X inactivation (Renault, Renault, Copeland, Howell, & Greer, 2011) as well as playing a role in controlling chromosome number (Solomon et al., 2013). *STAG2* participates in a multiprotein complex with *RAD21*, *SMC1A*, as well as *SMC3* and both up and downregulation of *STAG2* may disrupt the function of the protein complex. Interestingly, *SMC1A* is located near one of the differentially expressed ncRNAs in our study (*G088519*). *HTR2C* may explain part of the lower body mass index recorded in 47,XXX syndrome cohorts (Otter et al., 2010), as it plays a role in the regulation of appetite and eating behavior (Heisler et al., 2002). Lastly, *ZNF673* seem to be a factor in X-linked mental retardation and hence a possible factor in the lower IQ associated with 47,XXX syndrome (Lugtenberg et al., 2006). Our expression profiles provided data for three of the seven genes (*GEMIN8*, *AMOT*, and *STAG2*). Despite the methylation results 47,XXX and female controls showed similar expression levels for these three genes. This could suggest a compensatory regulation at a higher level, for example, through mRNA degradation or factors like histone modifications, maintaining the inactivation status in spite of methylation changes (Prestel, Feller, & Becker, 2010).

In keeping with the X chromosome expression analysis, only two autosomal genes (*PTPRS* and *ANKRD18A*) were differentially expressed in white blood cells. Differential expression of *PTPRS* seem to fit well with the 47,XXX phenotype as the gene is thought to play a major role in development and differentiation of the neuroendocrine system (Batt, Asa, Fladd, & Rotin, 2002). *ANKRD18A*, found near the differentially expressed ncRNA, G084562, has previously been reported to show increased promoter methylation and decreased mRNA expression in a study of Tourette's syndrome and OCD (Yu et al., 2015). OCD has been described in TS individuals in particular those with an X ring chromosome (Abd, Patton, Turk, Hoey, & Howlin, 1999).

We speculate that the overall comparable autosomal expression profile could be the result of methylation changes dampening the effect of an extra X chromosome. We identified four DMRs relating to four genes (*SPEG*, *MUC4*, *SP6*, and *ZNF492*). However, these genes were not expressed in the tissue of study here, preventing us from drawing conclusion on the effect of the methylation. Still, we found SPEG of particular interest since it interacts in a protein complex with MTM1, whose gene is located on the X chromosome and change in expression may disrupt stoichiometry (Agrawal et al., 2014).

Evidence accumulates that ncRNAs play roles in transcriptional regulation. ncRNAs are implicated in chromatin modifications, cis and trans gene regulation, RNA–protein interactions, as well as functioning as molecular scaffolds (Guttman & Rinn, 2012). *XIST* (cis-regulator) and *JPX* (trans-regulator) are prime examples of ncRNAs with fundamental roles in XCI. Intriguingly, recent research indicates that ncRNAs may recruit the de novo methyltransferase DNMT3b and hence repress transcription through DNA methylation. As such, ncRNAs may be involved in the perturbed DNA methylation profile of aneuploidy syndromes (Skakkebaek et al., 2018). We found differential expression of several ncRNAs. The relation of these ncRNAs to differentially expressed coding genes suggests that they may either have a regulatory role or that they are incidental byproducts of transcription of coding genes. Only one ncRNA (*G089798*) correlated significantly with expression of its nearby coding gene (*SEPT6*). Interestingly, *SEPT6* plays a role in acute myeloid leukemia with AML previously reported in TS patients (Manola et al., 2008), as well as coarctation of the aorta (Moosmann et al., 2015). We also found that other ncRNAs were differentially expressed and related to genes known to participate in protein–DNA and PPI (*RNF144A*), in the regulation of transcription (*KDM5C*, *EIF2S3*, and *ZFX*), and chromatin remodeling (*KDM5C*), possibly contributing to or alleviating the TS and 47,XXX phenotypes. Lastly, two ncRNAs (*EIF1AXP1* and *RPS4XP6*) which are pseudogenes of differentially expressed genes (*EIF1AX* and *RPS4X*) were differentially expressed, and we speculate that they participate in the regulation of their coding counterpart. *RPS4X* has previously been implicated in Turner syndrome (Fisher et al., 1990; Omoe & Endo, 1996; Rajpathak et al., 2014; Zhang et al., 2013). It encodes the ribosomal protein S4 and transcription is dosage sensitive (Fisher et al., 1990; Just, Geerkens, Held, & Vogel, 1992), and thus it could also be thought to play a role in 47,XXX syndrome.

The fact that a gene is not differentially expressed, does not exclude differential exon usage potentially giving rise to proteins with altered functionality. Hence, we performed whole genome differential exon usage analysis and found five X chromosome genes (*TKTL1*, *PIN4*, *VSIG4*, *MAP7D2*, and *DDX3X*) and seven autosomal genes (*DOCK7*, *CBLB*, *EGF*, *RP11-2H8.2*, *ALDH1A2*, *CBFA2T3*, and *ULK2*) showing differential exon usage. *DDX3X* links to mental retardation (Snijders et al., 2015) and *ULK2* to brain development and hence these two gene are candidates to explore in future studies.

Since the 47,XXX phenotype differs with respect to development both anthropometrically and mentally, we acknowledge tissue specific expression as a likely explanation for the lack of expression of some genes in our study. The differing tissue tolerance to aneuploidy (Fragouli, Wells, & Delhanty, 2011) also adds to this with the brain possibly being more dosage sensitive (Nguyen & Disteche, 2006). We chose leukocytes as tissue of choice due to the difficulty in gaining brain tissue and the invasive methods needed when obtaining muscle, fat, or liver biopsies. Histone modifications may change methylation status as well as regulate transcription directly and it would be interesting to study this in the future.

The network analysis identified *STAT4* as a hotspot seeding gene in the 45,X versus 47,XXX contrast. *STAT4* encodes a transcription factor that is required for development of Th1 cells from naive CD4+ T-cells and IFN-γ production in response to IL-12 (Bacon et al., 1995; Kaplan, 2005). The hotspot represented by *STAT4* has 12 gene members (*STAT4*, *IL18R1*, *IL18RAP*, *IL12RB2*, *NMI*, *IRF1*, *ZNF467*, *HLX*, *IFI35*, *BATF*, *SNAPC5*, *IL18*, and *IL18BP*) that are functionally enriched for "T-helper 1 type immune response" and "interleukin-18-mediated signaling pathway," which is relevant since disturbances in the immune system and autoimmune diseases have been associated with TS. This relation to immune processes was further highlighted by the hotspot represented by CD2 with seven gene members (*CD2*, *CD59*, *CD53*, *CD48*, *CD58*, *CD244*, and *NCR1*), all involved in immune processes. In addition, network analysis identified *AXIN1* as a hotspot seeding gene in the 46,XX versus 47,XXX contrast, with 32 gene members, especially members of the Wnt signaling family, being involved in multiple signal transduction pathways both during embryogenesis and also in cancer development.

There are limitations related to this study, and they include the lack of a validation cohort, and the fact that our study was restricted to DNA methylation and RNA expression in leucocytes from peripheral blood samples, and therefore further studies are needed to establish the impact of DNA methylation and gene expression on the phenotype in 47,XXX in different tissues, such as brain and ovaries. Although the three study groups had similar ages, there was not a complete match and this can of course have affected the methylation analyses, which should be seen as a limitation. In addition, functional validation of the putative genes found here, also needs to be performed. We are confident that such future studies will be able to closely link genotype of 47,XXX individuals with the phenotype and thus possibly explain some or all of the phenotypic traits related to the syndrome. This should provide a fertile seeding ground for the development of new clinical interventions and possibly drugs that can be used in the treatment of individuals with 47,XXX syndrome.

5 | CONCLUSION

In conclusion, we provide evidence of subtle differences in the methylation status, as assessed by DNA-methylation, affecting both coding and noncoding genes possibly involved in the 47,XXX phenotype, although additional functional verification will be necessary. We describe an overall comparable expression profile only differing with respect to the two autosomal gene *PTPRS* and *ANKRD18A* and at the level of differential exon usage.

ACKNOWLEDGMENTS

The authors thank Pamela Celis, Karen Matthiesen, Lone Andersen, Hanne Steen, and Lone Kvist for expert technical assistance. C. H. G. and A. S. are members of the European Reference Network on Rare Endocrine Conditions (ENDO-ERN), Project ID number 739543. This work was supported by a research grant from the Lundbeck Foundation, the Novo Nordisk Foundation (grant agreement NNF13OC0003234 and NNF15OC0016474), the Korning Foundation, "Fonden til lægevidenskabens fremme," and a PhD grant from Aarhus University. The funding bodies had no role in the design of the study, collection, analysis, interpretation of data, or in writing the manuscript.

CONFLICT OF INTEREST

The authors report no competing financial interests.

AUTHOR CONTRIBUTIONS

Claus Højbjerg Gravholt, Henrik Hornshøj, and Christian Trolle conceived the experiments. Søren Vang, Morten Muhlig Nielsen, Henrik Hornshøj, and Christian Trolle conducted the experiments. Søren Vang, Iver Nordentoft, and Jakob Hedegaard conducted the RNA experiments. Søren Vang preprocessed RNA results. Morten Muhlig Nielsen, Henrik Hornshøj, Søren Vang and Christian Trolle analyzed the results. Henrik Hornshøj, Søren Vang, Christian Trolle, Claus Højbjerg Gravholt, and Morten Muhlig Nielsen drafted the manuscript. Morten Muhlig Nielsen, Christian Trolle, Henrik Hornshøj, Anne Skakkebæk, Søren Vang, Jakob Hedegaard, Iver Nordentoft, Jakob Skou Pedersen, and Claus Højbjerg Gravholt reviewed the manuscript.

DATA AVAILABILITY STATEMENT

Sequence data will be deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG. Further information about EGA can be found on https://ega-archive.org.

ETHICS STATEMENT

Informed written consent was obtained. The ethical committee of The Central Denmark Region approved the study (reference number M20110235) and all clinical investigation was conducted according to the principles expressed in the Declaration of Helsinki.

ORCID

Anne Skakkebæk https://orcid.org/0000-0001-9178-4901

Claus Højbjerg Gravholt https://orcid.org/0000-0001-5924-1720

REFERENCES

Abd, S. E., Patton, M. A., Turk, J., Hoey, H., & Howlin, P. (1999). Social, communicational, and behavioral deficits associated with ring X Turner syndrome. *American Journal of Medical Genetics*, *88*, 510–516.

Agrawal, P. B., Pierson, C. R., Joshi, M., Liu, X., Ravenscroft, G., Moghadaszadeh, B., … Beggs, A. H. (2014). SPEG interacts with myotubularin, and its deficiency causes centronuclear myopathy with dilated cardiomyopathy. *American Journal of Human Genetics*, *95*, 218–226.

Anders, S., McCarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., & Robinson, M. D. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature Protocols*, *8*, 1765–1786.

Anders, S., Reyes, A., & Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Research*, *22*, 2008–2017.

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*, 1363–1369.

Bacon, C. M., Petricoin, E. F., III, Ortaldo, J. R., Rees, R. C., Larner, A. C., Johnston, J. A., & O'Shea, J. J. (1995). Interleukin 12 induces tyrosine phosphorylation and activation of STAT4 in human lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 7307–7311.

Batt, J., Asa, S., Fladd, C., & Rotin, D. (2002). Pituitary, pancreatic and gut neuroendocrine defects in protein tyrosine phosphatase-sigma-deficient mice. *Molecular Endocrinology*, *16*, 155–169.

Bellott, D. W., Hughes, J. F., Skaletsky, H., Brown, L. G., Pyntikova, T., Cho, T. J., … Page, D. C. (2014). Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*, *508*, 494–499.

Berglund, A., Viuff, M. H., Skakkebaek, A., Chang, S., Stochholm, K., & Gravholt, C. H. (2019). Changes in the cohort composition of turner syndrome and severe non-diagnosis of Klinefelter, 47,XXX and 47,XYY syndrome: A nationwide cohort study. *Orphanet Journal of Rare Diseases*, *14*, 16–0976.

Bishop, D. V., Jacobs, P. A., Lachlan, K., Wellesley, D., Barnicoat, A., Boyd, P. A., … Scerif, G. (2011). Autism, language and communication in children with sex chromosome trisomies. *Archives of Disease in Childhood*, *96*, 954–959.

Bishop, D. V. M., Brookman-Byrne, A., Gratton, N., Gray, E., Holt, G., Morgan, L., … Thompson, P. A. (2019). Language phenotypes in children with sex chromosome trisomies. *Wellcome Open Research*, *3*, 143.

Buchan, J. G., Alvarado, D. M., Haller, G., Aferol, H., Miller, N. H., Dobbs, M. B., & Gurnett, C. A. (2014). Are copy number variants associated with adolescent idiopathic scoliosis? *Clinical Orthopaedics and Related Research*, *472*, 3216–3225.

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., … Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, *123*, 581–592.

Corbitt, H., Morris, S. A., Gravholt, C. H., Mortensen, K. H., Tippner-Hedges, R., Silberbach, M., & Maslen, C. L. (2018). TIMP3 and TIMP1 are risk genes for bicuspid aortic valve and aortopathy in Turner syndrome. *PLoS Genetics*, *14*, e1007692.

Cotton, A. M., Price, E. M., Jones, M. J., Balaton, B. P., Kobor, M. S., & Brown, C. J. (2015). Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Human Molecular Genetics*, *24*, 1528–1539.

Du, P., Zhang, X., Huang, C. C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, *11*, 587. https://doi.org/10.1186/1471-2105-11-587.:587-11

Fisher, E. M., Beer, R. P., Brown, L. G., Ridley, A., McNeil, J. A., Lawrence, J. B., … Page, D. C. (1990). Homologous ribosomal protein genes on the human X and Y chromosomes: Escape from X inactivation and possible implications for Turner syndrome. *Cell*, *63*, 1205–1218.

Fragouli, E., Wells, D., & Delhanty, J. D. (2011). Chromosome abnormalities in the human oocyte. *Cytogenetic and Genome Research*, *133*, 107–118.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., … Zhang, J. (2004). Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, *5*, R80.

Goswami, R., Goswami, D., Kabra, M., Gupta, N., Dubey, S., & Dadhwal, V. (2003). Prevalence of the triple X syndrome in phenotypically normal women with premature ovarian failure and its association with autoimmune thyroid disorders. *Fertility and Sterility*, *80*, 1052–1054.

Gravholt, C. H., Chang, S., Wallentin, M., Fedder, J., Moore, P., & Skakkebaek, A. (2018). Klinefelter syndrome—Integrating genetics, neuropsychology and endocrinology. *Endocrine Reviews*, *39*, 389–423.

Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature*, *482*, 339–346.

Heisler, L. K., Cowley, M. A., Tecott, L. H., Fan, W., Low, M. J., Smart, J. L., … Elmquist, J. K. (2002). Activation of central melanocortin pathways by fenfluramine. *Science*, *297*, 609–611.

Helmy, K. Y., Katschke, K. J., Jr., Gorgani, N. N., Kljavin, N. M., Elliott, J. M., Diehl, L., … van Lookeren, C. M. (2006). CRIg: A macrophage complement receptor required for phagocytosis of circulating pathogens. *Cell*, *124*, 915–927.

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., … Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, *13*, 86. https://doi.org/10.1186/1471-2105-13-86.:86-13

Ji, B., Higa, K. K., Kelsoe, J. R., & Zhou, X. (2015). Over-expression of XIST, the master gene for X chromosome inactivation, in females with major affective disorders. *eBioMedicine*, *2*, 909–918.

Jiao, Y., Widschwendter, M., & Teschendorff, A. E. (2014). A systems-level integrative framework for genome-wide DNA methylation and gene expression data identifies differential gene expression modules under epigenetic control. *Bioinformatics*, *30*, 2360–2366.

Just, W., Geerkens, C., Held, K. R., & Vogel, W. (1992). Expression of RPS4X in fibroblasts from patients with structural aberrations of the X chromosome. *Human Genetics*, *89*, 240–242.

Kaplan, M. H. (2005). STAT4: A critical regulator of inflammation in vivo. *Immunologic Research*, *31*, 231–242.

Kelkar, A., & Deobagkar, D. (2010). Methylation profile of genes on the human X chromosome. *Epigenetics*, *5*, 612–618.

Lugtenberg, D., Yntema, H. G., Banning, M. J., Oudakker, A. R., Firth, H. V., Willatt, L., … van Bokhoven, H. (2006). ZNF674: A new Kruppel-associated box-containing zinc-finger gene involved in nonsyndromic X-linked mental retardation. *American Journal of Human Genetics*, *78*, 265–278.

Maclary, E., Buttigieg, E., Hinten, M., Gayen, S., Harris, C., Sarkar, M. K., … Kalantry, S. (2014). Differentiation-dependent requirement of Tsix long noncoding RNA in imprinted X-chromosome inactivation. *Nature Communications*, *5*, 4209. https://doi.org/10.1038/ncomms5209.:4209

Manola, K. N., Sambani, C., Karakasis, D., Kalliakosta, G., Harhalakis, N., & Papaioannou, M. (2008). Leukemias associated with Turner syndrome: Report of three cases and review of the literature. *Leukemia Research*, *32*, 481–486.

Mehta, A., Malek-Jones, M., Bolyakov, A., Mielnik, A., Schlegel, P. N., & Paduch, D. A. (2012). Methylation-specific PCR allows for fast diagnosis of X chromosome disomy and reveals skewed inactivation of the X

chromosome in men with Klinefelter syndrome. *Journal of Andrology*, *33*, 955–962.

Moosmann, J., Uebe, S., Dittrich, S., Ruffer, A., Ekici, A. B., & Toka, O. (2015). Novel loci for non-syndromic coarctation of the aorta in sporadic and familial cases. *PLoS One*, *10*, e0126873.

Mortensen, K. H., Erlandsen, M., Andersen, N. H., & Gravholt, C. H. (2013). Prediction of aortic dilation in Turner syndrome - enhancing the use of serial cardiovascular magnetic resonance. *Journal of Cardiovascular Magnetic Resonance*, *15*, 47.

Nguyen, D. K., & Disteche, C. M. (2006). High expression of the mammalian X chromosome in brain. *Brain Research*, *1126*, 46–49.

Omoe, K., & Endo, A. (1996). Relationship between the monosomy X phenotype and Y-linked ribosomal protein S4 (Rps4) in several species of mammals: A molecular evolutionary analysis of Rps4 homologs. *Genomics*, *31*, 44–50.

Otter, M., Schrander-Stumpel, C. T., & Curfs, L. M. (2010). Triple X syndrome: A review of the literature. *European Journal of Human Genetics*, *18*, 265–271.

Otter, M., Schrander-Stumpel, C. T., Didden, R., & Curfs, L. M. (2012). The psychiatric phenotype in triple X syndrome: New hypotheses illustrated in two cases. *Developmental Neurorehabilitation*, *15*, 233–238.

Papp, B., Pal, C., & Hurst, L. D. (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature*, *424*, 194–197.

Pessia, E., Engelstadter, J., & Marais, G. A. (2014). The evolution of X chromosome inactivation in mammals: The demise of Ohno's hypothesis? *Cellular and Molecular Life Sciences*, *71*, 1383–1394.

Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., & Marais, G. A. (2012). Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 5346–5351.

Peters, T. J., Buckley, M. J., Statham, A. L., Pidsley, R., Samaras, K., Lord, V., … Molloy, P. L. (2015). De novo identification of differentially methylated regions in the human genome. *Epigenetics & Chromatin*, *8*, 6. https://doi.org/10.1186/1756-8935-8-6.eCollection@2015.:6-8

Prestel, M., Feller, C., & Becker, P. B. (2010). Dosage compensation and the global re-balancing of aneuploid genomes. *Genome Biology*, *11*, 216–211.

Price, M. E., Cotton, A. M., Lam, L. L., Farre, P., Emberly, E., Brown, C. J., … Kobor, M. S. (2013). Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics & Chromatin*, *6*, 4–6.

Rajpathak, S. N., Vellarikkal, S. K., Patowary, A., Scaria, V., Sivasubbu, S., & Deobagkar, D. D. (2014). Human 45,X fibroblast transcriptome reveals distinct differentially expressed genes including long noncoding RNAs potentially associated with the pathophysiology of Turner syndrome. *PLoS One*, *9*, e100076.

Raznahan, A., Parikshak, N. N., Chandran, V., Blumenthal, J. D., Clasen, L. S., Alexander-Bloch, A. F., … Geschwind, D. H. (2018). Sex-chromosome dosage effects on gene expression in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 7398–7403.

Renault, N. K., Renault, M. P., Copeland, E., Howell, R. E., & Greer, W. L. (2011). Familial skewed X-chromosome inactivation linked to a component of the cohesin complex, SA2. *Journal of Human Genetics*, *56*, 390–397.

Robinson, A., Lubs, H. A., Nielsen, J., & Sorensen, K. (1979). Summary of clinical findings: Profiles of children with 47,XXY, 47,XXX and 47,XYY karyotypes. *Birth Defects Original Article Series*, *15*, 261–266.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*, 139–140.

Sharma, A., Jamil, M. A., Nuesgen, N., Schreiner, F., Priebe, L., Hoffmann, P., … El-Maarri, O. (2015). DNA methylation signature in peripheral blood reveals distinct characteristics of human X chromosome numerical aberrations. *Clinical Epigenetics*, *7*, 76.

Skakkebaek, A., Nielsen, M. M., Trolle, C., Vang, S., Hornshoj, H., Hedegaard, J., … Gravholt, C. H. (2018). DNA hypermethylation and differential gene expression associated with Klinefelter syndrome. *Scientific Reports*, *8*, 13740–31780.

Smyth, G. K. (2005). Limma: linear models for microarray data. In V. C. Gentleman, R. Dudoit, R. Irizarry, & W. Huber (Eds.), *Bioinformatics and computational biology solutions using R and bioconductor*. New York, NY: Springer.

Snell, D. M., & Turner, J. M. A. (2018). Sex chromosome effects on male-female differences in mammals. *Current Biology*, *28*, R1313–R1324.

Snijders, B. L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R., … Kleefstra, T. (2015). Mutations in DDX3X are a common cause of unexplained intellectual disability with gender-specific effects on Wnt signaling. *American Journal of Human Genetics*, *97*, 343–352.

Solomon, D. A., Kim, J. S., Bondaruk, J., Shariat, S. F., Wang, Z. F., Elkahloun, A. G., … Waldman, T. (2013). Frequent truncating mutations of STAG2 in bladder cancer. *Nature Genetics*, *45*, 1428–1430.

Stochholm, K., Juul, S., & Gravholt, C. H. (2010). Mortality and incidence in women with 47,XXX and variants. *American Journal of Medical Genetics. Part A*, *152A*, 367–372.

Swerdlow, A. J., Schoemaker, M. J., Higgins, C. D., Wright, A. F., & Jacobs, P. A. (2005). Mortality and cancer incidence in women with extra X chromosomes: A cohort study in Britain. *Human Genetics*, *118*, 255–260.

Tartaglia, N. R., Howell, S., Sutherland, A., Wilson, R., & Wilson, L. (2010). A review of trisomy X (47,XXX). *Orphanet Journal of Rare Diseases*, *5*, 8.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., … Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*, 562–578.

Trolle, C., Nielsen, M. M., Skakkebaek, A., Lamy, P., Vang, S., Hedegaard, J., … Gravholt, C. H. (2016). Widespread DNA hypomethylation and differential gene expression in Turner syndrome. *Scientific Reports*, *6*, 34220.

van Rijn, S., Stockmann, L., Borghgraef, M., Bruining, H., van Ravenswaaij-Arts, C., Govaerts, L., … Swaab, H. (2014). The social behavioral phenotype in boys and girls with an extra X chromosome (Klinefelter syndrome and trisomy X): A comparison with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*, 310–320.

van Rijn, S., Stockmann, L., van Buggenhout, G., van Ravenswaaij-Arts, C., & Swaab, H. (2014). Social cognition and underlying cognitive mechanisms in children with an extra X chromosome: A comparison with autism spectrum disorder. *Genes, Brain, and Behavior*, *13*(5), 459–467.

van Rijn, S., & Swaab, H. (2015). Executive dysfunction and the relation with behavioral problems in children with 47,XXY and 47,XXX. *Genes, Brain, and Behavior*, *14*, 200–208.

Viuff, M. H., Stochholm, K., Uldbjerg, N., Nielsen, B. B., the Danish Fetal Medicine Study Group, & Gravholt, C. H. (2015). Only a minority of sex chromosome abnormalities are detected by the Danish national prenatal screening program for Down syndrome. *Human Reproduction*, *30*, 2419–2426.

Vogt, L., Schmitz, N., Kurrer, M. O., Bauer, M., Hinton, H. I., Behnke, S., … Bachmann, M. F. (2006). VSIG4, a B7 family-related protein, is a negative regulator of T cell activation. *The Journal of Clinical Investigation*, *116*, 2817–2826.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*, 2 ed., (1–253). New York, NY: Springer-Verlag.

Wigby, K., D'Epagnier, C., Howell, S., Reicks, A., Wilson, R., Cordeiro, L., & Tartaglia, N. (2016). Expanding the phenotype of triple X syndrome: A comparison of prenatal versus postnatal diagnosis. *American Journal of Medical Genetics. Part A*, *170*, 2870–2881.

Wu, Z., & Aryee, M. J. (2010). Subset quantile normalization using negative control features. *Journal of Computational Biology*, *17*, 1385–1395.

Yu, D., Mathews, C. A., Scharf, J. M., Neale, B. M., Davis, L. K., Gamazon, E. R., … Pauls, D. L. (2015). Cross-disorder genome-wide analyses suggest a complex genetic relationship between Tourette's syndrome and OCD. *The American Journal of Psychiatry*, *172*, 82–93.

Zhang, H., Meltzer, P., & Davis, S. (2013). RCircos: An R package for Circos 2D track plots. *BMC Bioinformatics*, *14*, 244.

Zhang, R., Hao, L., Wang, L., Chen, M., Li, W., Li, R., … Wu, J. (2013). Gene expression analysis of induced pluripotent stem cells from aneuploid chromosomal syndromes. *BMC Genomics*, *14*(Suppl 5), S8. https://doi.org/10.1186/1471-2164-14-S5-S8

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Nielsen MM, Trolle C, Vang S, et al. Epigenetic and transcriptomic consequences of excess X-chromosome material in 47,XXX syndrome—A comparison with Turner syndrome and 46,XX females. *Am J Med Genet Part C*. 2020;1–15. https://doi.org/10.1002/ajmg.c.31799