ORIGINAL INVESTIGATION

Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation

Allison M. Cotton · Lucia Lam · Joslynn G. Affleck · Ian M. Wilson ·
Maria S. Peñaherrera · Deborah E. McFadden · Michael S. Kobor ·
Wan L. Lam · Wendy P. Robinson · Carolyn J. Brown

Abstract X-chromosome inactivation (XCI) results in the differential marking of the active and inactive X with epigenetic modifications including DNA methylation. Consistent with the previous studies showing that CpG island-containing promoters of genes subject to XCI are approximately 50% methylated in females and unmethylated in males while genes which escape XCI are unmethylated in both sexes; our chromosome-wide (Methylated DNA ImmunoPrecipitation) and promoter-targeted methylation analyses (Illumina Infinium HumanMethylation27 array) showed the largest methylation difference ($D = 0.12$, $p < 2.2$ E−16) between male and female blood at X-linked CpG islands promoters. We used the methylation differences between males and females to predict XCI statuses in blood and found that 81% had the same XCI status as previously determined using expression data. Most genes (83%) showed the same XCI status across tissues (blood, fetal: muscle, kidney and nerual); however, the methylation of a subset of genes predicted different XCI statuses in different tissues. Using previously published expression data the effect of transcription on gene-body methylation was investigated and while X-linked introns of highly expressed genes were more methylated than the introns of lowly expressed genes, exonic methylation did not differ based on expression level. We conclude that the XCI status predicted using methylation of X-linked promoters with CpG islands was usually the same as determined by expression analysis and that 12% of X-linked genes examined show tissue-specific XCI whereby a gene has a different XCI status in at least one of the four tissues examined.

A. M. Cotton · L. Lam · J. G. Affleck · M. S. Peñaherrera ·
M. S. Kobor · W. P. Robinson · C. J. Brown (✉)
Department of Medical Genetics, University of British
Columbia, Vancouver, BC, Canada
e-mail: cbrown@interchange.ubc.ca

A. M. Cotton · J. G. Affleck · C. J. Brown
Molecular Epigenetics Group, Life Sciences Institute,
Vancouver, BC, Canada

I. M. Wilson · W. L. Lam
British Columbia Cancer Research Centre,
Vancouver, BC, Canada

I. M. Wilson · W. L. Lam
Pathology and Laboratory Medicine,
University of British Columbia,
Vancouver, BC, Canada

L. Lam · M. S. Peñaherrera · D. E. McFadden ·
M. S. Kobor · W. P. Robinson
Child and Family Research Institute,
Vancouver, BC, Canada

L. Lam · M. S. Kobor
Centre for Molecular Medicine and Therapeutics,
Vancouver, BC, Canada

Introduction

To achieve dosage compensation with 46,XY males, 46,XX mammalian females undergo X-chromosome inactivation (XCI), which results in one inactive X chromosome (Xi) and one active X chromosome (Xa). Based predominantly on studies using somatic cell hybrids the majority of X-linked genes have been determined to be subject to XCI and are only expressed from the Xa, while approximately 15% of genes escape XCI and are expressed

from both the Xa and the Xi (Carrel and Willard 2005). Genes which escape XCI are enriched in the more evolutionarily recently diverged regions of the short arm of the X chromosome, including all of the genes examined in the Xp pseudoautosomal region; however, additional escapees are distributed throughout the X chromosome (Carrel and Willard 2005). The overall differences in expression between the Xa and the Xi are reflected in the association of the Xi with heterochromatic histone marks as well as a lack of euchromatic histone marks (reviewed in Chow and Heard 2009). The facultative heterochromatin of the X chromosome provides an excellent system to study epigenetic silencing, and one of the first epigenetic marks proposed to play a role in XCI was DNA methylation (Riggs 1975).

In somatic cells, DNA methylation is found almost exclusively at CpG dinucleotides (Lister et al. 2009), which are underrepresented across the human genome but are enriched at the promoters of 60% of genes, resulting in regions known as CpG islands (Bird 1980; Weber et al. 2007). Definitions of what constitutes a CpG island differ (Gardiner-Garden and Frommer 1987; Takai and Jones 2002); however, the use of three levels of CpG density: high CpG density (HC), intermediate CpG density (IC) and low CpG density (LC) allows for the unique properties of CpG islands to be dissected (Weber et al. 2007). While autosomal promoters associated with CpG islands are typically unmethylated in all tissues, a subset of autosomal promoters has been found to show tissue-specific methylation. The regions surrounding CpG islands, named CpG island shores, show the largest methylation differences between tissues (Eckhardt et al. 2006; Irizarry et al. 2009; Illingworth et al. 2008; Schilling and Rehli 2007). Genes on the X chromosome that are subject to XCI typically have CpG islands that are unmethylated in males, but show partial methylation in females; reflecting that CpG island promoters on the Xa are unmethylated, similar to autosomal CpG island promoters, while CpG island promoters on the Xi are methylated (Cotton et al. 2009; Wolf et al. 1984a, b). Genes with CpG islands that escape XCI appear to be unmethylated on both the Xa and Xi (Goodfellow et al. 1988). The consistent relationship between X-linked CpG island promoter methylation and XCI status has been found in multiple individual gene studies (Hansen and Gartler 1990; Carrel et al. 1996; Anderson and Brown 2002) and a study of neutrophils to propose novel genes which escape XCI (Yasukochi et al. 2010). Multiple lines of evidence support that methylation is important to maintain XCI including the reactivation of human X-linked genes upon treatment with DNA methyltransferase inhibitors (Venolia et al. 1982) and the high reactivation frequency for X-linked genes from the hypomethylated marsupial Xi

during cell culture (Loebel and Johnston 1996; Kaslow and Migeon 1987; Rens et al. 2010).

Approximately 50% of CpG islands overlap a known transcription start site (TSS), with the remaining 50% of CpG islands, called orphan CpG islands, being distributed throughout the intragenic and the intergenic regions of the genome (Illingworth et al. 2008, 2010). These orphan CpG islands are often associated with promoter histone modifications and RNA Pol II, suggesting that at least 40–60% of orphan CpG islands are acting as promoters for currently unknown genes (Illingworth et al. 2010). While the methylation status of CpG island promoters has been well examined, fewer studies have examined non-promoter regions. On the X chromosome the use of in situ nick translation suggested that overall the Xi is hypermethylated compared to the Xa; whereas the Xi is reported to be hypomethylated compared with the Xa using methylation-sensitive restriction enzymes (Prantera and Ferraro 1990; Viegas-Pequignot et al. 1988). It appears that X-linked gene bodies are hypermethylated on the Xa compared to the Xi (Cotton et al. 2009; Weber et al. 2005; Hellman and Chess 2007) and it has been proposed that this is due to gene transcription on the Xa (Jones 1999). The link between transcription and gene body methylation has not just been detected on the X chromosome but across the genome with highly expressed genes showing more gene body methylation than genes with low expression (Aran et al. 2011). Within genes, the 5′ most exons show methylation which, like CpG island promoters, correlates with gene silencing, while both internal exons and introns show variable methylation (Brenet et al. 2011; Edwards et al. 2010). CpG island promoters only represent approximately 1% of the DNA of the X chromosome; therefore, the study of methylation across other regions is important to gain a complete picture of X chromosome-wide methylation levels.

To study the distribution of X-linked methylation we used the Illumina Infinium HumanMethylation27 array and analysed methylation at 777 X-linked promoters in human blood and fetal somatic tissues (muscle, brain, spinal cord and kidney). Methylated DNA ImmunoPrecipitation (MeDIP) and hybridization to a NimbleGen 2.1M array allowed for chromosome-wide methylation analysis in human blood. Examination of methylation of X-linked non-promoter regions revealed methylation differences between the introns, but not exons, of highly versus lowly expressed genes in both males and females. Our analysis of X-linked promoters confirmed that promoter methylation differences between males and females are found primarily at CpG islands on the X chromosome and not on the autosomes. As has been previously shown on a single gene level (Hansen and Gartler 1990; Carrel et al. 1996; Anderson and Brown 2002), we demonstrated that

X-linked promoter methylation correlates well with the XCI status of genes with CpG islands which in turn suggests that the majority of unannotated CpG islands on the X chromosome may be the promoters of genes subject to XCI. Translating X-linked promoter methylation into a genic XCI status across different tissues suggests that 12% of genes show tissue-specific XCI in which the gene is predicted to be subject to XCI in at least one tissue, but also predicted to escape XCI in at least one tissue.

Materials and methods

Sample collection and DNA extraction

Collection of samples was approved by the ethics committees of the University of British Columbia and the Children's and Women's Health Centre of British Columbia. Whole blood samples (female $n = 59$ and male $n = 36$) were collected from anonymous donors (ethics approval number H08-02773) and peripheral blood mononuclear cells (PBMCs) isolated using BD Cell Preparation Tubes as per manufacturer's instructions. DNA was extracted using Qiagen AllPrep DNA/RNA mini kits as per standard conditions. Fetal tissues (muscle: female $n = 6$ and male $n = 4$, spinal cord: female $n = 2$ and male $n = 1$, brain: female $n = 4$ and male $n = 4$, kidney: female $n = 6$ and male $n = 5$) were chromosomally normal and collected from biopsied abortuses from anonymous pregnancies terminated for medical reasons (ethics approval number H06-70085). Genomic DNA was extracted using a standard salting-out method as outlined in Papageorgiou et al. (2009).

Illumina Infinium HumanMethylation27 array

Genomic DNA was bisulfite modified with the EZ DNA Methylation Kit (Zymo Research) as per the manufacturer's instructions and 180–200 ng of bisulfite DNA was then amplified, fragmented and hybridized to Illumina Infinium HumanMethylation27 beadarray chips (Illumina, Inc) using Illumina supplied reagents and conditions. The arrays were scanned on the Illumina iScan system and imported into GenomeStudio for further analysis (2010.2). Results were subjected to a background normalization using BeadStudio (versions 3.1.3.0 Illumina, Inc). Quantile normalization was performed in R 2.11.0 using the limma package (Bolstad et al. 2003). Although beta-values are compressed when less than 0.2 and greater than 0.8 and both of these ranges show high heteroscedasticity, since we were interested in large methylation differences between males and females for the purposes of this paper beta-value was considered equivalent to percent methylation (Du et al. 2010).

CpG density definitions

We used CpG density classifications based on those used by Weber et al. (2007) to define three CpG densities: HC, IC and LC. The program CpGIE (Wang and Leung 2004) was used to define and locate HC and IC islands on the X chromosome, chromosomes 20, 21 and 22. HCs had a GC content greater than 55%, an observedCpG/expectedCpG greater than 0.75 and were at least 500 bp in length. ICs had a GC content greater than 50%, an observedCpG/expectedCpG greater than 0.48 and were at least 200 bp in length. Those ICs which overlapped with an HC were excluded from the IC category but their HC component remained in the HC category. In addition, all ICs which overlapped with repetitive elements, as defined by RepeatMasker (Fujita et al. 2011; Smit et al. 1996–2010), were not included in the IC category. LCs were all those regions which were not HC or IC.

Illumina Infinium HumanMethylation27 composition and probes removed from analysis

The Illumina Infinium HumanMethylation27 array is a promoter array with all probes located in close proximity to an annotated TSS. Approximately 45% of X-linked promoters are represented on the array and those promoters which overlap CpG islands (HC and IC) represent nearly three quarters of the probes on the array despite the fact that only 5% of CpGs on the X chromosome are located in islands. The BLAST program (Altschul et al. 1990) from NCBI was used to determine whether a probe sequence mapped to a single unique location in the genome or to multiple sites. Due to the large number of genes on the X chromosome which have homologs on the Y chromosome, probes which mapped to the Y chromosome as well the X chromosome were not removed from the analysis. 153 X-linked and 134 autosomal (chr 20, 21 and 22) probes were removed from the analysis due to mapping to more than one location in the genome. 137 X-linked probes located in the promoters of the cancer-testis (CT) family of genes were removed from the analysis since they are known to be methylated in all tissues except testis regardless of CpG density (De Smet et al. 1999). To determine if probes were located in repetitive elements, probe locations were compared against the location of known repetitive elements from RepeatMasker for UCSC (Fujita et al. 2011; Smit et al. 1996–2010) which resulted in the removal of 88 X-linked and 220 autosomal probes.

Statistical analysis

Statistical analysis of the Illumina Infinium HumanMethylation27 array was performed using the Mann–Whitney

test as calculated by Graphpad Prism. Statistical analysis of MeDIP data was calculated in R (Team 2010) using the Kolmogorov–Smirnov Test. Intrasex variation was calculated for each sex by comparing all combinations of samples using the Kolmogorov–Smirnov test. Due to the large sample size, p values < 0.0001 were considered significant and p values < 1.0 E-10 highly significant. In order for the results to be considered significant and to ensure the difference between the average male and average female methylation was larger than any differences within the sexes, we required that the p value resulting from the comparison of the average male and average female methylation was smaller than the intrasex p values.

Decision tree to predict XCI status

Probes were predicted to escape XCI when the male average and female average were unmethylated ($<0.15\%$ methylated) and when males and females showed a similar range of methylation (either the range of male and female methylation overlapped or, if the ranges did not overlap, the difference between the male and female average was less than 10%). Probes were predicted to be subject to XCI when males and females showed a different range of methylation where the difference between the male average and female average was greater than 10%. Probes were predicted to variably escape XCI when although the difference between the male and female average was greater than 10% there was also an overlap in the range of male and female methylation. When the male and female averages were greater than 15% and/or the difference between the male and female average was less than 10% probes were defined as unclassifiable. Supplementary Figure S1 outlines the decision tree.

RNA extraction and Q-PCR

RNA from four somatic cell hybrids containing a human Xi (t75-2maz34-4a, t48-1a-1Daz4A, t86-B1maz1b-3a and t11-4Aaz5), two somatic cell hybrids containing a human Xa (AHA-11aB1 and t60-12) and a control female cell line (GM7350) was extracted using Trizol (Invitrogen) as per the manufacturer's protocol and 5 μg converted to cDNA via a standard RT-PCR reaction using M-MLV (Invitrogen) at 42°C for 2 h followed by a 5 min incubation at 95°C. Real-time quantitative PCR (Q-PCR) was performed using the StepOnePlus™ Real-Time PCR System (Applied Biosystems, Darmstadt, Germany) on each sample in triplicate with the following conditions: 95°C (5 min), [95°C (30 s), 60°C (30 s), variable annealing temperature (30 s)] for 40 cycles and melting curve analysis [95°C (15 s), 60°C (60 s) then fluorescence was measured every 0.3°C per until 95°C]. Primer sequences and annealing

temperatures are listed in Supplementary Table S1. The average of three triplicate Ct values were corrected based on the average efficiency for each assay, as calculated by LinReg (Ramakers et al. 2003; Ruijter et al. 2009) and delta Ct values calculated for TSR2 and ZRSR2 compared to ZFX. The negative and positive error was calculated based on the sum of the standard deviation for the test (TSR2 or ZRSR2) and the control (ZFX) assay for each sample (in triplicate). All assays were found to not amplify mouse gDNA (data not shown).

MeDIP and whole genome amplification

MeDIP of male ($n = 3$) and female ($n = 3$) blood was performed as outlined in Vucic et al. (2009). Briefly, three reactions of 1 μg of genomic DNA were sonicated then 200 ng of input removed. The remaining 800 ng of DNA was denatured (95°C, 10 min) then 5 μg of anti-5′-methyl-cytosine mouse mAb (CalBiochem) added before incubating at 4°C for 2 h. 30 μL of Dynabeads M-280 Sheep anti-Mouse IgG (Dynal Biotech, Invitrogen) were then added followed by a 2 h incubation at 4°C. Two rounds of washing were performed to remove the Dynabeads then 100 μg of proteinase K was added and left overnight at 50°C. A phenol:chloroform clean-up was performed the next day and DNA resuspended in 10 μL H$_2$O. Whole genome amplification was performed using the Genome-Plex Complete Whole Genome Amplification Kit (Sigma) according to the manufacturer's instructions.

NimbleGen array processing and analysis

Three reactions using 1 μg of whole genome amplified DNA for each sample were labeled using Cy3-9mer primers for input and Cy5-9mer primers (TriLink Biotechnologies, Inc.) for IP. Labeling was performed as outlined in the NimbleGen Arrays User's Guide: ChIP-chip Analysis v3.0 (Roche NimbleGen, Inc) then samples were sent to the Fred Hutchinson Cancer Research Center (Seattle, WA, USA) for hybridization to a Human ChIP-chip 2.1M Whole-Genome Tiling, array number 10 (Roche NimbleGen, Inc). Files of the scanned arrays were processed according to the NimbleGen Arrays User's Guide: DNA Methylation Analysis v5.0 and the resulting ratio files subjected to BATMAN (Bayesian Tool for Methylation Analysis) (Down et al. 2008) to correct the effect of CpG density of MeDIP efficiency. The average standard deviation of the three samples was 0.05 in males and females. To ensure that samples were more similar within a sex than between, the 6 blood samples were combined in all 18 possible combinations of three and the average standard deviation (0.06) of all 18 combinations was always greater than that observed within each sex. Galaxy

(Goecks et al. 2010; Blankenberg et al. 2010) was then used to calculate the frequency of probes in the various genomic elements examined.

Expression data

Expression ratios on log base 2 scale from Su et al. (2002) in whole blood were downloaded from http://symatlas.gnf. org. Genes were divided by chromosome and ranked from lowest to highest expression. The top and bottom 20% of genes from each chromosome were used to represent those genes with the lowest and highest expression levels in blood.

Results

X-linked promoters show differences in methylation dependent on sex and CpG density

To determine how X-linked promoter methylation differed between the sexes we applied DNA from 36 male and 59 female bloods to the Illumina Infinium HumanMethylation27 array containing 1,085 X-linked probes. A total of 308 X-linked probes were removed from analysis as they were located in repetitive elements, mapped to more than one location in the genome and/or were located in the promoters of the cancer-testis family of genes (summarized in Supplementary Figure S1). To detect the large methylation differences between males and females previously reported at most X-linked promoters, we created three broad methylation classes: unmethylated (0–0.15 beta value), intermediate (0.15–0.60 beta value) and methylated (0.60–1.00 beta value) and then separated the remaining 777 X-linked probe results into one of the three methylation classes. The majority (67%) of the X-linked probes in male blood were shown to be unmethylated whereas the majority (66%) of probes in female blood had intermediate methylation. As CpG density is known to influence methylation we sub-divided probes based on their location within HC and IC islands or LCs. X-linked promoter probes in HC and IC islands were generally those which were unmethylated in male blood and intermediately methylated in female blood, whereas X-linked promoter probes in LCs were usually methylated in both sexes (Fig. 1a).

To ensure that the observed X-linked methylation differences between males and females were not simply due to differences in overall methylation between the sexes, autosomal methylation levels were compared between the same male and female blood samples. Male and female probes (1,843 probes located on chromosomes 20, 21 and 22) were compared after removal of 307 autosomal probes
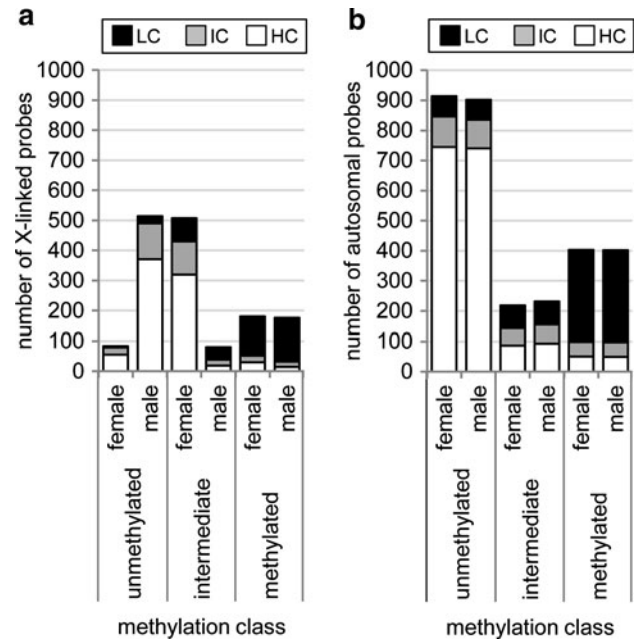

Fig. 1 Promoter methylation analysed in blood (female: $n = 59$ and male: $n = 36$) revealed X chromosome sex-specific methylation differences as well as differences based on CpG density. Probes were divided by CpG density (LC *black*, IC *gray*, HC *white*) and classified as unmethylated (0–15% methylated), intermediate (15–60% methylated) or methylated (60–100% methylated) in males and females. **a** Methylation levels in males and females were significantly different (p value < 0.0001, Mann–Whitney test) across all X-linked probes. The majority of HC and IC promoter probes ($n = 560$) on the X chromosome were unmethylated in males and intermediate in females. X-linked LCs probes ($n = 217$) were mostly methylated regardless of sex. **b** Methylation levels in males and females were not significantly different (p value = 0.2779, Mann–Whitney test) across autosomal probes. The majority of promoter probes on chromosomes 20, 21 and 22 were unmethylated. Probes in HC and ICs ($n = 1088$) were mostly unmethylated whereas LC ($n = 448$) probes were mostly methylated. Males and females showed no differences in their methylation classes

that were located in repetitive elements and/or mapped to multiple locations in the genome. The majority (98%) of all autosomal probes showed the same methylation level in males and females regardless of CpG density. Furthermore, those probes at which males and females had different methylation classes (unmethylated, intermediate or methylated) showed only a 2% difference in methylation and were not significantly different in their methylation, making it unlikely that this difference was biologically functional. We therefore conclude that the differences between male and female methylation is mostly unique to the X chromosome and occurred primarily at X-linked promoter probes located in HC and IC islands (Fig. 1b). Given that 10% of X-linked probes in HC and IC islands were unmethylated in females, and that unmethylated X-linked promoters have previously been found at genes which escape XCI (Carrel et al. 1996) we wanted to

examine whether the detection of unmethylated probes in HC and IC islands in males and females reflected escape from XCI as was previously proposed by (Yasukochi et al. 2010).

Assessment of promoter methylation as a predictor of XCI status

Before we could evaluate how effectively X-linked promoter methylation might predict the XCI status of a gene, we had to establish a consistent method of translating male and female methylation results from multiple probes into a single genic XCI status. We developed a decision tree examining the average male and average female methylation levels, the difference between these averages and the range of methylation observed in males and females to call the XCI status of each probe (Supplementary Figure S1). The majority of probes in genes with HC and IC islands (77%) predicted the gene was subject to XCI, 11% predicted escape from XCI, 3% variable escape and 8% were unclassifiable. The 560 probes in HC and IC islands were found in 343 X-linked genes (145 genes were represented by only 1 probe), therefore probes from the 198 genes that contained more than one probe in an HC or IC island were combined to create a single predicted XCI status for each gene. For 90% of genes, all probes (if multiple probes were present) predicted the same XCI status. Of the remaining 10% of genes, over half had one probe where an XCI status

was predicted with the other probe being unclassifiable. These genes were, therefore, given the XCI status of the probe which had predicted an XCI status (subject, escape or variable escape). In only 5 genes out of the 343 X-linked genes was there a conflict in which 1 probe predicted an XCI status of escape and the other subject to XCI was found (Supplementary Table S2). Interestingly, all five of these were found in genes that had previously been reported to escape, or variably escape XCI (Carrel and Willard 2005). We predicted that the majority (81%) of genes with probes in HC or IC islands were subject to XCI, 10% escaped XCI, 2% variably escaped XCI and 5% of genes remained unclassifiable (Fig. 2a). Of the 19 unclassifiable genes, 13 were methylated in both males and females. Overall we were able to use methylation to predict an XCI status for 95% of examined X-linked genes with probes in HC or IC islands and could therefore compare these predictions to the XCI status of the same genes previously determined by expression.

To examine whether X-linked promoter methylation was effective at predicting XCI status, we analyzed genes at which the XCI status had previously been established and determined the degree to which our predicted XCI status agreed. We compared our predicted XCI status in blood with the XCI status derived from reported studies of somatic cell hybrids by Carrel and Willard (2005). Since the bulk of X-linked genes are subject to XCI we first
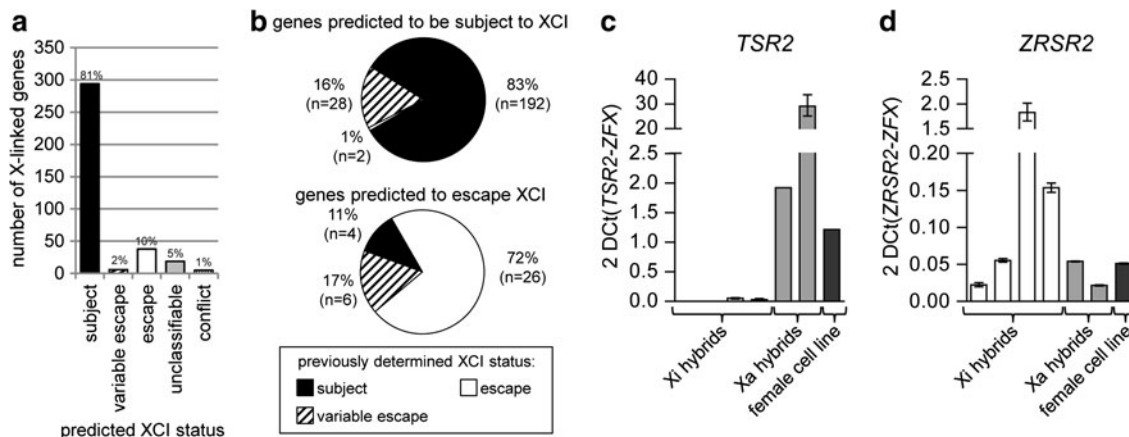

Fig. 2 The XCI status predicted using methylation in blood corresponds with previously determined XCI status. **a** The XCI status (subject *black*, variable escape *diagonal stripes*, escape *white*, unclassifiable *gray*, conflicts *dotted*) of 372 X-linked genes with probes in HC and IC islands was predicted using methylation. The percentage of the total X-linked genes with probes in HC and IC islands is given for each predicted XCI status. In blood, the majority (81%) of genes are predicted to be subject to XCI. **b** The XCI status previously determined by Carrel and Willard (2005) in somatic cell hybrids (subject *black*, variable escape *diagonal stripes*, escape *white*) for those genes predicted by methylation to be subject to XCI (*black bar* in **a** and *top pie chart*) and those genes predicted to escape XCI (*white bar* in **a** and *bottom* pie chart). **c–d** Q-RT-PCR confirmation in

somatic cell hybrids of predicted XCI status based on methylation. The expression level of two genes (*TSR2* and *ZRSR2*) in four somatic cell hybrids containing a human Xi (*white* t75-2maz34-4a, t48-1a-1DAZ4A, t86-B1maz1b-3a and t11-4Aaz5), two somatic cell hybrids containing a human Xa (*light gray* AHA-11aB1 and t60-12) and a control female cell line (*dark gray* GM7350) were compared to confirm that methylation could predict XCI status. Test genes (*TSR2* and *ZRSR2*) were normalized against a gene known to escape XCI (*ZFX*). *Error bars* represent the positive and negative error between three replicate PCRs. **c** *TSR2* was unmethylated in male blood and intermediate in female blood and was predicted to be subject to XCI. **d** *ZRSR2* was unmethylated in male and female blood and was predicted to escape XCI

examined the genes with probes in HC or IC islands which we predicted to be subject to XCI. After the removal of genes not examined by Carrel and Willard (2005), 83% ($n = 192$) of the genes predicted by X-linked promoter methylation to be subject to XCI were also found by Carrel and Willard (2005) to be subject to XCI. Given our interest in using X-linked promoter methylation to predict escape from XCI we also examined those genes with probes in HC or IC islands for which our methylation data had predicted escape from XCI. Here we found that 72% of genes predicted by X-linked promoter methylation to escape XCI were also shown by Carrel and Willard (2005) to escape XCI (Fig. 2b). To further address the ability of promoter methylation to predict XCI status, the expression patterns of two genes (*TSR2* AND *ZRSR2*) not examined by Carrel and Willard (2005) were examined by Q-PCR in somatic cell hybrids (four hybrids containing a human Xi and two containing a human Xa) as well as a control female cell line. Based on promoter DNA methylation, *TSR2* was predicted to be subject to XCI and this was confirmed, as none of the Xi hybrids showed expression comparable to the Xa hybrids or the female cell line. *ZRSR2* was predicted to escape XCI and this too was confirmed with all Xi hybrids showing expression at least as high as the Xa hybrids and the female cell line (Fig. 2c, d).

This validation, along with the high degree of agreement between previously determined XCI status and our prediction using X-linked promoter methylation, led us to believe that the methylation of probes in HC and IC islands X-linked promoters can be used to predict XCI status and, therefore, we can propose an XCI status for 62 genes (see Supplementary Table S3). A few of these have been described in other studies and our prediction of XCI is in agreement (Brinkman et al. 2006; Yasukochi et al. 2010; Lopes et al. 2006). While our results suggest that methylation is an effective predictor of XCI status, 59 genes were shown by Carrel and Willard (2005) to have a different XCI in somatic cell hybrids than was predicted by our analysis of methylation in blood. Tissue-specific escape from XCI has been reported in mouse (Yang et al. 2006) and therefore we wished to investigate the extent to determine if tissue-differences could be a substantial contributor to the 15% discordance we observed between the XCI status in somatic cell hybrids and blood (Carrel and Willard 2005).

Tissue-specific XCI is observed at 12% of genes

We extended our Illumina Infinium HumanMethylation27 array analysis to fetal tissues (muscle, kidney, brain and spinal cord) to determine if all tissues showed the same male and female methylation, and therefore the same predicted XCI status. We first confirmed that, as with blood,

there was a sex-specific methylation difference that was limited to the X chromosome and not the autosomes (Supplementary Figure S2). We then examined fetal muscle and fetal kidney and combined fetal brain and fetal spinal cord into one fetal "neural" tissue category. The same process of predicting XCI status as was used in blood, again demonstrated that although the level of methylation was significantly different between tissues ($p < 0.0001$), the majority of X-linked CpG-island genes showed a pattern of methylation consistent with being subject to XCI (unmethylated males and intermediate females) regardless of the tissue examined (blood = 81%, fetal muscle = 74%, fetal neural = 66%, fetal kidney = 73%). Interestingly, a larger proportion of X-linked genes showed a pattern of methylation that we considered predictive of escape from XCI in fetal tissues (muscle = 15%, neural = 17%, kidney = 15%) compared to blood (10%), despite the considerably smaller sample size, which led us to compare the predicted XCI in all tissues to determine how often the same XCI status was predicted in all tissues (Fig. 3a).

We compared the predicted XCI status across all tissues and found that the majority (78%) of X-linked genes showed the same predicted XCI status in all tissues examined. An additional 6% of genes showed the same predicted XCI status in all but one tissue (which was designated unclassifiable). However, at 12% of X-linked genes, promoter methylation resulted in a different predicted XCI status in different tissues (Fig. 3b) and we designated these genes as showing tissue-specific XCI. Of the genes which showed tissue-specific XCI, nearly half (48%) show more escape in the fetal tissues compared to blood. Supplementary Table S4 lists genes which displayed tissue-specific XCI, and the locations of these genes are shown in Fig. 4 along with the location of genes that showed consistent XCI patterns. The distribution of these genes is influenced by the choice of probes on the array, and notably no pseudoautosomal probes were included. Our finding that X-linked promoter methylation differs across 12% of genes examined suggests tissue-specific XCI in these genes.

To investigate if tissue-specific XCI was consistent between females we examined six different females each with at least two different fetal tissues. We compared the predicted XCI status in fetal tissues (muscle, neural tissue and kidney) from four females, fetal muscle and fetal kidney from one female and fetal neural and fetal kidney from another female. We used the individual female's methylation value along with the average male methylation in the same tissue to predict XCI status. In each female examined, 84–86% (see Supplementary Table S5) of the total X-linked genes examined were predicted to escape or be subject to XCI across all tissues, while 8–14% of genes
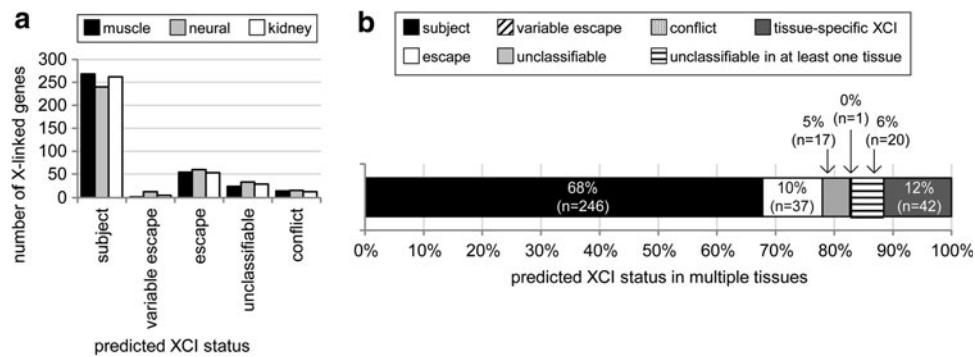
Fig. 3 Most genes show the same predicted XCI status in all tissues examined while 12% of genes show tissue-specific XCI. **a** Male and female methylation was used to predict XCI status (as outlined in Supplementary Figure S1) of genes with probes in HC and IC islands in fetal muscle (*black*, female $n = 6$; male $n = 4$), fetal neural tissue (*gray*, female $n = 6$; male $n = 5$) and fetal kidney (*white*, female $n = 6$; male $n = 5$). **b** The combined predicted XCI status in all four tissues examined (blood, fetal muscle, fetal neural and fetal kidney).

The majority of genes showed the same XCI status (subject *black*, variable escape *diagonal stripes*, escape *white*, unclassifiable *gray*, conflicts *dotted*) in all tissues. 6% of genes were unable to predict an XCI status in at least one tissue but predicted same XCI status in all other tissues (*horizontal stripes*). 12% of genes showed tissue-specific methylation differences which resulted in at least one tissue having a different predicted XCI from the other tissues (*dark gray*)

were unclassifiable and 1–7% of genes showed tissue-specific XCI. In females with multiple tissues, fetal muscle showed the fewest genes with tissue-specific escape while fetal neural tissue showed the most. We found that when escape from XCI was predicted by X-linked promoter methylation in one tissue, it was generally predicted in all tissues (listed in Supplementary Table S6). Overall, DNA methylation-based evidence for tissue-specific XCI was found in all females examined, with the highest degree of tissue-specific escape always observed in fetal neural tissue but with a great deal of variability between females.

X-linked non-island methylation is a poor predictor of XCI status

Having established that the X-linked promoter methylation of probes in HC and IC islands was highly predictive of XCI status, we were interested if the same methodology could be applied to LC probes (those not located in HC and IC islands). It should be noted that the majority of probes on the Illumina Infinium HumanMethylation27 array are located in CpG islands (65%) associated with promoters. Some X-linked LC promoters have been reported to exhibit methylation that correlates with gene silencing on the Xi [such as *TIMP1* (Anderson and Brown 2002), *CHM* (Carrel and Willard 1999), and *OTC* (Yorifuji et al. 1998)]. The same decision tree (Supplementary Figure 1) used to predict the XCI status of probes in HC and IC islands was applied to LC probes to evaluate what proportion of LC probes showed a methylation status which could predict an XCI status. Approximately one quarter (27%) of all LC probes examined were located in promoters (±1 kb around the TSS) which also included an HC or IC island while the remaining LC probes were located in promoters which

lacked an HC or IC island. LC probes were generally unclassifiable (82%) due to high methylation regardless of whether a CpG island was present within the promoter region. The remaining probes showed methylation patterns classifiable as escape (4%), variable escape (4%) or subject to XCI (genes with CpG islands: 21%, genes without CpG islands: 7%). Those LC probes with a methylation status which predicted an XCI status of subject, variable escape or escape, along with any HC or IC probes in the same gene are listed in Supplementary Table S7. We compared the XCI status predicted using methylation to that determined by Carrel and Willard (2005) and found that approximately 40% of LC probes predicted the same XCI status as Carrel and Willard (2005) regardless of whether the LC probe was in the promoter of a gene with a CpG island or not. Given the low concordance between the predicted XCI status based on LC probe methylation and that previously determined by Carrel and Willard (2005), we conclude that LC probes are not usually reliable as a predictor of XCI status.

X-linked HC and IC promoters show the strongest sex-specific methylation difference

The data we analysed from the Illumina Infinium HumanMethylation27 array examined only approximately 45% of X-linked promoters and did not examine any non-promoter elements such as the intragenic and intergenic regions of the chromosome. To expand the study of X-linked methylation beyond promoters, MeDIP was performed on DNA isolated from male ($n = 3$) and female ($n = 3$) blood followed by hybridization to a NimbleGen 2.1M array to analyse chromosome-wide methylation of chromosomes 20, 21 and 22 along with the X chromosome.
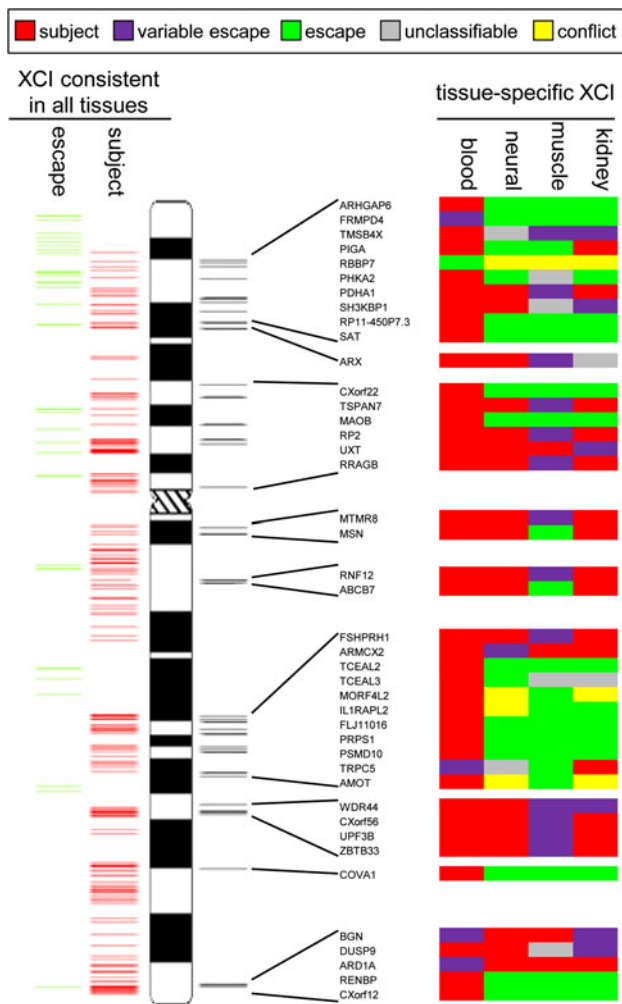
Fig. 4 Genes predicted to show tissue-specific XCI based on methylation from the Illumina Infinium HumanMethylation27 array as found across the X chromosome. The genomic locations of genes which showed the same predicted XCI status in all tissues examined (escape *green*, subject *red*) are shown to the left of the X chromosome ideogram. On the right are the genomic locations of genes in which at least one tissue had a predicted XCI status different from the other tissues. The predicted XCI status (subject *red*, variable escape *purple*, escape *green*, unclassifiable *gray*, conflict *yellow*) in each tissue examined (blood, fetal neural, fetal muscle and fetal kidney) is shown along with the names of all genes which show tissue-specific XCI

To correct for the effect of CpG density on MeDIP efficiency, BATMAN (Down et al. 2008) was used to convert the ratio of IP:IN into a methylation value from 0 to 1. BATMAN was performed on all samples and the resulting scores averaged to create one average male score and one average female score, in subsequent analyses only male versus female differences that were greater than intrasex differences were considered for statistical significance. Methylation histograms were compiled to assess the distribution of methylation on different DNA elements of interest.

To determine if the X-linked sex-specific methylation difference found using the Illumina Infinium HumanMethylation27 array could also be observed via MeDIP, the first elements we examined were X-linked promoters. Promoters were defined as the probes within 1 kb up and downstream of all TSS, therefore the presence of an HC island in a promoter, or the presence of an IC but not an HC, resulted in the classification of HC or IC promoter, respectively. LC promoters were those promoters which had neither an HC nor IC island in the region 1 kb upstream and downstream of the TSS. X-linked HC promoters showed a higher frequency of unmethylated probes in the male than the female with a significantly different ($D = 0.12$, $p < 2.2$ E−16) distribution between the sexes (Fig. 5a). IC promoters on the X chromosome showed a significantly different ($D = 0.07$, $p = 2.5$ E−10) distribution between the sexes and were slightly more unmethylated on the male X chromosome than on the female X chromosomes. In addition, X-linked IC promoters also had a higher percentage of both male and female probes being intermediate or fully methylated than was observed in HC promoters (Supplementary Figure S3a). On the autosomes, neither HC nor IC promoters were significantly different between males or females, however, HC promoters were mostly unmethylated while IC promoters also showed intermediate methylation. X-linked LC promoters were mostly methylated and were not significantly different between males and females while autosomal LC promoters were slightly less methylated (Fig. 5b). By examining all known X-linked promoters we were able to show that sex-specific methylation differences were highly significant at X-linked HC promoters, slightly significant at X-linked IC promoters but not significant at X-linked LC promoters or on the autosomes.

Our definition of promoter elements comprised only approximately 2% of base pairs on the X chromosome; therefore, determining the methylation status at non-promoter elements was of critical importance if an overview of chromosome-wide methylation was to be established. Intragenic and intergenic regions showed similar methylation in males and females; however, on the X chromosome these regions were bimodally distributed whereas on the autosomes they were not (Fig. 5c, d). Across the genome there are CpG islands not currently associated with known genes. Males and females displayed significantly different methylation at HC islands not associated with a known TSS on the X chromosome ($D = 0.09$, $p = 2.4$ E−9) but not on the autosomes nor at X-linked or autosomal IC islands not associated with a known TSS. The IC islands were more methylated in both sexes than IC promoters on either the X chromosome or the autosomes (Supplementary Figure S3b and c). Having compared elements across the entire X chromosome we confirmed that although HC
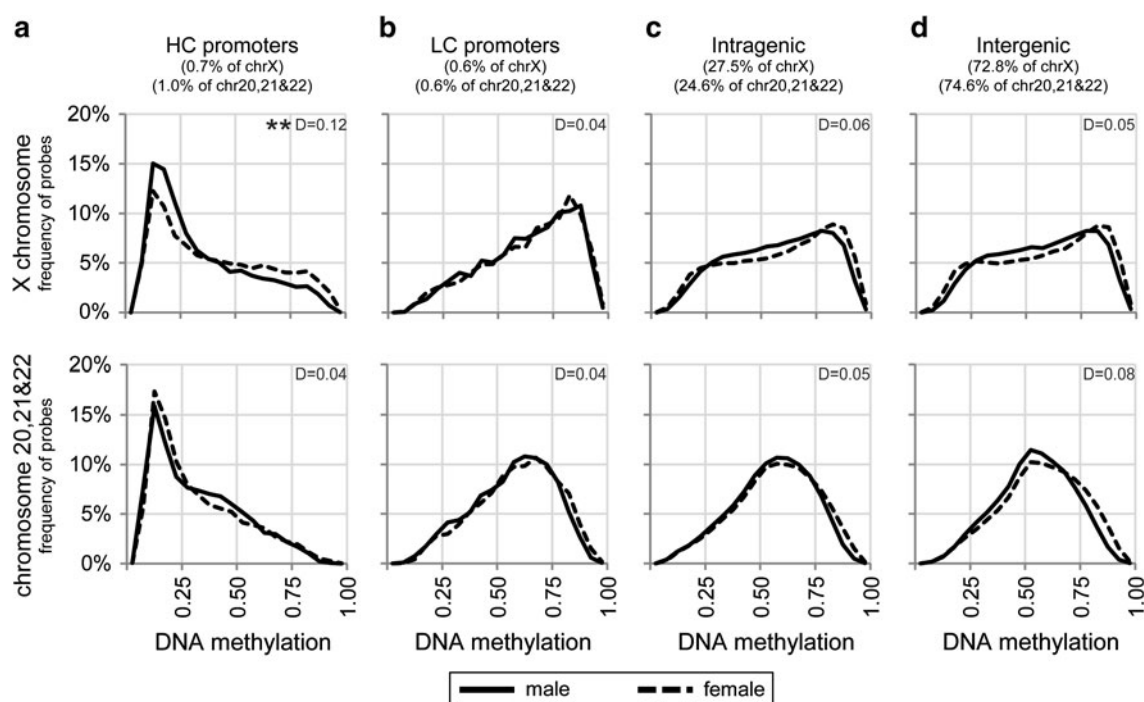
Fig. 5 Methylation histograms reveal X-linked HC promoters show the largest X-linked sex-specific methylation difference. The average male and average female methylation from probes representing four different genomic elements was used to create methylation histograms by determining the frequency at which probes were at a specific level of methylation (20 bins from 0 to 1.0 methylated). Female methylation frequencies are shown as *dotted lines* and males as *solid lines* with methylation frequencies from the X chromosome displayed on the *upper row* and the autosomal average from chromosomes 20, 21 and 22 on the *bottom*. The percentage of the total chromosomal DNA each element represents is given for the X chromosome and the autosomes (chromosomes 20, 21 and 22). Significance was calculated comparing the distribution of average male and average female methylation using the Kolmogorov–

Smirnov test. When *p* values were greater than 0.0001 they were not significant, however, *p* values between 0.0001 and 1.0 E−10 (*asterisk*) and *p* values < 1.0 E−10 (*double asterisk*) were considered significantly different. **a–b** Promoter elements (the 1 kb up and downstream of all TSS) showed differences in methylation frequencies based on CpG density. HC promoters (**a**) showed males were hypomethylated compared to females on the X chromosome but not the autosomes. LC promoters (**b**) (contained neither an HC nor IC island) showed no sex-specific methylation difference on either the X chromosome or the autosomes. **c–d** Non-promoter elements tended to be methylated on the male and female X chromosome and intermediately methylated on the autosomes in both intragenic (**c**) and intergenic (**d**) regions

promoters compose a small fraction of the X chromosome they are the element which showed the strongest degree of X-linked sex-specific methylation.

X-linked genes with high expression show high gene-body methylation

It has previously been shown that the intragenic regions of highly expressed genes are more methylated than those of lowly expressed genes (Aran et al. 2011) and on the X chromosome gene-bodies of the Xa have been found to be more methylated than on the Xi (Hellman and Chess 2007). We therefore used published expression data (Su et al. 2002) to separate genes with high expression levels (top ranking 20%) from those with low expression (bottom 20%) to allow for a male:female comparison of gene-body methylation levels relative to expression levels. No significant differences between the distribution of male and female methylation were found at exons or introns on

either the X chromosome or the autosomes. However, X-linked introns of highly expressed genes were more methylated than those of lowly expressed genes in both males and females (Fig. 6a, b). Interestingly, X-linked exons did not show this differences between highly and lowly expressed X-linked genes. Overall, the division of genes based on expression did not demonstrate a significant difference in the distribution of male:female methylation although X-linked intronic methylation was greater in more highly expressed genes compared to lowly expressed genes.

Discussion

The presence of methylation at X-linked CpG island promoters on the Xi is classically associated with genes subject to XCI (Cotton et al. 2009; Jamieson et al. 1996). We found that in all tissues examined (blood, fetal muscle, fetal
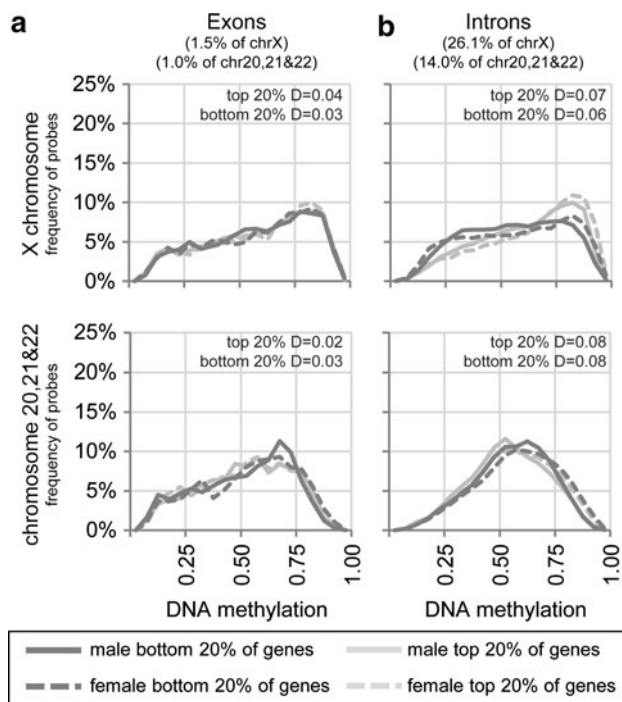
Fig. 6 X-linked introns but not exons show differences in methylation based on expression level. The average male and average female methylation from probes representing four different genomic elements was used to create methylation histograms by determining the frequency of probes at a specific level of methylation (20 bins from 0 to 1.0 methylated). Female methylation frequencies are shown as *dotted lines* and males as *solid lines* with methylation frequencies from the X chromosome displayed on the *upper row* and the autosomal average from chromosomes 20, 21 and 22 on the *bottom*. X-linked and autosomal (chromosomes 20, 21 and 22) genes were separated based on expression (determined in Su et al. 2002) and the *top* (*light gray*) and *bottom* (*dark gray*) 20% divided into those which correspond to either the exons (**a**) or introns (**b**). Significance was calculated comparing the distribution of male and female methylation using the Kolmogorov–Smirnov test. When *p* values were greater than 0.0001 they were not significant (ns), however, *p* values between 0.0001 and 1.0 E−10 (*asterisk*) and *p* values < 1.0 E−10 (*double asterisk*) were considered significantly different. While exons (**a**) were similarly methylated regardless of sex or expression level on both the X chromosome and the autosomes, introns (**b**) were more methylated in highly expressed X-linked genes than lowly expressed X-linked genes in both sexes. Autosomal introns showed no methylation difference in either sex

kidney, fetal brain and fetal spinal cord), the majority of X-linked promoter probes in HC and IC islands were unmethylated in males and intermediately methylated in females which is the pattern of methylation typically associated with genes subject to XCI. In support of this sex difference being reflective of XCI, nearly all autosomal probes (over 95%) showed the same methylation status in males and females, regardless of CpG density. Genes which escape XCI have previously been found to be unmethylated in both males and females (Carrel et al. 1996) and this unique property has been used to propose novel

genes which escape XCI (Yasukochi et al. 2010) in neutrophils. We extended the search for genes which escape XCI to blood, fetal muscle, fetal kidney and fetal neural tissue using the DNA methylation pattern for genes with probes in HC and IC islands. We found a high degree of concordance (81%) with the XCI status previously determined by Carrel and Willard (2005) in somatic cell hybrids; however, for 19% of genes there was discordance between our methylation-based prediction in blood and results from expression in hybrids.

Previous studies which have examined XCI status have typically used either somatic cell hybrids or females with clonal XCI who are heterozygous for known SNPs (Carrel and Willard 2005). Similar to our results, previous comparisons between expression in hybrids and in female tissues have shown discordancies (Carrel and Willard 2005; Stabellini et al. 2009). We propose several different reasons for the differences between the XCI status we predicted using methylation and that of Carrel and Willard (2005). First, methylation may not always be an accurate predictor of XCI status. This might occur in regions where other epigenetic marks, such as histone modifications, are more important to maintain XCI. A second possible explanation is that due to the proposed decrease in stability of XCI of somatic cell hybrids (Gartler and Goldman 1994; Stabellini et al. 2009) genes which are typically subject to XCI in blood now escape XCI in somatic cell hybrids. If the differences in XCI were caused by a decrease in stability of XCI in somatic cell hybrids then any conflicts in XCI status should involve a higher degree of escape from XCI in the somatic cell hybrids. 15% of genes examined showed more escape from XCI in the somatic cell hybrids than in blood supporting this hypothesis, however, 4% of genes showed more escape in blood than the somatic cell hybrids. These conflicts cannot be explained by a decrease in the stability of XCI in somatic cell hybrids, suggesting that hybrid instability is not the full explanation.

A third possibility is that somatic cells hybrids and blood actually have different XCI statuses at a subset of genes. We attempted a direct comparison of methylation in hybrids (data not shown) with expression status for individual genes; however, we observed considerable variability of methylation between hybrids, even in Xa hybrids (data not shown) and were thus not able to compare methylation to expression in the hybrids. We therefore examined male and female methylation levels in different human tissues to determine if tissue-specific methylation changes were frequent. While most genes had the same predicted XCI status in all tissues examined, we detected potential tissue-specific XCI in 12% of genes, the majority of which reflected genes being subject to XCI in blood while at least one other tissue was not subject to XCI. We also found that over 50% of genes showing tissue-specific

XCI were found within 1MB of each other suggesting a possible regional effect causing tissue-specific XCI. We caution that when examining X-linked genes the XCI status should always be confirmed in the tissue of interest. The degree of predicted tissue-specific XCI differed between the six examined females and between tissues, with neural tissue showing the highest degree of predicted tissue-specific escape from XCI. Studies examining expression amongst all X-linked genes have consistently shown brain to have one of the highest X:autosome ratios, regardless of the technique being used (Xiong et al. 2010). The X chromosome contains an overrepresentation of genes expressed in the brain (Nguyen and Disteche 2006; Vicoso and Charlesworth 2006) and many X-linked genes are known to play a role in X-linked mental retardation [reviewed in (Ropers 2006)] which is significantly more common in males than in females (Turner and Turner 1974; Croen et al. 2001). Expression of genes from the Xi when the Y homolog is no longer functional could lead to a dosage difference between males and females, and might contribute to sex-specific differences in disease susceptibility.

On the autosomes, tissue-specific methylation differences in CpG islands have previously been detected across a number of tissues (Rakyan et al. 2008; Eckhardt et al. 2006; Irizarry et al. 2009; Illingworth et al. 2008; Schilling and Rehli 2007) and it has been proposed that the majority of tissue-specific differentially methylated regions are located in the regions surrounding CpG islands known as CpG island shores (Irizarry et al. 2009). We found that the majority of probes which showed tissue-specific XCI (83%) were located in HC islands rather than shores. The criteria we used to detect sex-differences on the X chromosome were designed to identify large changes in methylation associated with the XCI status of the gene which may explain why the tissue-specific methylation we observed on the X chromosome was mostly located in the CpG islands and not in the shores as was previously reported on the autosomes (Irizarry et al. 2009). Our analysis of X-linked HC non-promoters (HC islands not associated with a known promoter) revealed a similar hypomethylation in male blood compared to female blood. The presence of a sex-specific methylation difference is evidence that these HC islands may be the promoters of unannotated X-linked genes that are subject to XCI. This is in agreement with a previous report in which the majority of genome-wide orphan CpG islands were predicted to be associated with the promoters of unknown genes based on histone modifications and the presence of RNA Pol II (Illingworth et al. 2010). The X-linked IC non-promoter we examined lacked a significant sex-specific methylation difference suggesting that it is less likely that these CpG islands are associated with unknown genes. To confirm that the X-linked non-promoters islands we were predicting to be promoters were in fact not enhancers, we examined the histone modifications typically associated with enhancers (Heintzman et al. 2007) and did not find any significant enrichment (data not shown).

Across the genome, the most widely expressed genes tend to have a promoter CpG island along with a smaller subset of tissue-specific genes (Gardiner-Garden and Frommer 1987). On the X chromosome, some genes, notably androgen receptor, which has been widely used to examine XCI skewing, also have tissue-specific expression (Su et al. 2002) yet show consistent methylation (males: unmethylated, females: $\sim 50\%$ methylated) even in tissues where they are not expressed (Bittel et al. 2008). Consistent with this observation, data from the Illumina Infinium HumanMethylation27 array showed that female X-linked promoters had no differences in methylation between highly and lowly expressed genes at any CpG density while males showed a slight significance at X-linked IC promoters. Chromosome-wide methylation analysis revealed that the HC promoters of highly expressed X-linked genes maintained a significant difference between males and females (data not shown), where males were more hypomethylated than females. On both the X chromosome and the autosomes, all other promoters showed no significant difference between the distribution of males and female methylation at either highly or lowly expressed genes.

The association between promoter methylation and transcriptional silencing is well established (Saxonov et al. 2006; Gardiner-Garden and Frommer 1987; Bird 1986); however, the interaction between gene-body methylation and transcription, as well as the methylation status of intergenic regions, is less clear. In general, the distribution of methylation in intragenic and intergenic regions of the X chromosome is different from the autosomes, likely reflecting the unique sequence composition of the sex chromosomes. This difference is less apparent at exons where the distribution of methylation on the X chromosome is more similar to the autosomes. When X-linked introns are examined the methylation of the top 20% of expressing genes differs from the bottom 20% of genes, whereas on the autosomes expression does not greatly affect the distribution of methylation. The shift of methylation of highly expressed X-linked introns yields a distribution of methylation very similar to that found at all X-linked exons. Although we do not observe a significance difference between the distributions of male and female methylation we do see that highly expressed X-linked introns are more methylated than lowly expressed introns. A role for transcription in gene-body methylation is supported by a recent genome-wide study which showed that early replicating genes have more gene body methylation

than late replicating genes (Aran et al. 2011) while another study showed methylation of the gene body was more likely to be found in highly expressed genes (Brenet et al. 2011). Differences in autosomal exon and intron methylation have previously been found, with first exons typically being unmethylated (especially if the gene is expressed) (Brenet et al. 2011; Edwards et al. 2010) while internal exons and introns tend to show variable methylation (Brenet et al. 2011). While the difference between X-linked male and female gene-body methylation was small, this is consistent with our previous analysis (Cotton et al. 2009) and suggests that X-linked male:female gene-body methylation differences may not be as large as other studies have suggested.

There are several features of exons and introns which may explain the observed differences in methylation. Firstly, although exons typically make up a smaller portion of genes than introns, exons have a higher GC content and CpG fraction than any region of the genome other than promoters (Saxonov et al. 2006). The difference in size between exons and introns may have also influenced the observed methylation as the smaller exons will be more affected by the surrounding methylation than the larger introns. CpG density is known to have an effect on the pull down success of techniques such as MeDIP and while BATMAN is designed to correct of the effect of CpG density of pull down efficiency (Down et al. 2008), the methylation differences observed between exons and introns may in part be due to differences in CpG density. Exons have been shown to be enriched, compared to introns, for histone modifications associated with the transcription of active genes (Hodges et al. 2009); our data suggests that X-linked exons maintain their methylation status regardless of expression while it is introns which show increased methylation with higher expression suggesting that transcription may affect exons and introns differently.

The nature of the Xa and Xi provides a unique system to compare methylation between active and inactive chromatin domains. We conclude that the largest difference in X-linked methylation between males and females is found at CpG island promoters. Therefore, we proposed that methylation differences between the sexes could be used to predict XCI status and found a overall good concordance with XCI statuses previously determined by expression analysis. Most genes showed similar methylation, and therefore the same predicted XCI status across tissues, thus, our results support that discrepancies between the XCI statuses we predicted using methylation and those previously determined may be due to tissue-specific XCI, as 12% of genes showed methylation patterns suggestive of tissue-specific XCI in the four tissues we examined. Using methylation to predict XCI status would allow for

examination of a gene that is not expressed and would not require extraction of RNA or restrict studies to females with clonal XCI. Outside of CpG islands chromosome-wide methylation analysis revealed differences between exons and introns suggesting that the effects of transcription on gene-body methylation may affect exons and introns differently.

Conflict of interest The authors declare no conflict of interest.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Anderson CL, Brown CJ (2002) Variability of X chromosome inactivation: effect on levels of TIMP1 RNA and role of DNA methylation. Hum Genet 110:271–278

Aran D, Toperoff G, Rosenberg M, Hellman A (2011) Replication timing-related and gene body-specific methylation of active human genes. Hum Mol Genet 20(4):670–680. doi:10.1093/hmg/ddq513

Bird A (1980) DNA methylation and the frequency of CpG in animal DNA. Nuc Acids Res 8:1499–1504

Bird AP (1986) CpG-rich islands and the function of DNA methylation. Nature 321:209–213

Bittel DC, Theodoro MF, Kibiryeva N, Fischer W, Talebizadeh Z, Butler MG (2008) Comparison of X-chromosome inactivation patterns in multiple tissues from human females. J Med Genet 45(5):309–313. doi:10.1136/jmg.2007.055244

Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J (2010) Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol Chapter 19:Unit 19 10 11–21. doi:10.1002/0471142727.mb1910s89

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19(2):185–193

Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND, Scandura JM (2011) DNA methylation of the first exon is tightly linked to transcriptional silencing. PLoS ONE 6 (1):e14524. doi:10.1371/journal.pone.0014524

Brinkman AB, Roelofsen T, Pennings SW, Martens JH, Jenuwein T, Stunnenberg HG (2006) Histone modification patterns associated with the human X chromosome. EMBO Rep 7(6):628–634. doi:10.1038/sj.embor.7400686

Carrel L, Willard HF (1999) Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes

X inactivation in some human cell lines but is inactivated in others. Proc Natl Acad Sci USA 96:7364–7369

Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434(7031):400–404

Carrel L, Clemson CM, Dunn JM, Miller AP, Hunt PA, Lawrence JB, Willard HF (1996) X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme E1 and PCTAIRE-1 genes in human and mouse. Hum Mol Genet 5(3):391–402

Chow J, Heard E (2009) X inactivation and the complexities of silencing a sex chromosome. Curr Opin Cell Biol 21(3):359–366. doi:10.1016/j.ceb.2009.04.012

Cotton AM, Avila L, Penaherrera MS, Affleck JG, Robinson WP, Brown CJ (2009) Inactive X chromosome-specific reduction in placental DNA methylation. Hum Mol Genet 18(19):3544–3552. doi:10.1093/hmg/ddp299

Croen LA, Grether JK, Selvin S (2001) The epidemiology of mental retardation of unknown cause. Pediatrics 107(6):E86

De Smet C, Lurquin C, Lethé B, Martelange V, Boon T (1999) DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. Mol Cell Biol 19(11):7327–7335

Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, Graf S, Johnson N, Herrero J, Tomazou EM, Thorne NP, Backdahl L, Herberth M, Howe KL, Jackson DK, Miretti MM, Marioni JC, Birney E, Hubbard TJ, Durbin R, Tavare S, Beck S (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nat Biotechnol 26(7):779–785. doi:10.1038/nbt1414

Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, Lin SM (2010) Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 11:587. doi:10.1186/1471-2105-11-587

Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K, Beck S (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet 38(12):1378–1385

Edwards JR, O'Donnell AH, Rollins RA, Peckham HE, Lee C, Milekic MH, Chanrion B, Fu Y, Su T, Hibshoosh H, Gingrich JA, Haghighi F, Nutter R, Bestor TH (2010) Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. Genome Res 20(7):972–980. doi:10.1101/gr.101535.109

Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, Diekhans M, Dreszer TR, Giardine BM, Harte RA, Hillman-Jackson J, Hsu F, Kirkup V, Kuhn RM, Learned K, Li CH, Meyer LR, Pohl A, Raney BJ, Rosenbloom KR, Smith KE, Haussler D, Kent WJ (2011) The UCSC Genome Browser database: update 2011. Nucleic Acids Res 39 (database issue):D876–882. doi:10.1093/nar/gkq963

Gardiner-Garden M, Frommer M (1987) CpG islands in vertebrate genomes. J Mol Biol 196:261–282

Gartler SM, Goldman MA (1994) Reactivation of inactive X-linked genes. Dev Genet 15:504–514

Goecks J, Nekrutenko A, Taylor J (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol 11 (8):R86. doi:10.1186/gb-2010-11-8-r86

Goodfellow PJ, Mondello C, Darling SM, Pym B, Little P, Goodfellow PN (1988) Absence of methylation of a CpG-rich region at the 5' end of the MIC2 gene on the active X, and inactive X, and the Y chromosome. Proc Natl Acad Sci USA 85:5605–5609

Hansen RS, Gartler SM (1990) 5-Azacytidine-induced reactivation of the human X chromosome-linked PGK1 gene is associated with a large region of cytosine demethylation in the 5' CpG island. Proc Natl Acad Sci USA 87(11):4174–4178

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39(3):311–318

Hellman A, Chess A (2007) Gene body-specific methylation on the active X chromosome. Science 315(5815):1141–1143

Hodges E, Smith AD, Kendall J, Xuan Z, Ravi K, Rooks M, Zhang MQ, Ye K, Bhattacharjee A, Brizuela L, McCombie WR, Wigler M, Hannon GJ, Hicks JB (2009) High definition profiling of mammalian DNA methylation by array capture and single molecule bisulfite sequencing. Genome Res 19(9):1593–1605. doi:10.1101/gr.095190.109

Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C, Bird A (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 6 (1):e22. doi:10.1371/journal.pbio.0060022

Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr AR, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genet 6 (9). doi:10.1371/journal.pgen.1001134

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41(2):178–186. doi:10.1038/ng.298

Jamieson RV, Tam PP, Gardiner-Garden M (1996) X-chromosome activity: impact of imprinting and chromatin structure. Int J Dev Biol 40(6):1065–1080

Jones PA (1999) The DNA methylation paradox. Trends Genet 15(1):34–37

Kaslow DC, Migeon BR (1987) DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. Proc Natl Acad Sci USA 84:6210–6214

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462(7271):315–322. doi:10.1038/nature08514

Loebel DA, Johnston PG (1996) Methylation analysis of a marsupial X-linked CpG island by bisulfite genomic sequencing. Genome Res 6(2):114–123

Lopes AM, Ross N, Close J, Dagnall A, Amorim A, Crow TJ (2006) Inactivation status of PCDH11X: sexual dimorphisms in gene expression levels in brain. Hum Genet 119(3):267–275. doi:10.1007/s00439-006-0134-0

Nguyen DK, Disteche CM (2006) Dosage compensation of the active X chromosome in mammals. Nat Genet 38(1):47–53

Papageorgiou EA, Fiegler H, Rakyan V, Beck S, Hulten M, Lamnissou K, Carter NP, Patsalis PC (2009) Sites of differential DNA methylation between placenta and peripheral blood: molecular markers for noninvasive prenatal diagnosis of aneuploidies. Am J Pathol 174(5):1609–1618

Prantera G, Ferraro M (1990) Analysis of methylation and distribution of CpG sequences on human active and inactive X chromosomes by in situ nick translation. Chromosoma 99(1):18–23

Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Gräf S, Tomazou EM, Bäckdahl L, Johnson N, Herberth M, Howe KL, Jackson DK, Miretti MM, Fiegler H, Marioni JC, Birney E, Hubbard TJ, Carter NP, Tavaré S, Beck S (2008) An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). Genome Res 18(9):1518–1529

Ramakers C, Ruijter JM, Deprez RH, Moorman AF (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339(1):62–66

Rens W, Wallduck MS, Lovell FL, Ferguson-Smith MA, Ferguson-Smith AC (2010) Epigenetic modifications on X chromosomes in marsupial and monotreme mammals and implications for evolution of dosage compensation. Proc Natl Acad Sci USA 107 (41):17657–17662. doi:10.1073/pnas.0910322107

Riggs AD (1975) X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet 14:9–25

Ropers HH (2006) X-linked mental retardation: many genes for a complex disorder. Curr Opin Genet Dev 16(3):260–269. doi:10.1016/j.gde.2006.04.017

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res 37 (6):e45. doi:10.1093/nar/gkp045

Saxonov S, Berg P, Brutlag D (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci USA 103(5):1412–1417

Schilling E, Rehli M (2007) Global, comparative analysis of tissue-specific promoter CpG methylation. Genomics 90(3):314–323. doi:10.1016/j.ygeno.2007.04.011

Smit AFA, Hubley R, Green P (1996–2010) RepeatMasker Open-3.0. http://www.repeatmasker.org.

Stabellini R, de Mello JC, Hernandes LM, Pereira LV (2009) MAOA and GYG2 are submitted to X chromosome inactivation in human fibroblasts. Epigenet Off J DNA Methylation Soc 4(6):388–393

Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, Patapoutian A, Hampton GM, Schultz PG, Hogenesch JB (2002) Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci USA 99 (7):4465–4470. doi:10.1073/pnas.012025199

Takai D, Jones PA (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci USA 99(6):3740–3745

Team RDC (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Turner G, Turner B (1974) X-linked mental retardation. J Med Genet 11(2):109–113

Venolia L, Gartler SM, Wassman ER, Yen P, Mohandas T, Shapiro LJ (1982) Transformation with DNA from 5-azacytidine-reactivated X chromosomes. Proc Natl Acad Sci USA 79:2352–2354

Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7(8):645–653. doi:10.1038/nrg1914

Viegas-Pequignot E, Dutrillaux B, Thomas G (1988) Inactive X chromosome has the highest concentration of unmethylated Hha I sites. Proc Natl Acad Sci USA 85:7657–7660

Vucic EA, Wilson IM, Campbell JM, Lam WL (2009) Methylation analysis by DNA immunoprecipitation (MeDIP). Methods Mol Biol 556:141–153. doi:10.1007/978-1-60327-192-9_10

Wang Y, Leung FC (2004) An evaluation of new criteria for CpG islands in the human genome as gene markers. Bioinformatics 20(7):1170–1177. doi:10.1093/bioinformatics/bth059

Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37(8):853–862

Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 39(4):457–466

Wolf SF, Dintzis S, Toniolo D, Persico G, Lunnen KD, Axelman J, Migeon BR (1984a) Complete concordance between glucose-6-phosphate dehydrogenase activity and hypomethylation of 3′ CpG clusters: implications for X chromosome dosage compensation. Nucl Acids Res 12(24):9333–9348

Wolf SF, Jolly DJ, Lunnen KD, Friedmann T, Migeon BR (1984b) Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. Proc Natl Acad Sci USA 81:2806–2810

Xiong YY, Chen XS, Chen ZD, Wang XZ, Shi SH, Wang XQ, Zhang JZ, He XL (2010) RNA sequencing shows no dosage compensation of the active X-chromosome. Nat Genet 42 (12):1043–U1029. doi:10.1038/Ng.711

Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ (2006) Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res 16(8):995–1004. doi:10.1101/gr.5217506

Yasukochi Y, Maruyama O, Mahajan MC, Padden C, Euskirchen GM, Schulz V, Hirakawa H, Kuhara S, Pan XH, Newburger PE, Snyder M, Weissman SM (2010) X chromosome-wide analyses of genomic DNA methylation states and gene expression in male and female neutrophils. Proc Natl Acad Sci USA 107 (8):3704–3709. doi:10.1073/pnas.0914812107

Yorifuji T, Muroi J, Uematsu A, Tanaka K, Kiwaki K, Endo F, Matsuda I, Nagasaka H, Furusho K (1998) X-inactivation pattern in the liver of a manifesting female with ornithine transcarbamylase (OTC) deficiency. Clin Genet 54(4):349–353